



KU LEUVEN
FACULTY OF ECONOMICS AND BUSINESS

DISSERTATION PRESENTED TO OBTAIN
THE DEGREE OF DOCTOR IN ECONOMICS

Information asymmetries,
agricultural technology adoption,
and public service delivery:
Experimental evidence from Uganda

Author

Caroline MIEHE

Supervisor

Prof. Dr. Johan SWINNEN

Daar de proefschriften in de reeks van de Faculteit Economie en Bedrijfs-wetenschappen het persoonlijk werk zijn van hun auteurs, zijn alleen deze laatste daarvoor verantwoordelijk.

The views expressed in dissertations published by the Faculty of Economics and Business are those of their authors, and do not necessarily represent those of the University of Leuven.

Committee

Supervisor	Prof. Dr. Johan Swinnen	KU Leuven, IFPRI
Members	Dr. Bjorn Van Campenhout	KU Leuven, IFPRI
	Prof. Dr. Eva-Marie Meemken	ETH Zürich
	Prof. Dr. Francesco Cecchi	Wageningen University
Chair	Prof. Dr. Patrick Van Cayseele	KU Leuven

Acknowledgments

To Jo, for introducing me to the world of development economics – a world that you revealed to me, where I discovered my true passion and purpose. Your decision to provide me with the opportunity to pursue a PhD at LICOS forever changed the course of my life. I will always be grateful for that, and for countless evenings at Café Caroline, Casa Del Popolo, and Gambrinus, filled with captivating stories and hearty laughter.

To Bjorn, for your constant support and genuine care throughout this journey. You have generously shared your knowledge and expertise with me, and have taught me the tools of our trade, ranging from complex experimental design, over your commitment to the highest ethical and scientific standards, to the challenges of taking our experiments to the field. More importantly, you have been patient and understanding, and truly appreciated me for who I am, providing me with the confidence to grow as a researcher and as a person. You also introduced me to Uganda and the African continent. Witnessing its beauty and its challenges gave me a sense of purpose and dedication, and sparked a curiosity within me that has become a driving force behind my work. Thank you for your genuine interest in my development and success, there is no doubt that your mentorship has shaped my academic and personal growth, and I will never forget the role you have played in my life.

To David Spielman and Robert Sparrow, my favorite co-authors, for letting me join the Seed Systems project that shifted my research interest towards agricultural technology adoption, and for making me feel like an appreciated part of the team.

To Eva-Marie Meemken and Francesco Cecchi, for the time you dedicated to reading and evaluating my work. Your thoughtful comments improved my dissertation, and I am grateful to have benefited from your expertise.

To my colleagues at LICOS: Adriana, for always being up to a friendly and interesting chat; Andreas, for countless conversations about random and serious topics – your presence next door has become a comforting certainty, and I will miss you; Anusha, we have gone through the most memorable moments of the PhD life together – assignments in econometrics, our joint paper, conferences in Göttingen, Oxford, and Kigali – I am grateful to have shared these experiences with you and for your support along the way; Dorien, for being my role model in terms of positivism and academic excellence – you are a fascinating person, genuinely interested in the most diverse set of subjects, activities, and people, and I am happy that you will keep on inspiring others in your new role as a professor; Elfriede, thank you, not only for your administrative and organizational superpowers, but for the reliability and compassion you brought to LICOS; Fernando, for speaking your mind; Gero, you joining LICOS has been a gift – you are a genuinely good person with a strong sense of community, always offering to help – I will miss the positive energy you bring to all our gatherings, and our chats standing in the doorway, just checking in and enlightening the mood; Giulia, you are one of the warmest persons I know and your laughter is infectious – having you around feels like a hug, but not only that, you would stand up for us if necessary, and I will never forget your support; Hannah, whenever I felt overwhelmed, I would knock on your door – thank you for the mental support, always welcoming my emotions, never judging them; Hyejin, I simply enjoy being around you and feel like we developed quite a bond throughout the years; Julie, for your dedication to the cause; Matthieu, the newest member of our LICOS family, even though it feels like you have been here for years; Theresa, I hope you find a way to see what I see; and my former LICOS colleagues: both Joachims, Liz, Nele, Nik, Rob, Saule, and Senne, for welcoming me with open arms and guiding me through my first months as a PhD student.

To the Ugandan team, for your hard work and for going out of your way to make me feel comfortable in the field. Especially to Charles, for many hours of honest conversations at Hotel Continental in Iganga, expanding my perspective and deepening my understanding.

To 16,855 study participants, for contributing their time and energy to the interviews that enabled this research.

To Laurenz, my brother and best friend, for our inseparable bond. We are connected at heart, two halves of one whole. Having you by my side during childhood, youth, and young adulthood has been the strongest support, and made every challenge I faced, including this PhD, so much more feasible. Thank you for never leaving my side, for knowing me better than I know myself, for your genuine interest in my work and always asking the right questions, and for believing in me since we were children. I am blessed to have you in my life and so proud of the person you have become.

To my parents, for raising me to become independent and fearless, and for providing me with space and time to explore. I remember the first years of my life as if they were stories in an Astrid Lindgren book – Laurenz and I were exploring mysterious forests, climbing tall trees, building secret hideouts, swimming in cold lakes, and sitting by crackling campfires under the stars – without a care in the world. Thank you for letting me experience this freedom, for allowing me to take risks, fall, and learn. This childhood provided me with the confidence and strength to pursue whichever path I would choose in life. Thank you for trusting my judgment and believing in my capabilities, from those days when I was a child playing in the forest until today. I remember our large garden and our uniquely shaped house, and that the two of you would sit at our kitchen table and drink tea. It was a warm place, and I know that growing up in a cozy home will forever influence my life. Many things have changed since then but I know that I can count on your support in every scenario. When I considered moving to Belgium, my mother was the one who selflessly told me to trust my feelings, even if that would mean that I would leave Berlin for good. This is one of many great pieces of advice she gave me over the years and until today, I would not take an important decision without consulting her first, knowing that she is the best judge of any situation.

To my grandparents Elke and Klaus, for being perfect role models in living a happy life, and for being supportive, yet never demanding. To my grandmother Elisabeth, a strong woman ahead of her time and another figure to look up to.

To Fré, for filling my life with love and joy. You are truly in peace with who you are and the most positive person I know, turning our journey into a constant flow of happy highlights. We met in a park in Barcelona on a warm summer evening in 2014 and shared our lives ever since. Our journeys are so intertwined that it is impossible to say whom we would have become if our paths would not have crossed. Meanwhile, we met every version of each other, built a home, and an amazing circle of friends. I very much enjoy spending time with you, chatting about our days while having dinner on calm weekday evenings, drinking countless cups of tea and discussing our view of life on slow weekend mornings, dancing until the sun comes up, hiking in the most beautiful scenery, traveling and exploring this world together. Thank you for all these happy moments that contrasted the challenging ones during my PhD journey, for your mental and practical support over the past years, for setting boundaries when I would not find them myself, for reminding me to not lose sight of what really matters, for your openness to follow me wherever life will take us, and for your unconditional love.

To my family in law, especially Diane, for treating me as one of their own since day one.

To all my friends in Belgium, Berlin, and around the globe. To Bette, Emma, Yaël, and Wendy, my source of happiness during light days and my shelter during dark ones. Thank you for loving me for who I am, for the honest conversations that we shared, and for the most fun nights of my life.

To the teachers who believed in me and inspired me, especially the first and most relevant one, Frau Potyka.

Contents

Introduction	3
The impact of community-based monitoring on public service delivery: A randomized control trial in Uganda	5
1.1 Introduction	5
1.2 Background of the baraza program	9
1.2.1 Nature of the baraza institution	9
1.2.2 Ugandan baraza program	10
1.3 Research objectives	12
1.3.1 Impact of (sub-county-level) barazas	12
1.3.2 Information mechanism	12
1.3.3 Deliberation mechanism	13
1.3.4 Jurisdictional tier	14
1.4 Experimental design	15
1.5 Treatments	17
1.6 Implementation challenges	18
1.6.1 Balance between assigned-to-be-treated-but-not-treated sub-counties and control sub-counties	20
1.6.2 Selection of control sub-counties for end-line survey . . .	22
1.7 Results	25
1.8 Discussion	28
1.8.1 Problems with the theory of change	28
1.8.2 Political context	32
1.8.3 Methodological issues	33
1.9 Conclusion	34
1.10 Acknowledgments	36

Gender bias in customer perceptions: The case of agro-input dealers in Uganda	37
2.1 Introduction	37
2.2 Research question and relation to the literature	40
2.3 Context and data	42
2.3.1 Study context	42
2.3.2 Sample	42
2.3.3 Descriptive statistics	43
2.3.4 Measuring perceptions	44
2.4 Empirical strategy	47
2.5 Results	52
2.5.1 Between-dealers models	52
2.5.2 Farmer fixed-effects models	60
2.6 Conclusion and policy implications	63
2.7 Acknowledgments	70
2.8 Appendix	71

Miracle seeds: Biased expectations, complementary input use, and the dynamics of smallholder technology adoption	73
3.1 Introduction	73
3.2 Related literature	78
3.3 Theoretical framework	80
3.4 Intervention	86
3.5 Data and empirical strategy	88
3.5.1 Sample	88
3.5.2 Adoption	89
3.5.3 Empirical strategy	91
3.6 Results	91
3.6.1 Impact on knowledge	91
3.6.2 Impact on adoption	94
3.6.3 Impact on expectations and harvest	96
3.6.4 Impact on use of inputs and practices	97
3.6.5 Heterogeneity analysis: remote households	99
3.7 Conclusion	101
3.8 Acknowledgments	107
3.9 Appendix	108

The (perceived) quality of agricultural technology and its adoption: Experimental evidence from Uganda	109
4.1 Introduction	109
4.2 Experimental design	117
4.3 Interventions	118
4.3.1 Agro-input dealer training	118
4.3.2 Information clearinghouse	119
4.4 Empirical strategy	123
4.5 Data	125
4.5.1 Sample	125
4.5.2 Descriptive statistics	125
4.5.3 Orthogonality tests of randomization balance	126
4.6 Results	127
4.6.1 Impact on agro-input dealers	130
4.6.2 Impact on smallholder farmers	136
4.7 Causal chain and mechanisms	140
4.7.1 Dealer knowledge	140
4.7.2 Dealer efforts, services, and practices	144
4.7.3 Farmer switching	146
4.7.4 Farmer perceptions	150
4.8 Attrition	153
4.9 Conclusion	154
4.10 Acknowledgments	158
4.11 Appendix	159
4.11.1 Can farmers assess maize seed quality?	159
4.11.2 Rating computation: details	162
4.11.3 Rating dissemination: details	164
4.11.4 Outcome variables and results as they were pre-registered	165
4.11.5 Multiple choice questions to measure dealer knowledge .	170
Bibliography	173

Introduction

Markets with asymmetric information caught the gaze of economists in the 1970s, when George Akerlof, Michael Spence and Joseph Stiglitz established the foundations for this line of research. Their contribution was rewarded with the *Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel* in 2001. We now acknowledge that many markets are characterized by asymmetric information, meaning that one actor in the market has more or better information than the other, which explains several economic and social phenomena that would otherwise be difficult to understand (Stiglitz, 2000). Today, models with imperfect information are indispensable instruments in an economist's toolbox, but information asymmetries are not only academic abstractions: they have highly concrete implications for our understanding of developing economies. When Nobel Prize winner Joseph Stiglitz himself visited the developing world for the first time in 1967, he was particularly struck by the imperfections of information he encountered (Stiglitz, 2002).

And indeed, the poor often lack critical pieces of information. They do not know the benefits of immunizing their children, the returns to education, how much fertilizer to use, how to avoid Human Immunodeficiency Virus (HIV) infection, or what their politicians do (Duflo and Banerjee, 2011). Information affects decision making and consequently, people in poverty often end up making sub-optimal decisions, sometimes with dramatic consequences. Examples of sectors that are troubled by asymmetric information problems and affect the poor include health, veterinary medicine, education, agricultural credit, and civil services (Leonard et al., 2013). Also George Akerlof pointed to the prevalence and importance of information asymmetries in developing economies: one of his illustrative examples of adverse selection was the Indian credit market in the 1960s (Akerlof, 1970).

Given the pervasive lack of information in developing countries, providing the economically disadvantaged with this information can effectively improve their decision-making and behavior. This dissertation provides evidence from four interventions that attempt to reduce information asymmetries in public service and agricultural input markets in Uganda:

As the delivery of public services remains an important problem in many developing countries, Chapter 1 deals with asymmetric information between policy makers, civil servants, and citizens. Through community-based monitoring, information asymmetries can be reduced, and beneficiaries of public services can apply bottom-up pressure to under-performing service providers and their political leadership. In this spirit, the Government of Uganda organizes community fora—popularly known as *barazas*—where citizens receive information from government officials and get the opportunity to challenge them. We designed a cluster randomized control trial to assess the impact of this large-scale, high-profile policy intervention on public service delivery in agriculture, health, education, and infrastructure. Following a pre-specified confirmatory analysis, we do not find that the intervention has significant effects on general public service delivery, even though public services in the agricultural sector do improve. We discuss some of the potential reasons for this finding, including assumptions underlying the impact pathways, the political context in which the program was implemented, and challenges related to the evaluation of large policy programs.

Chapter 2 explores information asymmetries in agricultural input markets. Faced with incomplete and imperfect information, economic actors rely predominantly on perceptions and often base decisions on heuristics prone to bias. Gender bias in perceptions favoring men has been found in a wide variety of settings and may be an important reason why some sectors remain dominated by men and gender gaps in terms of benefits persist. Using ratings of agro-input dealers provided by smallholder farmers in their vicinity, we test if farmers perceive male-managed shops differently than shops managed by women. After explicitly controlling for quality differences between male- and female-managed agro-input shops and including fixed-effects to account for farmer-level heterogeneity, we find that farmers rate male-managed agro-input outlets higher on a range of attributes related to the dealership in general, as well as on the quality of inputs sold by the dealer. Our results show that gender bias in customer perceptions persists and continues to be a severe comparative disadvantage and

an important entry barrier for female agro-input dealers. The bias affects social outcomes like women’s capabilities, aspirations, and empowerment in seed systems but also impairs development at more aggregate levels: as a considerable share of agro-input shops is managed by women, this finding may impose challenges for varietal turnover, hindering agricultural productivity, food security, and rural transformation. Policies and interventions designed to challenge gender norms and customs are needed to correct this bias.

Chapter 3 points to another set of information asymmetries in agriculture. To fully benefit from new agricultural technologies like improved seed varieties, significant investments in complementary inputs such as fertilizers and pesticides, and practices such as systematic planting, irrigation, and weeding are required. Farmers may fail to recognize the importance of these complements, leading to unsatisfactory crop yields and outputs and, eventually, to dis-adoption of the variety. We provide a simple model of biased expectations, complementary input use, and technology adoption and test its predictions using a field experiment among smallholder maize farmers in eastern Uganda. We find that pointing out the importance of complementary investments using an information intervention in the form of a short, engaging video effectively deters some farmers from using commercial improved varieties. Consistent with the theoretical model, we find some evidence that this behavior change is driven by increased knowledge and expectations that are more in line with realized outcomes.

Chapter 4 examines the imperfections of information regarding the quality of agricultural technologies. It presents findings from a randomized control trial that tests two hypotheses on how the (perceived) quality of agricultural inputs affects adoption among smallholder farmers in Uganda. First, poor quality could be caused by agro-input dealers’ lack of knowledge about proper handling and storage. An information intervention in the form of a training is expected to improve input quality and subsequent adoption. Second, information asymmetries could crowd out the market for quality inputs—a classic lemons problem. Here, we implement an information clearinghouse based on crowd-sourced reviews similar to yelp.com. We find that agro-input dealers and farmers benefit from the clearinghouse, but not from the training.

The impact of community-based monitoring on public service delivery: A randomized control trial in Uganda

This chapter is co-authored with Tewodaj Mogues (International Monetary Fund, United States), Bjorn Van Campenhout (IFPRI and KU Leuven, Belgium), and Nassul Kabunga (Uganda Bureau of Statistics, Uganda). Earlier versions of this chapter are published as IFPRI Discussion Paper 1933 and 3ie Impact Evaluation Report 136.

1.1 Introduction

In many developing countries, the delivery of public services remains problematic. Public infrastructure, such as roads or boreholes for drinking water, is poor. The quality of service provided in hospitals and schools is low. Absenteeism and corruption are endemic. This is also the case in Uganda, where the public service sector suffers from high levels of elite capture, ineffective monitoring, and weak accountability (Reinikka and Svensson, 2004). In response to this, the Office of the Prime Minister (OPM) implemented the baraza program,

an initiative of the president of Uganda, to improve public service delivery by enhancing public involvement in holding the government accountable for service delivery in relation to resources spent (OPM, 2013). The program brings politicians, civil servants, and citizens together in town hall-type meetings (barazas) to share information and engage with each other.

Since 2009, barazas have been organized in sub-counties throughout Uganda. However, no rigorous impact evaluation was undertaken until now. A first objective of this paper is thus to assess if the program had an impact on public service delivery and associated outcomes. Furthermore, in 2012, the OPM decided to organize subsequent barazas at a higher administrative level (district) instead of at the sub-county level, mainly to reduce cost and speed up roll-out. A second major objective of this study is to inform policy makers on the consequences of this change in terms of effectiveness of the program. A third objective of the study is to assess the relative importance of the two main components of a generic baraza event and differentiate between, on the one hand, the impact of arming citizens with information related to budgeting and spending, and, on the other hand, the impact of allowing citizens to engage with public servants and politicians in moderated question-and-answer sessions. We answer these research questions using a cluster randomized control trial (RCT).

There are several studies that look at the impact of community-based monitoring on public service delivery, many of them implemented in Uganda. A landmark study is Björkman and Svensson (2009), who analyze the impact of a community-driven local accountability project in primary health care provision in Uganda. They find that community-based monitoring resulted in significant improvements in health care delivery, utilization, and health outcomes (most notably child mortality and weight-for-age z-scores) after one year, and confirm in Björkman Nyqvist, de Walque, and Svensson (2017) that these effects also hold in the longer run. More recently, however, Raffler, Posner, and Parker-son (2020) come to different conclusions when testing an intervention closely modeled on the one of Björkman and Svensson (2009) in a similar setting. The study, involving a three-wave panel of more than 14,000 households, shows that the intervention was able to change the behavior of public service providers. However, it did not find that it directly increased community monitoring, nor generated improvements in health outcomes, at least in the short run.

Our study contributes to this literature in various ways. First, it is one of the few studies that consider the role of the jurisdictional tier on the effectiveness of

community-based monitoring. The level at which the town hall-style meetings are organized may affect their effectiveness in opposing ways: interventions at a more local level may result in more relevant issues being scrutinized. However, qualitative explorations suggest that, often, issues raised in lower level barazas are beyond the operational jurisdiction of the officials that are present (Van Campenhout et al., 2018). This may be less of a problem if barazas are organized at district-level. Most other studies consider interventions that were placed at local jurisdictional levels. For instance, the intervention in Raffler, Posner, and Parkerson (2020) was implemented in health centers and their associated catchment areas, defined as the three villages closest to the health center.

Second, our study evaluates the impact of a large-scale, high-profile policy intervention that receives broad support within the government, among opinion leaders and among citizens in Uganda. Such interventions may trigger an entirely different set of dynamics than interventions that are organized by local or international Non-governmental organizations (NGOs), as in Björkman and Svensson (2009) and Raffler, Posner, and Parkerson (2020). For instance, many of the actors involved may find that NGOs are not mandated when it comes to public services such as health and education. This is also consistent with suggestive evidence in Raffler, Posner, and Parkerson (2020), who find that the presence of sub-county officials during their community-based monitoring intervention boosted its impact.

Third, barazas take a comprehensive, multi-sector approach, enabling cross-sectoral planning and potentially allowing for re-allocation of resources across sectors to solve problems that were identified during these meetings. Some of the problems most mentioned by users, such as hygiene in health centers or accessibility, involve cooperation between heads of different sectors (e.g., health and infrastructure to get access to water in health centers). Bringing civil servants from different sectors together and confronting them with the priorities of citizens may increase information sharing and cooperation between them. Most other studies focus on a single sector; health in particular seems to be a popular sector for community monitoring interventions (for example, Björkman and Svensson, 2009; Raffler, Posner, and Parkerson, 2020; Arkedis et al., 2021).

This paper also serves as an illustration of the challenges of evaluating real-world policy interventions and how to deal with them. While we feel that field

experiments with policy interventions provide greater external validity than the carefully designed and closely monitored ones implemented by NGOs that are commonly published, real-world policy experimentation is challenging: researchers need to overcome political resistance to randomization and commit to an hands-off approach, even if things do not go as planned (de Souza Leão and Eyal, 2019). Our study is no exception, as the OPM faced various challenges that affected the roll-out of the intervention, including budgetary constraints and disruptions related to the general elections of 2016. Four years after the baseline survey, with about 50% of the planned interventions completed, we needed to trade-off waiting for the remaining barazas to be implemented and conducting the end-line after partial roll-out. We decided to proceed with end-line data collection and employ various strategies to diagnose and mitigate problems that may have been introduced by the partial roll-out.

We find that barazas do not impact public service delivery as measured by a pre-registered summary index of various outcomes in four key sectors. If we zoom into these sectors, sub-county-level barazas seem to be effective in agriculture, while the other sectors remain unaffected by the intervention. This is in line with recent overview articles, meta-analyses, and replications of influential studies on the benefits of community-based monitoring such as the one by Raffler, Posner, and Parkerson (2020). The emerging consensus is that community-based monitoring can be effective, but only if certain conditions are met, and that effects may differ for different groups and outcomes. For instance, the perceptions of citizens related to the intentions and legitimacy of the organizing entity may affect impact (Hickey and King, 2016). Furthermore, individuals must have both the power and incentives to act on information they receive during interactions with officials (Kosec and Wantchekon, 2020). Effects on some outcomes (such as citizen engagement) may be stronger than effects on other outcomes (such as child mortality), depending on where in the causal chain impact is measured and on the scope for further improvement in the outcome (Raffler, Posner, and Parkerson, 2020).

Building on this literature, we look at potential reasons why barazas are not more effective in our specific case. We start by exploring the theory of change behind the baraza program, such as the alignment of priorities between citizens and officials, and the subsequent translation of priorities into strategies and actions. We then discuss how the political context in which community-based monitoring is embedded can affect effectiveness and point to various red

flags in the Ugandan case. Finally, we point out methodological challenges inherent to the evaluation of interventions where interactions between citizens and government officials lead to treatment heterogeneity (as opposed to the theoretical ideal where experimental units are assumed to be passive recipients of a uniform and homogeneous treatment) and argue that this heterogeneity is likely to dilute average treatment effects.

1.2 Background of the baraza program

1.2.1 Nature of the baraza institution

Baraza is a Swahili word that refers to a set of places where people meet (e.g., a veranda in front of a house or some benches under a simple shelter near the local mosque), and to clubs, unions, and associations, but the word also describes all kinds of meetings and gatherings of people, from simple informal assemblies to formal councils (Loimeier, 2005). Throughout East Africa, barazas are traditional forms of community assemblies (Naanyu et al., 2011). In Zanzibar, they are essential features of the social discourse and have facilitated discussions ranging from local deliberations and disputes to international matters. Here, barazas facilitate negotiations of all kinds and serve to share information and knowledge. They can be public or private, formal or informal, but often follow a distinct set of rules and a particular code of behavior (Loimeier, 2005). In Burundi, barazas have traditionally been used to settle differences and discuss community issues. In Tanzania, barazas include formal councils and corporations like the National Swahili Council “Baraza la Kiswahili la Taifa” or the Tanzania News Broadcast Corporation “Baraza la Habari Tanzania” (Naanyu et al., 2011). In Kenya, barazas refer to assemblies where politicians and bureaucrats engage with citizens on a public stage. These meetings are usually held outdoors, individually licensed by the state, and range in size from smaller gatherings to large assemblies with several thousand participants who are addressed by the president (Haugerud, 1995). Kenyan barazas used to be the main meeting ground between “the ruled” ordinary citizens on the one hand, and state officials and bureaucrats as “the rulers” on the other hand. While some barazas only serve to legitimize a political ideology or aim at pro forma popular consensus, others lead to honest and real debates between citizens and politicians (Haugerud, 1995). Until today, all Kenyan chiefs are required by

law to convene at least two barazas per month.

The above shows that barazas have a long history in the region. The attempt of the Government of Uganda to leverage such an important institution to increase its own effectiveness is interesting and may lead to quite different outcomes than other community monitoring interventions. At the same time, community fora organized by the government may crowd out or even replace alternative community fora that are potentially more inclusive and more critical of those in power.¹ In the limiting case, government-run barazas may completely bypass the initial aim of increasing accountability and simply become instruments of propaganda.

1.2.2 Ugandan baraza program

In the mid-1980s, after attaining relative stability, the Government of Uganda, supported by development partners, initiated far reaching liberalization efforts and introduced a decentralized system of governance, in which district and sub-county administrations received considerable responsibilities (Francis and James, 2003). This process was assumed to bring representative governance closer to the people and to align public services to citizen needs. However, two decades after these ambitious reforms, public services in Uganda continued to be characterized by misaligned priorities, substandard services, elite capture, ineffective monitoring, and weak accountability (Reinikka and Svensson, 2004).

In response to this, the Government of Uganda, under the stewardship of the OPM, initiated community-based monitoring and accountability fora (or barazas) with the general objective of “enhancing public involvement in holding the government accountable for service delivery in relation to the resources spent” (OPM, 2013). Barazas were conceived as platforms that enhance information sharing between policy makers (the client), public service providers (the implementer), and beneficiaries of public goods and services (the users). In addition, barazas were designed to provide users with an opportunity to ask questions to the client and the implementer and to deliberate among themselves.

¹For instance, in Uganda, there used to be meetings called *ebimeeza* which allowed people to discuss and raise issues publicly or quasi-publicly. These gatherings have been taking place for a considerable period, varying in terms of the degrees of institutionalization. The Government of Uganda prohibited these *ebimeeza* meetings around the same time when the baraza program was implemented. We are grateful to an anonymous referee for pointing this out.

Barazas were piloted in the financial year 2009/2010. Since then, efforts have been underway to roll out barazas in all sub-counties in the country. By the last quarter of the 2011/2012 financial year, 267 of the country's total 1,340 sub-counties, spread over 112 districts, had held a baraza meeting. At the beginning of 2012/2013, however, changes in implementation were suggested: subsequent barazas would target districts instead of sub-counties to accelerate roll-out, increase participation at a higher jurisdictional level, and reduce costs.

A typical baraza meeting is initiated from the center, with the OPM mobilizing district and sub-county officials. These include the Chief Administrative Officer (CAO) as the head of public service providers at the district level, the Resident District Commissioner (RDC) as a direct representative of the president, the District Local Council chairperson (LC5) as the representative of political leadership at the district level, and the sector heads responsible for public services in each sector (agriculture, education, infrastructure, and health). For sub-county-level barazas, the equivalent of the CAO at the sub-county level (the sub-county chief) and the Sub-county Local Council chairperson (LC3) also have important roles. The OPM, in consultation with the district leaders (RDC, CAO, and LC5) and other stakeholders, agrees on a date and a neutral location. Again, in consultation with the district leaders, viable moderators are identified to guide the baraza forum. Village mobilizers and community resource persons publicize the event. These community mobilization efforts are further reinforced by adverts in the local media in the form of radio announcements, printed banners, posters and fliers, and mobile public address systems, a few days before the baraza event.

A baraza meeting is chaired by the RDC's office. In front of the audience, including local citizens, invited opinion leaders, elders, and journalists, the moderators seek feedback from the head of each major sector (agriculture, education, health, and infrastructure). Sector heads present which services were planned to be delivered in the sub-county (or district for district-level barazas), which services were actually delivered, which issues and challenges have emerged, and what is the way forward. The moderators then seek feedback from citizens on whether what has been presented is what they understand was planned and actually implemented. Sector heads are then given another opportunity to clarify, or react to, any issues raised by the citizens. At the end of the process, a report to the OPM is prepared, indicating issues that arose in the baraza meeting. This report points out policy and program implementation

weaknesses and challenges, which is expected to further feed into the general government performance management system. Usually, a minister of state is also present at the baraza. Baraza events often become emotional and attract considerable media attention, also from national newspapers.

1.3 Research objectives

1.3.1 Impact of (sub-county-level) barazas

The baraza intervention is a community-based monitoring intervention that combines the provision of information with the opportunity for citizens to engage with each other and with decision makers at a local level. Community-based monitoring has become a popular tool to improve public service delivery. However, not all interventions seem to achieve success (Olken, 2007; Waddington et al., 2019). A key question is therefore related to the effectiveness of a typical baraza intervention as originally conceived and organized by the OPM. We thus first test the impact of sub-county-level barazas that combine information provision with a deliberation component on public service delivery and associated outcomes.

It is reasonable to expect meaningful interactions between the information and deliberation components of a baraza. For example, in a generic baraza, the information component is primarily designed to inform citizens about the activities of the service providers. To some extent, citizens are passive recipients of this information, and officials report what they consider relevant, or may even attempt to misrepresent the facts. If citizens can also engage with policy makers and civil servants, they may request information that is relevant to them. It may also result in information flows in the opposite direction as government officials learn about the priorities and concerns of citizens.

1.3.2 Information mechanism

The relationship between citizens and government officials is reminiscent of the principal-agent problem. In essence, there are three players (elected politicians, civil servants, and citizens) with only partly overlapping information sets and

potentially competing interests.² Bringing these stakeholders together in town hall-type meetings is assumed to reduce information asymmetries and may be an effective way to improve the quality of public services by (1) allowing citizens to monitor and apply bottom-up pressure on under-performing civil servants and by (2) revealing the discrepancy between what was promised and what was actually delivered to politicians, increasing top-down pressure on under-performing civil servants. A second aim of this study is therefore to isolate the effect of the information provision component within the broader baraza intervention.

There is some evidence that channeling information about the quantity, modality, and quality of public services, as well as about the investments and policy decisions by politicians, bureaucrats, and service providers to citizens, can increase their ability to hold leaders accountable and thus improve service provision. For example, Pandey, Goyal, and Sundararaman (2009) establish that a community-based information campaign has a positive impact on school performance, using from a cluster RCT in India. Grossman and Michelitch (2018) disseminate information about the job performance of randomly selected Ugandan politicians. While this increases job performance on a range of criteria, they find no impact on public service provision. Banerjee et al. (2018) find that targeted information reduces elite capture in Indonesia: mailing cards with program information to targeted beneficiaries increased the subsidy they received from a subsidized rice program. A recent review of 48 empirical studies on the impact of information on governance and service delivery suggests that the availability of information alone may not be sufficient. Information must be deemed relevant to its recipient, and individuals must have both power and incentives to act on the information (Kosec and Wantchekon, 2020).

1.3.3 Deliberation mechanism

Deliberation may improve the quality of public services in different ways. First, it has a legitimizing effect on decisions arrived at in this fashion. Effective deliberation assumes equal voices of marginal and advantaged agents, and the role of evidence that supports the positions articulated (Mustalahti and Rakotonarivo,

²As the public servant must be responsive to the needs of both the client and the community at the same time, the problem can be characterized as a multiple or common agency problem, which adds a collective action component to the standard principal-agent problem (Bernheim and Whinston, 1986).

2014). Second, deliberation can distill social choice more effectively than simple voting and majoritarian rule, in part by building consensus both among citizens and between public servants and citizens (Rosenberg, 2007). Third, deliberation has been found to positively influence the vigor and breadth of subsequent citizen involvement in community affairs (Björkman Nyqvist, de Walque, and Svensson, 2017).

The effectiveness of deliberation has also been the subject of empirical analyses. For example, experimental evidence shows that deliberative processes can reduce clientelism (Fujiwara and Wantchekon, 2013). López-Moctezuma et al. (2020) show that town hall-style meetings like barazas during election campaigns increase voter awareness on the issues parties campaigned on. The third key question thus deals with the relative importance of the deliberative component of a typical baraza.

1.3.4 Jurisdictional tier

Baraza interventions can also be distinguished by the administrative level at which they are implemented. Barazas were originally implemented at the sub-county level, but from 2012 onwards, many were implemented at the district level.³ This points to a potential trade-off between achieving breadth of coverage (through district-level barazas), and depth of coverage (through sub-county-level barazas). While conducting one district-level baraza is likely to be cheaper than conducting sub-county-level barazas in all sub-counties in that district, it is not clear a priori how these cost savings justify potential reductions in effectiveness.

Whether placement at a higher or lower level is more effective is likely to depend on the outcome and the context. For instance, it has been argued that engaging small groups can be more effective because they can be coordinated more easily, but large groups may be preferable if the desired outcome would be enjoyed by a broader group (Donato and Mosqueira, 2019). Furthermore, action may be more likely if an issue is brought by a large group instead of a small group of people that complains about a highly localized issue (Banerjee, Deaton, and Duflo, 2004). It may also be that issues highlighted at the local level fall under the responsibility of higher-level authorities and vice versa. Therefore, a final objective of this study is to assess the impact of district-level

³An average district in Uganda consists of 15 to 20 sub-counties.

barazas and compare it to the effectiveness of barazas organized at the lower sub-county level.

1.4 Experimental design

To answer the above research questions, we designed a field experiment that covers districts, sub-counties, and households across the four regional blocks of Uganda (Northern, Western, Central, and Eastern). Each regional block has unique characteristics in terms of ethnicity, geographical and agro-ecological conditions, as well as cultural history. As noted in Section 1.2, a small share of sub-counties, albeit located throughout all of Uganda’s 112 districts across the four regions, had already received a baraza prior to the study.⁴ We selected our sample of districts from “eligible districts,” and our sample of sub-counties from “eligible sub-counties.” An “eligible district” is defined as a district in which no district-level baraza had been implemented prior to the start of the study. An “eligible sub-county” is defined as a sub-county to which two conditions applied: (i) no sub-county-level baraza had been conducted yet, and (ii) the sub-county is in an “eligible district”.⁵

This study employed a two-step randomization design, illustrated in Figure 1.1. In a first step, we randomly allocated eligible districts to treatment and control conditions. Some eligible districts were selected to receive district-level barazas (D^{ID}), while other districts would not receive a baraza at this level (D^0).⁶ In a second step, we randomly allocated each eligible sub-county to one of four conditions: about one quarter of all eligible sub-counties sampled from D^0 serves as a pure control and would not receive a baraza at any level (S_0^0); about one quarter would receive a sub-county-level baraza that combines both information and deliberation treatment (S_{ID}^0); about one quarter would receive a sub-county-level baraza that consists of officials providing information

⁴Since the beginning of our study, the number of districts has increased to about 140, primarily by splitting larger districts into smaller ones. We worked with the districts and sub-county boundaries as they were defined at the start of the study.

⁵Restricting ourselves to eligible districts and sub-counties has obvious consequences for the study population. For instance, the three districts that constitute Karamoja, one of the poorest and most remote areas located in the northeast of Uganda, had already received district-level barazas. As a result, this area was excluded from the analysis. While the impact evaluation had national coverage, the fact that we evaluate an ongoing policy project needs to be kept in mind when interpreting results and extrapolating conclusions.

⁶We do not differentiate between the information and deliberation components of barazas organized at the district level. As such, they are the district-level equivalents of sub-county-level barazas that combine both information and deliberation treatment.

but limited opportunity for citizens to engage (S_I^0); and about one quarter would receive a sub-county-level baraza with a focus on citizens engaging with each other and with officials but with limited information provision (S_D^0).⁷ We also took a random sample of sub-counties from the D^{ID} districts that would receive a district-level baraza (S_0^{ID}). Within each sub-county, we randomly sampled five villages, and in each village, we randomly selected ten households.

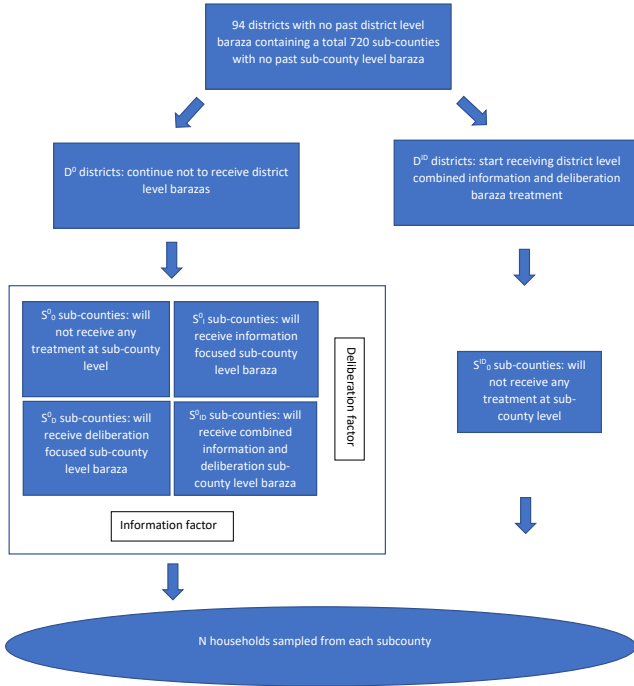


Figure 1.1: Experimental design

⁷The experiment at the sub-county level was set up as a 2^2 factorial design, so that pooling treatment cells would boost power. However, as we find that the interaction effects between the information and deliberation treatments are often important, we estimate fully saturated models as explained in the results section. As a result, the design at the sub-county level is equivalent to a parallel design with a control group, an information only group, a deliberation only group and an information+deliberation group.

To determine the number of districts, sub-counties, and households to include in the study, we ran an extensive series of power calculations using data from the Uganda National Household Survey of 2009/2010 and the Demographic and Health Survey of 2011 to estimate standard errors of and correlations between the outcome variables. In Online Appendix A.1, we show summary statistics for a range of characteristics of about 12,500 households that were included in the study, collected between June and September 2015. We also test for baseline balance between the different treatment groups in the original design. While we find significant differences for three out of 40 comparisons, we do not think that this indicates structural imbalance.

1.5 Treatments

The treatments are based on barazas as designed and implemented by the government before this study. A generic sub-county-level baraza was already described in Section 1.2. From such a generic baraza, we either removed the information component or the deliberation component to test their relative importance. To standardize the treatments, we developed detailed scripts that RDCs and facilitators were expected to follow, and invited them for training. We will summarize the main differences between an information baraza and a deliberation baraza in this section. The training manuals can be found here.

For information barazas, templates to gather information were filled by officials and posted in central locations in each parish of the sub-county (or district for district-level barazas) two weeks before the baraza event took place. The template was designed to inform citizens about planned and actual public expenditures in the previous fiscal year, about achievements and challenges during that year, and about planned expenditures and targets for the next fiscal year. This information was provided by the sub-county chief for each of the four sectors (agriculture, infrastructure, health, and education).⁸ On the day of the baraza event, the CAO introduced the sector heads and gave a brief presentation on the overall budget for the fiscal year, and the main achievements and challenges in service delivery. After a brief intervention by the OPM, officials responsible for each sector presented the information on

⁸While the preparation and distribution of these posters were the responsibility of the RDC, there were research assistants that closely monitored the implementation. We also assisted in printing the template posters.

the templates. An information-focused baraza allowed for only ten clarifying questions, to be collected and asked by the facilitator.

For deliberation barazas, posters were also mounted in each parish of the sub-county (or district for district-level barazas), but only announced that a baraza was going to be held at a particular date and place. At the baraza, after a brief introduction by the OPM, citizens were instructed to break into four groups by sector, discuss problems they faced and draw up a list of priority issues that needed to be addressed. Facilitators in each group were expected to focus the discussion on what was done well, on the problems citizens encountered during the past year, and on what needs to be done in the next fiscal year. They were also required to anonymously collect the resulting information. After the break-out sessions, officials reacted to the specific comments and requests.

A sub-county-level baraza in the crossed treatment group combined both elements. District-level barazas were similar to (information and deliberation) sub-county-level barazas, except for the fact that they were organized at the district headquarters, that the chiefs and LC3's of each sub-county were expected to attend in case questions related to their sub-county arose, and that all citizens in the district were invited.

1.6 Implementation challenges

Our implementing partner OPM faced substantial challenges in the roll-out of sub-county-level barazas.⁹ At the start of 2018, and almost two and a half years after baseline data was collected, only about 25% of the planned sub-county-level interventions had happened. At that point, costs and benefits of waiting until the OPM finished all barazas or collecting baseline information despite incomplete roll-out were assessed. After an additional six months, with still only 56 of 155 sub-county barazas done, we decided to collect end-line data before all sub-counties were treated, given that the government's financial constraints made it unlikely that all barazas would be held anytime soon.

End-line data collection after partial roll-out will affect statistical power. More worryingly, it may introduce selection bias in the sub-sample of sub-counties that were assigned to the treatment. It may be that, from the sub-

⁹All district-level barazas were implemented as planned, therefore this section focuses on sub-county-level barazas.

counties randomly assigned to treatment, particular sub-counties were treated first, and others postponed. For instance, the implementing partner may have started with sub-counties near the capital for logistical reasons, resulting in a treatment group that is not directly comparable to the control group. Or the OPM may have decided to prioritize sub-counties that perform the worst on certain indicators, such as the number of boreholes or student enrollment rates.

We acknowledge the partial roll-out as a threat to internal validity that should not be ignored. At the same time, the fact that we started from a cluster RCT still provides substantial advantages over studies that are based on observational data. For instance, potential selection emanating from partial roll-out is restricted to the sub-sample of sub-counties that was assigned to be treated, substantially reducing the scope for bias. In addition, the list of sub-counties to be treated that we shared with the implementing partner was organized by treatment group (information, deliberation, and combined treatment), and in each treatment group, sub-counties were listed alphabetically. Looking at this list, we get the impression that the OPM started at the top and worked its way down: relatively more sub-counties that were assigned to the information baraza were treated, and relatively more sub-counties at the top of each treatment group were treated. This pattern is confirmed when we regress the likelihood of a sub-county being treated on its rank in the list. This is a first indication that OPM officials did not deliberately select certain sub-counties.

In this section, we diagnose potential selection bias by assessing baseline balance between assigned-to-be-treated-but-not-treated sub-counties and control sub-counties. We then explain the end-line data collection strategy that further reduces potential selection bias: selecting control units by matching them to treatment units on a range of sub-county characteristics that may have affected how the intervention was rolled out. In Online Appendix A.2, we investigate the impact of the partial roll-out on statistical power. To do so, we use simulation methods to estimate minimum detectable effect sizes at different power levels for the reduced sample size associated with the partial roll-out and conclude that we maintain reasonable power. Online Appendix A.3 further shows that attrition is low and independent of the treatment assignment.

1.6.1 Balance between assigned-to-be-treated-but-not-treated sub-counties and control sub-counties

One way to test if selection bias was introduced by the partial roll-out is to compare outcomes in sub-counties that were randomly allocated to function as controls to outcomes in sub-counties that were randomly allocated to receive treatment but did not end up receiving treatment. If the roll-out was random, these sub-counties could be interchanged. Finding no significant differences between these two groups would therefore support the hypothesis that the partial roll-out has not introduced selection bias. If the incomplete roll-out has introduced selection bias, comparing these groups will also be informative to assess the direction and magnitude of the bias.¹⁰

Table 1.1 presents differences between assigned-to-be-treated-but-not-treated sub-counties and control sub-counties for each of the three treatment arms affected by the partial roll-out. It is equivalent to the original balance table (Table A.1 in the Online Appendix), but after dropping sub-counties that were actually treated. We find that the imbalance related to the distance to the nearest all-weather road as shown in Online Appendix Table A.1 becomes more pronounced. This may indicate that the OPM prioritized the treatment of areas that are less remote. Similarly, the imbalance for the number of children in public schools is further accentuated, suggesting that information barazas were first organized in areas with fewer children in public schools prior to the study. At the same time, it should be noted that we only find significant differences for comparisons that rejected balance also in the original balance table.

¹⁰Alternatively, one could restrict attention to sub-counties that were randomly allocated to receive treatment and compare the ones that ended up being treated to the ones that did not. However, in our case, this comparison would be based on fewer observations.

Table 1.1: Balance between assigned-to-be-treated-but-not-treated sub-counties and control sub-counties

	mean	sc	baraza	information	deliberation
Household size	6.324 (2.825)	0.012 (0.171)	0.388* (0.170)	0.022 (0.140)	
Age of household head (years)	46.501 (14.615)	0.357 (0.714)	0.698 (0.663)	0.553 (0.808)	
Household head is woman (1=yes)	0.191 (0.393)	0.008 (0.017)	-0.019 (0.016)	-0.003 (0.017)	
HH head finished primary education (1=yes)	0.213 (0.410)	-0.007 (0.019)	-0.007 (0.027)	-0.003 (0.022)	
Thatched grass roof (1=yes)	0.298 (0.457)	-0.002 (0.029)	0.000 (0.024)	-0.036 (0.027)	
Traditional mud wall (1=yes)	0.424 (0.494)	0.007 (0.049)	-0.057 (0.047)	0.044 (0.044)	
Distance to nearest all weather road (km)	0.906 (0.915)	0.284** (0.131)	0.010 (0.100)	0.187 (0.110)	
Access to extension (1=yes)	0.108 (0.310)	0.005 (0.015)	0.008 (0.016)	0.007 (0.015)	
Village health team present in village (1=yes)	0.854 (0.353)	-0.007 (0.035)	-0.010 (0.028)	-0.015 (0.028)	
Number of children in public schools	2.478 (2.074)	0.043 (0.112)	0.249* (0.115)	0.076 (0.100)	
Number of observations	12545	4293	7842	8391	

Note: Column (1) reports sample means (and standard deviations below); column (2) reports differences between households that were planned to receive a sub-county information+deliberation baraza but did not receive any baraza, and those that were planned to not receive one (and standard errors below); column (3) reports differences between households that were planned to receive a sub-county information baraza but did not receive one, and those that were planned to not receive one (and standard errors below); column (4) reports differences between households that were planned to receive a sub-county deliberation baraza but did not receive one, and those that were planned to not receive one (and standard errors below); **, * and + denote significance at 1, 5 and 10% levels. Reported standard errors are clustered at the level of randomization. Distance to nearest all weather road is trimmed at 5% and transformed using the inverse hyperbolic sine transformation.

1.6.2 Selection of control sub-counties for end-line survey

Table 1.2 presents the factorial design that underlies the sub-county-level baraza impact evaluation. It shows that, based on the original power calculations, we planned to have between 51 and 53 sub-counties (corresponding to about 2,550 to 2,650 households) in each treatment cell. Just before the end-line data collection, 29 of 51 sub-counties that were supposed to receive the information treatment were treated, 18 of 51 sub-counties that were supposed to receive the deliberation treatment were treated, and 20 of 53 sub-counties that were supposed to receive the crossed (information + deliberation) treatment were treated. All eight districts that were supposed to be treated were treated, from which 40 sub-counties (corresponding to about 2,000 households) were sampled. Overall, this means that about 55% of all sub-counties that were planned to receive any form of baraza were treated, corresponding to about 5,350 households.

Table 1.2: Factorial design

	control	information
control	planned: 51	planned: 51
	included: 40	treated: 29
deliberation	planned: 51	planned: 53
	treated: 18	treated: 20

Statistical power is optimal when the number of treatment units is equal to the number of control units, and while adding more control sub-counties will increase power, these gains in power must be weighed against the costs of collecting additional data. As only a subset of sub-counties that were planned to be treated were treated, it would not be cost-effective to collect end-line data in all sub-counties that were not treated. We decided to collect end-line data in 40 control sub-counties, which is equal to the number of sub-counties with a district-level baraza (the largest number of sub-counties among the different treatment cells).

This implies that we had to decide in which of the potential control sub-counties (including those that were allocated to the control group and those that ended up not being treated) we should collect end-line data. One reasonable suggestion would be to pick them randomly. However, if the roll-out was not

random, this strategy may lead to a biased estimate of the causal impact of the intervention. For instance, if the OPM prioritized the treatment of less remote areas, randomly selecting control sub-counties may result in a sample in which less remote sub-counties are relatively under-represented and more remote sub-counties are relatively over-represented in the control group. A better strategy may involve matching each treated sub-county with a control sub-county that exhibits similar pre-treatment characteristics, which the planner was likely to know or have access to during the intervention roll-out (Bertsimas, Johnson, and Kallus, 2015; Kasy, 2016).

We matched each treated sub-county to a control sub-county that was similar in terms of a range of sub-county characteristics that may have affected the planner’s decision how to roll out the intervention. In particular, we match on the following characteristics that were obtained by surveying village chairs and CAOs of all sub-counties at baseline: geographic location (GPS coordinates of the headquarter), road infrastructure (kilometers of tarmac and all-weather gravel roads), share of households with electricity, share of households with an iron roof or tiles, number of health centers, female primary school dropout rate, number of Universal Primary Education schools, share of farmers using improved seed, and political connections of the sub-county (defined as having a minister or member of parliament coming from the sub-county). These characteristics are used in a probit regression to predict the probability that a sub-county is treated. We then match each treated sub-county to a candidate control sub-county that is closest in terms of the predicted probability of being treated. Note that we completed the matching before collecting end-line data and pre-registered the result here. Table 1.3 shows baseline balance for the resulting sample. Two significant differences in 40 comparisons can be expected from chance alone, and based on this, we conclude that we maintain balance between the treatment and control groups in the final sample.

Table 1.3: Balance between treated and control sub-counties (final sample)

	mean	sc baraza	information	deliberation	dist baraza
Household size	6.411 (2.855)	-0.186 (0.169)	0.065 (0.152)	-0.302 (0.166)	0.062 (0.248)
Age of household head (years)	47.009 (14.542)	1.096 (1.012)	-0.215 (0.731)	0.574 (1.038)	1.554 (0.998)
Household head is woman (1=yes)	0.191 (0.393)	0.025 (0.017)	-0.006 (0.018)	0.022 (0.024)	0.011 (0.015)
HH head finished primary education (1=yes)	0.208 (0.406)	0.005 (0.029)	-0.016 (0.025)	0.014 (0.035)	-0.018 (0.031)
Thatched grass roof (1=yes)	0.262 (0.440)	0.015 (0.030)	0.044 (0.030)	-0.007 (0.022)	0.037 (0.042)
Traditional mud wall (1=yes)	0.444 (0.497)	0.086 (0.058)	0.031 (0.053)	0.062 (0.058)	-0.008 (0.114)
Distance to nearest all weather road (km)	0.909 (0.912)	-0.279* (0.136)	0.027 (0.140)	-0.104 (0.135)	-0.229 (0.112)
Access to extension (1=yes)	0.105 (0.307)	0.011 (0.014)	0.000 (0.012)	0.012 (0.020)	0.018 (0.016)
Village health team present in village (1=yes)	0.865 (0.342)	0.020 (0.051)	0.019 (0.036)	0.090* (0.039)	0.075 (0.041)
Number of children in public schools	2.507 (2.072)	-0.089 (0.118)	0.001 (0.097)	-0.188 (0.111)	0.078 (0.154)
Number of observations	7340	2949	5298	5298	3999

Note: Column (1) reports sample means (and standard deviations below); column (2) reports differences between households that received a sub-county information+deliberation baraza, and those that did not receive any baraza (and standard errors below); column (3) reports differences between households that received a sub-county information baraza, and those that did not receive one (and standard errors below); column (4) reports differences between households that received a sub-county deliberation baraza, and those that did not receive one (and standard errors below); column (5) reports differences between households that received a district information+deliberation baraza, and those that did not receive any baraza (and standard errors below); **, * and + denote significance at 1, 5 and 10% levels. Reported standard errors are clustered at the level of randomization. Distance to nearest all weather road is trimmed at 5% and transformed using the inverse hyperbolic sine transformation.

While the ex-ante matching strategy may reduce bias resulting from incomplete roll-out, it may come at a cost. First, if the roll-out introduced selection, matching may further reduce the external validity of the study, as now also the control sub-counties are not a random sample of the study population anymore. Second, the decrease in potential bias for hypotheses related to sub-county-level barazas should be traded off against an increase in potential bias when testing differences between district-level barazas and control. As the sub-county-level analysis weighs higher in terms of research objectives, we decided to prioritize the reduction in potential bias at this level. However, both issues are only relevant if selection bias was introduced due to the partial roll-out.

1.7 Results

Before end-line data collection, we prepared a “mock report,” which contains the complete analysis on simulated end-line data and was pre-registered here. In this report, we committed to specific measures to judge impact. Overall impact and impact at sector level are assessed using indices that are composed of various outcomes in each sector following Anderson (2008).¹¹ Pre-registration and mock reports have been suggested to reduce intentional and unintentional selection of outcome variables and specifications that yield positive findings, leading to an unreliable body of published research (Humphreys, De la Sierra, and Van der Windt, 2013).¹²

Impact is assessed as a simple treatment-control comparison, implemented using an ANCOVA model that also controls for the region (which was used for stratification) and the baseline outcome. In particular, to estimate the impact of sub-county-level barazas with both information and deliberation components, we restrict the sample to pure control sub-counties (S_0^0 in Figure 1.1) and sub-counties that received the combined information and deliberation baraza (S_{ID}^0 in Figure 1.1) and estimate:

¹¹These indices are weighted means of the standardized values of the outcome variables. The weights maximize the amount of information captured in the index by giving less weight to outcomes that are highly correlated with each other. The variables that are included in the sector indices can be found in Tables A.3-A.6 in Online Appendix A.4. The use of indices is considered to be an effective strategy to account for multiple hypothesis testing.

¹²To further increase transparency, this article was prepared using LyX, an open source document processor, and all LyX and R code to replicate the analysis is placed under revision control using Git. The Git repository can be found here.

$$y_{h sr} = \alpha + \beta T_{sr} + \gamma \bar{y}_{h sr} + \tau_r + \varepsilon_{h sr} \quad (1.1)$$

Here, $y_{h sr}$ is the outcome variable (or outcome index) of interest for household h in sub-county s in region r , $\bar{y}_{h sr}$ is the corresponding outcome at baseline, T_{sr} is a binary indicator indicating if the sub-county received a combined information and deliberation sub-county-level baraza, τ_r are region fixed effects and $\varepsilon_{h sr}$ is an error term. In this equation, β is the parameter of interest and gives us the difference in average outcomes between households that live in sub-counties that received the baraza and those that live in sub-counties that did not, after controlling for baseline outcome and regional fixed effects. Similar comparisons are made between control sub-counties and sub-counties that received an information-focused baraza (S_I^0 in Figure 1.1, with T_{sr} as a binary indicator indicating if the sub-county received an information baraza) to estimate the effect of the information component; and between control sub-counties and sub-counties that received a deliberation-focused baraza (S_D^0 in Figure 1.1, with T_{sr} as a binary indicator indicating if the sub-county received a deliberation baraza) to estimate the effect of the deliberation component. Finally, an equation similar to equation (1.1) is estimated to assess the impact of district-level barazas but here, s refers to districts instead of sub-counties.¹³

Figure 1.2 provides a summary of the findings. It covers the four main hypotheses—the impact of sub-county-level barazas, the relative effectiveness of the information component, the relative effectiveness of the deliberation component, and the impact of district-level barazas—on four sectors: agriculture, infrastructure, health, and education.¹⁴ As mentioned above, these graphs are based on indices that are composed of various outcomes in each sector according

¹³We also compare district-level barazas to sub-county-level barazas by directly comparing outcomes of households that were sampled from S_{ID}^{ID} to outcomes of households that were sampled from S_{ID}^0 . In light of the shift from sub-county-level barazas to district-level barazas from 2012 on, this comparison is more relevant from a policy perspective, and it was also pre-registered. However, the partial roll-out of the intervention (see Section 1.6) reduced the number of sub-counties in S_{ID}^0 that ended up being treated. Comparisons of outcomes between areas that received a district-level baraza and areas that did not receive any baraza rely on more observations, leading to more statistical power.

¹⁴The agriculture index is composed of the variables: used inorganic fertilizer, used improved seed, extension visit at home, visited extension office/demo site/model farmer, NAADS/OWC in village, support in marketing from marketing committee, support in marketing from cooperative.

The infrastructure index is composed of the variables: used unprotected water source during dry season, distance to water source (km), average waiting time at source (min), water user committee in village, distance to nearest all weather road (km).

to Anderson (2008). Point estimates are standardized effects. We also combine the four indices in one overall index that assesses the impact on public service delivery in general.

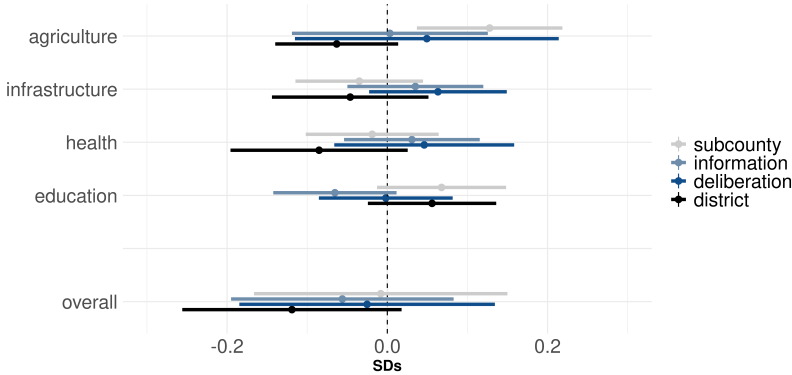


Figure 1.2: Summary of effects

We do not find any significant impact of the baraza program on overall public service delivery. We do find that sub-county-level barazas improved public services in the agricultural sector..

The indices combine various outcome variables, and for some of them the expected direction of the effect is unclear a priori. For instance, an information baraza may improve the quality of services in a hospital or health center when judged by an objective measure such as waiting time. However, the information may also result in higher expectations from the users. As such, citizen perceptions of quality may have reduced as a result of information barazas. That is why it is also interesting to look beyond the indices and consider outcomes individually. In Online Appendix A.4, we provide a detailed and more exploratory analysis of the individual outcomes behind the indices, and of some other outcomes that were not pre-registered. This analysis suggests a minor impact of sub-county-level barazas in some areas (such as schooling infrastruc-

The health index is composed of the variables: sought treatment for fever in public health facility, went to public health facility to give birth, Village Health Team in village, distance to nearest government health facility (km), number of days work/school missed due to illness, waiting time before being attended (min), visited traditional health practitioner.

The education index is composed of the variables: number of children in Universal Primary or Secondary School, distance to public school (km), school has boundary fence, school has water facility, school has School Management Committee (SMC), is informed about SMC, inspectors visited school.

ture or waiting time at boreholes). The Online Appendix A.4 further shows that the significant treatment effect on the agriculture index is driven by three variables: in sub-counties that held a sub-county-level baraza, households are more likely to have had an extension visit at home and to have received marketing support from a cooperative, and National Agriculture Advisory Services (NAADS) and Operation Wealth Creation (OWC) are more present in villages that belong to treated sub-counties.

1.8 Discussion

The main conclusion from our confirmatory analysis is in line with what other studies find: we do not find a significant impact of the baraza policy program as a form of community-based monitoring on public service delivery as a whole (Olken, 2007; Raffler, Posner, and Parkerson, 2020; Waddington et al., 2019). In this section, we explore some of the potential reasons why barazas do not seem to translate into large overall or sector-wide improvements in public services. We categorize these explanations under three headings: problems with the theory of change, the political context in which the program was implemented, and methodological challenges in evaluating real world policy interventions.

1.8.1 Problems with the theory of change

Community-based monitoring interventions such as the baraza program can only improve public service delivery if there is room for improvement: if politicians are already responsive to the needs of citizens and civil servants are already delivering services of high quality, it would be unlikely that barazas have any significant additional effect on objectively verifiable outcomes such as waiting time in health centers or student-teacher ratios.¹⁵ We think it is hard to argue that public services in Uganda cannot be improved. While progress has been made in some areas such as child mortality, other areas have seen less improvement. For instance, only about 10% of households report that they have access to agricultural extension, see Table 1.1. Baseline data further shows that patients at government health centers must wait for medical examination for

¹⁵Indeed, Raffler, Posner, and Parkerson (2020) argue that child mortality decreased significantly in response to the community-based monitoring intervention in Björkman and Svensson (2009) but not in their study is likely due to the fact that child mortality improved substantially over the last few decades, leaving little room for further improvement.

almost two hours on average and are examined by a nurse instead of a doctor in many cases. Access to water remains problematic, especially in the dry season when 36% of households turn to potentially unsafe drinking water.

Alternatively, if there are no funds available for public service delivery, community-based monitoring will most likely not lead to the desired outcomes. Again, we think that a lack of funding is not the main problem. The liberalization and decentralization drive in the mid-eighties seduced donors across the globe to pour money into Ugandan public services. While some studies report rampant elite capture (for example, Reinikka and Svensson, 2004), we asked government officials in all sub-counties of our sample at baseline about the proportion of the public service budget that they actually received for the fiscal year 2013/2014, and they reported to have received on average 83% of their health and education budgets.

The impact of the baraza program will also likely be muted if the priorities of politicians and civil servants are not aligned with the priorities of citizens.¹⁶ For priorities to be aligned, it is important that what is discussed during barazas is also what citizens think is most important. This is not automatically the case, as the most vocal citizens may dominate the discussion due to group dynamics and attempt to skew decisions during the meeting in their favor. Alternatively, politicians or civil servants could influence the course of the baraza. We explore the alignment between what citizens perceive as important and what transpires during the baraza for ten random sub-counties in Table 1.4. To this end, we use baseline data on priorities obtained from citizens living in a sub-county (column (3)) and compare it to detailed information about which sector was most discussed during each sub-county-level baraza using reports redacted by the OPM (column (4)). We find that what is prioritized by citizens is also discussed in the baraza in four out of ten sub-counties, providing indicative evidence of some degree of alignment between what citizens deem important and what is discussed during barazas.

Even if what transpires during barazas reflects the priorities of citizens, there will be no impact if officials subsequently do not act on it. To explore this link in the causal impact chain, we add a column that reports in which sector most positive change was recorded during end-line for each of the ten

¹⁶Note that unaligned priorities cannot fully explain why barazas have no effect: services may still improve, but in sectors which are not most important for citizens. However, assuming that citizens know best where progress can be made, unaligned priorities will reduce the effectiveness of barazas.

Table 1.4: Alignment of priorities

District	Sub-county	Community's priority at baseline	Prioritized sector during baraza	Largest change at end-line
Hoima	Kyabigambire	Health	Health	Education
Kibaale	Kasambya	Health	Agriculture	Infrastructure
Masaka	Mukungwe	Health	Agriculture	Agriculture
Hoima	Kiziranfumbi	Health	Health	Health
Ssembabule	Lwemiyaga	Health	Education	Agriculture
Hoima	Buseruka	Education	Education	Health
Luwero	Katikamu	Health	Health	Infrastructure
Ntoroko	Kanara	Infrastructure	Education	Health
Pader	Puranga	Health	Infrastructure	Health
Kanungu	Kanungu TC	Health	Infrastructure	Education

sub-counties to Table 1.4.¹⁷ We find that in only two out of ten sub-counties, the sector that was discussed most during the baraza event was also the sector where most change was recorded.

Table 1.4 suggests that even though the priorities of citizens are to some extent discussed during barazas, it is less common that what was discussed during the events also gets implemented. Using focus group discussions, we looked for reasons why this link breaks down. Firstly, poor relationships between politicians and technical personnel (e.g., between a CAO and a LC5 Chairperson) can hinder the effectiveness of barazas. Secondly, if barazas mainly expose issues beyond the operational jurisdiction of the officials and service providers that are present, they will not improve services. Thirdly, if barazas are perceived as one-off events and no subsequent meetings are organized in the same area, officials may be less likely to implement what was decided. Optimism among citizens and other stakeholders can quickly turn into disappointment if none of the issues raised is addressed. To investigate the role of the timing of the intervention, we reran the analysis and added an interaction term between the treatment indicator and an indicator variable that takes the value of one if the baraza that the household was exposed to had happened more than one and a half years before the end-line data collection.¹⁸ Results are summarized in Figure 1.3 and show that the effects of barazas dissipate over time, especially for infrastructure and health. This suggests that the policy program is not effective because barazas are not organized more frequently and enthusiasm fades, plans are abandoned, and promises forgotten over time.

¹⁷This is based on the indices that we created earlier to measure the impact at sector level (see Section 1.7).

¹⁸The slow roll-out of the intervention introduced variation in the time that passed between treatment and end-line data collection. For instance, the first barazas were held around June 2016, so that more than three years have passed between treatment and end-line. For the most recent barazas, there are only a few months between treatment and end-line. The indicator is coded as zero for the control group.

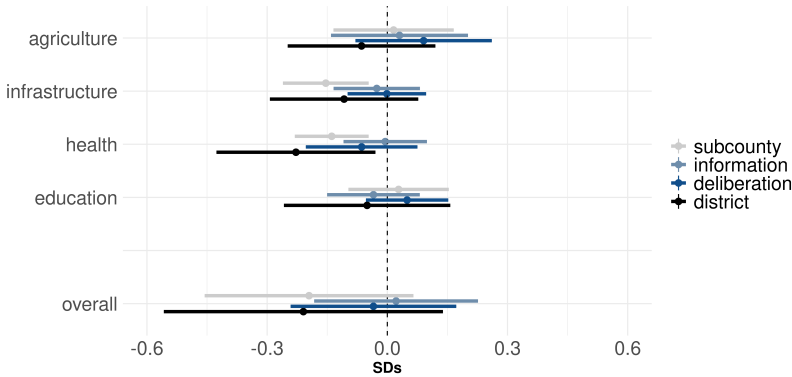


Figure 1.3: Summary of effects (heterogeneity: timing of intervention)

1.8.2 Political context

The political environment in which barazas are organized may also affect their effectiveness. Indeed, drawing on a systematic review of more than 90 social accountability interventions, Hickey and King (2016) argue that politics and contexts are critical for success. Raffler, Posner, and Parkerson (2020) note that citizens can only apply bottom-up pressure on political representatives if they are responsive to electoral pressure and that this condition is often not fulfilled in countries with uncompetitive, semi-democratic political systems.

Barazas are organized by the OPM and several officials who belong to this office have appeared in the news in the context of large corruption scandals over the last decade. We also saw in Subsection 1.2.2 that the meetings are chaired by the RDC's office. RDCs are direct representatives and appointees of the president, and this may affect how the public views barazas. Some may even argue that barazas cannot be characterized as a form of community-based monitoring or any collective bottom-up action at all because they are helicoptered into communities by the OPM and to some extent impose interactions between citizens, politicians, and civil servants. Instead, some groups in society (e.g., supporters of the opposition) may not perceive barazas as open fora but as highly politicized events, especially in comparison to the *ebimeeza* meetings which the government banned when the baraza program was rolled out (see Subsection 1.2.1). Government officials may abuse the meetings to pursue their own agenda, and indeed, some reports indicate that barazas turned

into political rallies, especially when elections were approaching.

To explore the importance of the political context and how the program was received in pro-government and pro-opposition areas, we look at treatment heterogeneity related to a sub-county's connection to people in power. The assumption is that a sub-county is more likely to be pro-government if it delivers higher-level politicians, such as ministers, members of parliament, heads of government agencies, RDCs, and other district- or central-level officials. One may expect that in these areas, barazas are more effective, since citizens can use their connections to the center to attract funding to issues raised, and politicians may want to reward their constituents. Figure 1.4 reports interaction effects between the different interventions and an indicator variable that takes the value of one if a sub-county is well connected to people in power. Based on the indices, we do not find significant interactions. However, coefficients are generally positive, especially for interactions with deliberation and district-level barazas.

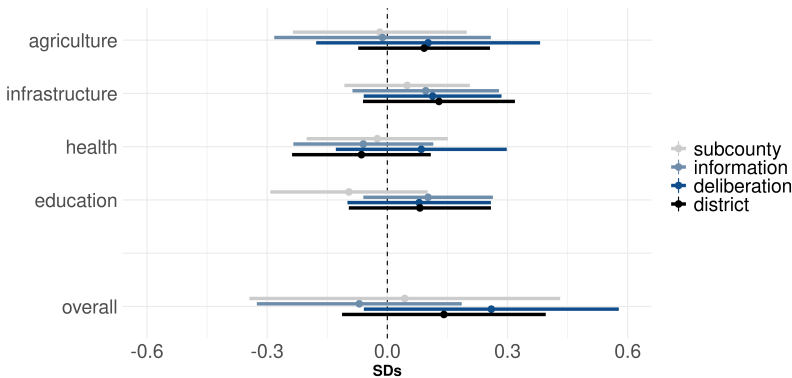


Figure 1.4: Summary of effects (heterogeneity: connection to people in power)

1.8.3 Methodological issues

We provided a detailed description of the implementation challenges surrounding the baraza program in Section 1.6. We also argued that the primary concerns related to the partial roll-out, namely reduced statistical power and potentially selective roll-out, are unlikely to affect our conclusions. In this subsection, we focus on treatment heterogeneity as a potential reason why we do not find more convincing effects. Focus group discussions with citizens indicated

that barazas were often perceived as effective: they provided different examples of changes in public services that they thought were due to, or accelerated by, the baraza event. Their stories suggest that the impact of barazas on public service delivery may be highly localized and context specific.

Barazas are fairly broad interventions that attempt to address a range of issues in heterogeneous settings, resulting in interventions that are unlikely to be standardized across treatment units, at least compared to typical treatments in biomedical sciences (Barrett and Carter, 2010). For example, a baraza in one sub-county may lead to significant improvements in smallholder access to extension services, but to no improvements in infrastructure. A baraza in another sub-county may accelerate the construction of an all-weather gravel road but agricultural public services remain the same (or even suffer because resources are re-allocated between sectors). Both barazas would be effective, yet outcomes are averaged over treatment units, and impact may be diluted.

1.9 Conclusion

To improve governance and public service delivery in Uganda, the OPM organizes town hall-style meetings—popularly known as barazas—where citizens receive information from government officials and get the opportunity to engage with them. In 2015, we designed a cluster RCT aimed at evaluating the effectiveness of this large-scale, high-profile policy intervention. The study set out to answer four research questions: (1) What is the impact of a baraza as originally conceived and implemented by the OPM? (2) What is the relative importance of the information component of a baraza? (3) What is the relative importance of the deliberation component of a baraza? (4) Should a baraza be organized at the district level or at the sub-county level? Baseline data from more than 12,500 households spread over almost 250 sub-counties in about 40 districts throughout Uganda was collected and the OPM started implementing barazas following our experimental design and protocols.

The OPM faced various challenges that affected the timely roll-out of the barazas, including budgetary constraints and disruptions related to the general elections of 2016, which resulted in the decision to collect end-line information from 6,700 households after partial roll-out. This paper therefore also serves as an illustration of the challenges of evaluating real-world policy interventions and how to deal with them, as we follow various strategies to diagnose, and

reduce the consequences of, potential selection bias and sample size reduction. We find that the partial roll-out did not introduce significant bias and that we retain sufficient statistical power.

Judged by an index of overall public service delivery, we find no significant impact of the program, even though sub-county-level barazas seem to have improved service delivery in the agricultural sector. We then consider various potential reasons why barazas did not have the expected sector-wide benefits. We find indications of problems in the theory of change: while priorities of citizens do seem to get discussed during barazas, solutions are not always implemented. Furthermore, the political context in which barazas are organized is likely to affect their effectiveness. We also reflect on the heterogeneity of community-based monitoring interventions, complicating the estimation of average treatment effects.

Our findings are in line with a recent study by Raffler, Posner, and Parkerson (2020) that failed to replicate Björkman and Svensson (2009). In general, initial high hopes about the potential of community engagement to improve public service delivery seems to have lowered as more empirical evidence becomes available. In a recent mixed-methods systematic review of participation, inclusion, transparency and accountability (PITA) initiatives, Waddington et al. (2019) suggest that citizen engagement may have positive effects on intermediate user engagement outcomes such as meeting attendance, access to, and quality of certain public services. At the same time, effects on public service outcomes such as education or child health are often limited. In many cases, citizen participation also does not lead to changes in service provider action outcomes such as public spending, staff motivation, and corruption. More research is needed to identify the conditions and complementary interventions needed to leverage the voices of citizens.

1.10 Acknowledgments

This research has been funded by the International Initiative for Impact Evaluation, Inc. (3ie) through the Global Development Network (GDN) under grant number RFQ PW2.18.UG.PG. This research was partly funded by the CGIAR Research Program on Policies, Institutions, and Markets (PIM), led by the International Food Policy Research Institute (IFPRI) and carried out with support from the CGIAR Fund contributors under grant number PIM 203002.002.221. This research was cleared by the Uganda National Council for Science and Technology (UNCST SS 5179), Makerere University School of Social Sciences Research Ethics Committee (MAKSS REC 05.19.291), as well as IFPRI's Institutional Review Board (DSG-19-1053). We would like to thank Alvina Erman, Jennifer Smart, Marc Charles Wanume, Wilberforce Walukano, Fiona Natterambo, Emmanuel Bizimungu, Leocardia Nabwire, Richard Ariong, and the OPM, Joseph Muserero in particular. The views expressed in this article are not necessarily those of 3ie or its members, or of GDN.

Gender bias in customer perceptions: The case of agro-input dealers in Uganda

This chapter is co-authored with Anusha De (KU Leuven, Belgium and University of Göttingen, Germany) and Bjorn Van Campenhout (IFPRI and KU Leuven, Belgium). An earlier version of this chapter is published as IFPRI Discussion Paper 2132.

2.1 Introduction

In the context of incomplete and imperfect information, economic actors rely predominantly on perceptions and use mental shortcuts to make decisions using limited data (Kahneman, 2017). Reliance on instincts and emotions becomes dominant if it is difficult to objectively assess the value of a commodity or service being bought and sold. However, perceptions and decision heuristics may suffer from a variety of cognitive biases such as stereotype thinking and availability bias and may be influenced by social and cultural phenomena such as homophily effects and prevailing norms and customs.

Agricultural inputs such as inorganic fertilizers or improved seed varieties (high yielding cultivars like open-pollinated or hybrid maize varieties), lie somewhere on the continuum between experience goods and credence goods. When farmers inspect products at the agro-input shop, they can assess quality only superficially from readily observable characteristics such as the homogeneity of the seed or by checking if the fertilizer package is intact. Even after farmers

used the commodity and observed the yield, it may be difficult for them to learn about the quality of the seed or fertilizer, as many other factors in addition to the input affect yield. That is why perceptions and emotions often take the upper hand when farmers acquire agricultural inputs.

In addition to the difficulty of judging the quality of agricultural inputs, several studies note that there is considerable heterogeneity in the actual quality of these inputs in the market. For instance, Bold et al. (2017) test agricultural inputs purchased in local markets in Uganda and find that 30% of nutrients are missing in fertilizer, and hybrid maize seed is estimated to contain less than 50% authentic seed. Also in Uganda, Ashour et al. (2019) test herbicides and find that the average bottle in their sample is missing 15% of the active ingredient and 31% of samples contained less than 75% of the ingredient advertised. While it remains unclear if quality-related issues are the result of deliberate adulteration or poor storage and handling, and at what point in the value chain quality starts to deteriorate, the resulting uncertainty makes reliance on perceptions and decision heuristics more likely (Barriga and Fiala, 2020).

In traditional societies with strong gender norms and customs, small businesses alongside food supply chains are often some of the few options for women to earn money independently from their husbands. While rapid urbanization has led to the emergence of fast-food restaurants, informal food vendors, who tend to be self-employed women, are still the main source of food for most households in sub-Saharan Africa (Giroux et al., 2021). And while supermarkets are emerging throughout the developing world, wet markets where mostly women sell products continue to account for most of the expenditure on fresh produce in many countries (Gorton, Sauer, and Supatpongkul, 2011). In Uganda, we find that a surprisingly large share of agro-input shops are operated and/or managed by women.

However, the same gender norms and customs also mean that perceptions may be stacked against women if they venture into areas such as agro-input provision. Farmers, both male and female, may believe modern agricultural technologies fall in the male domain. Furthermore, agro-input shops primarily deal in seed for semi-commercial crops such as maize or rice, as opposed to food security crops such as beans or cassava. Again, commercial crops are often considered to be the responsibility of men, while women are expected to take care of the household food supply (Dolan, 2001; Orr et al., 2016). A case in point is the recent study by Ntakyio and Van Den Berg (2022) which confirms

this traditional stance in the context of smallholder production in Uganda. The authors find a significant negative impact of a commercialization program on women empowerment in crop production and their control over income, clearly showing a power shift to men in rural households. Hence, we conjecture that female-managed agro-input shops may be disadvantaged when farmers form opinions about the quality of services rendered or goods sold, deeming women not to be fit for these commercial roles.

In this paper, we test if farmers perceive agro-input shops managed by women less favorably than agro-input shops under male management using a unique dyadic data set of farmer-dealer links.¹⁹ To operationalize perceptions, we asked farmers to rate agro-input dealers, on a scale of one to five, on a range of characteristics. We then make between-dealers comparisons, explicitly accounting for observable differences in the quality of male- and female-managed shops. Furthermore, we use the fact that a farmer has generally rated more than one agro-input dealer. If the same farmer rates both male and female-managed agro-input shops, we can exploit this within-farmer variation and control for farmer specific observable and unobservable confounders.

We find that farmers generally rate male-managed agro-input shops more favorably than shops managed by women. The difference in ratings is largest when farmers are asked to rate the agro-input dealership in terms of price competitiveness and in terms of reputation. We also find that the quality of seed from male-managed agro-input shops is rated higher than the quality of seed from shops managed by women. As the differences in ratings persist after explicitly controlling for the quality of the dealerships and the services and products they provide, we conclude that gender-biased customer perceptions persist and create comparative disadvantages and entry barriers for female-managed agro-input shops. This gender equity bias directly affects social outcomes like women's capabilities, aspirations, and their empowerment in agricultural and seed systems. Additionally, there are consequences at more aggregate levels: as almost half of the agro-input shops in our sample are managed by women, the finding that farmers do not trust these shops may impose challenges for varietal turnover, hindering agricultural productivity, food security, and rural transformation.

¹⁹The paper builds on earlier exploratory work published in Van Campenhout and De (2023) that looks at gender-related perceptions in Uganda's maize value chain more broadly, and prompted us to formulate a more specific hypothesis and collect data to test this using appropriate quantitative methods.

2.2 Research question and relation to the literature

We aim to test if gender related discrimination is present in the way small-holder maize farmers in southeastern Uganda perceive agro-input dealers in their neighborhood. In the wider literature, gender related discrimination is often referred to as gender equity bias—behavior that shows favoritism toward one gender over another. Gender equity bias has been confirmed in a wide range of contexts, usually when people are asked to assess the performance of another person. Stereotyping and role congruence are often catalysts for distorted perceptions and false beliefs about the abilities of groups of people. We highlight some of the most important studies that search for systematic bias related to the gender of the person being assessed.

Gender equity bias often surfaces when individuals decide on who to engage with, be it who to work with, who to elect as leaders, or who to consult. For example, when it comes to hiring decisions, managers must decide based on limited information. Discrimination in labor markets, including discrimination related to gender, has been documented in several studies. Wu (2020) uses data from an online forum for economists called “Economic Job Market Rumors” to measure gender bias in discussions about women versus men. Gender equity bias is also studied in the context of the wage gap, that is, when women appear to make substantially less money for the same work than their male counterparts. Often, this is also tied to gender equity bias in performance appraisals, where (often male) managers’ gendered beliefs and perceptions creep into evaluations of their subordinates (Correll et al., 2020).

Another area where gender equity bias has been studied extensively is in scientific publishing using peer review. For instance, Card et al. (2019) look at differences in rejection rates at four top economics journals. They compare male-authored papers to female-authored papers, using citations as a noisy measure of quality to account for potential sources of divergence, other than gender, between the two. They find that editors largely follow referees, resulting in a 1.7 percentage point lower probability of a revise and re-submit verdict for papers with female authors relative to a citation-maximizing benchmark. However, evidence on gender biases in the evaluation of economic research remains mixed. For example, Chari and Goldsmith-Pinkham (2017) find no disparity in the acceptance rates of female- and male-authored papers for National Bureau

of Economic Research conferences; Hospido and Sanz (2021) do find a significant advantage for male authors being accepted at three different European conferences. Gender equity bias has also been studied extensively in student evaluations of teaching. For instance, Mitchell and Martin (2018) find that the language students use in evaluations of male professors is significantly different from their language when evaluating female professors. They also show that a male instructor administering an identical online course as a female instructor receives higher ordinal scores in teaching evaluations.

Gender equity bias is also pervasive in politics. Pair et al. (2021) use Natural Language Processing to search for gender bias in Kenya’s leading newspaper and sentiment analysis to predict quantitative sentiment scores for sentences surrounding female leader names compared to male leader names. They find evidence of improvement in gender equality but also a backlash from increased female representation in high-level governmental leadership. Le Barbanchon and Sauvagnat (2021) find that female candidates obtain fewer votes in municipalities with higher gender earning gaps. Klein, Shtudiner, and Zwilling (2021) find that the adviser’s gender is one of the most important factors influencing a customer’s choice of financial adviser. The female advisers’ gender was found to have a negative effect on the desire to invest, and this negative attitude was found to be significantly higher among male respondents.

In the context of small and medium-sized enterprises (SMEs), in the agricultural sector in particular, we find few studies that look at gender equity bias. Alibhai et al. (2019) who study discrimination against female-led SMEs in Turkey come closest. Conducting a novel loan application experiment with 77 officers in banks, they find that 35% of the loan officers are biased against female applicants, with women receiving significantly smaller loans than men. The authors argue that loan officers may use gender bias as a heuristic device given limited information and risk aversion.

Gender bias features so prominently in areas such as labor markets, scholarly peer review, or teaching assessments partly because perceptions are made explicit in the process, for instance through review reports, student feedback, or hiring committees. However, in economic transactions, gender biases remain hidden as perceptions are never measured. As a result, differences in outcomes are often attributed to various other causes, such as differences in education or ability between men and women.

2.3 Context and data

2.3.1 Study context

The study was conducted in Uganda. As in many traditional agricultural societies, women's roles are mainly domestic, including housekeeping, child rearing, fetching water, cooking, and tending to community needs. Strong gender norms and stereotypes about the different capabilities of women and men imply that many women shy away from economic activities such as cash cropping or post-harvest processing. Women do participate to some extent in economic life through marketing as owners of small shops or vendors during market days. Even though the government signaled a willingness to mainstream gender as early as in 1997 with its first National Gender Policy, and women are reasonably represented at higher levels of government, norms and customs prevent women in most rural areas of Uganda from participating fully in economic life.

2.3.2 Sample

Our study area comprises eleven districts in southeastern Uganda, and roughly corresponds to the Busoga Kingdom. We include agro-input dealers located in trading centers and villages as well as smallholder maize farmers that live in the catchment areas of these key market sheds. The dealer sample was obtained by listing all agro-input shops in the area during a census, which resulted in 193 dealers. We collected information on their characteristics in September and October 2020.

After the census, these agro-input shops were grouped in catchment areas based on their location. A catchment area is defined as the area that is served by a dealer, the area where this dealer's customers live. If catchment areas of two or more dealers overlap because these dealers operate in the same town or trading center, they are assigned to the same catchment area. This is done based on their geographical location. Using GPS coordinates of the shops, the haversine function constructs an adjacency matrix, and shops that are less than five kilometers apart are assigned to the same catchment area. The five kilometer threshold was selected based on a visual inspection of the map, the size of an average village and the mean reported distance between farmers and dealers. The 193 agro-input dealers in Busoga were assigned to 65 catchment areas. In some catchment areas, there is a high density of shops, while in others

there are only one or two dealers. On average, there are 2.7 dealers in an area, with a minimum of one and a maximum of 18.

In each catchment area, we also sampled farmers in proportion to the number of agro-input dealers in the area. We connected shops to villages by asking every dealer where most of his or her customers come from. Enumerators were sent to these villages and instructed to randomly sample ten households that grow maize. Consequently, we sampled 1,931 smallholder maize farmers and collected information about their characteristics in April 2021. To increase the number of ratings, a second round was collected in January and February 2022. While these ratings were provided by the same farmers, only 1,893 of them were found during the second wave. These two rounds of surveys constitute the key sources of data for the study. Note that we only include shops in our analysis if the gender of their manager did not change between the first and the second round of rating collections.

2.3.3 Descriptive statistics

Table 2.1 describes the average agro-input dealer included in our study, differentiated by gender of the shop manager. When enumerators approached a shop, they tried to interview the manager, i.e., the person who is most knowledgeable about the day-to-day operations of the business, inventories, sales, and so on. It may be that the shop is owned by one person, but the owner employs another person to manage it. About 60% of managers are male. In 63% of the cases, the male respondent is also the owner of the shop, while only 47% of female managers are also the owner. If the gender of the shop manager is different from the gender of the owner, the question emerges who's gender affects perceptions of the shop as a whole. We feel that the person who manages the shop is most visible and as such most likely to affect perceptions, therefore the gender of the respondent determines if a shop is categorized as female or male-managed in our analyses.

There is substantial heterogeneity across shops. Some are small informal shops located in rural areas, which sell other goods and only stock seed during the planting season. Others are located in towns or trading centers and specialize in farm inputs and tools. The average shop has been in operation for about five to five and a half years. We already see signs of potential discrimination when we look at the number of customers: male managers report serving more

than 51 customers a day, female managers serve only 36 customers per day. A shop stocks on average three maize seed varieties. More statistics describing seed handling and storage practices, efforts, and services of agro-input dealers, are presented in Table 2.12 in the appendix. We see that female managers do not seem to handle or store seed in less appropriate ways than male managers. In fact, on many measures, female managers appear to do better than male managers.

Table 2.2 provides descriptive statistics of the farmers included in the study. The average farmer in our sample works on a small farm, with about 3.4 acres of land for crop production. Half of our sampled farmers indicate that they used improved maize seed, i.e., seed of a hybrid or open-pollinated variety, on at least one plot in the season preceding the survey, and of the farmers that used improved seed, two-thirds obtained it from an agro-input shop. However, fertilizer use is low. As a result, productivity is also low, with the average farmer harvesting only about 450 kg of maize per acre. Almost 70% of farmers believe that maize seed sold at agro-input shops is counterfeit.

2.3.4 Measuring perceptions

Quantifying perceptions of the quality of services provided by agro-input dealers and of the products they sell—improved maize varieties in particular—is central to our analysis. To do so, we asked farmers to rate agro-input dealers in the catchment area on a range of attributes. We broadly categorized the attributes into two families. A first set of indicators attempts to measure overall quality of agro-input dealers and the services they provide, while a second set of indicators has a narrower focus and asks about maize seed, a particular product sold by the agro-input dealer.

To measure the perceived quality of agro-input dealers, farmers were asked to rate these dealers on a scale of one (worst) to five (best) on their general quality, location (convenience, accessibility, closeness to clients), price (competitive pricing, discounts), seed quality, stock (availability of seed, number of varieties in stock), and reputation (what do other farmers think about the dealer). We also compute an average of these six dealer-level ratings. For these indicators, farmers were asked to rate the shop as a whole.

To measure the perceived quality of seed, farmers were asked to rate the improved maize varieties that dealers sell on a scale of one to five on their gen-

Table 2.1: Descriptive agro-input dealer statistics

	Male					Female				
	mean	min	max	SD	obs.	mean	min	max	SD	obs.
Dealer's age in years	32.114	15	59	10.951	114	30.519	18	50	8.474	79
Dealer finished secondary education	0.405	0	1	0.493	111	0.364	0	1	0.484	77
Dealer owns shop	0.632	0	1	0.485	114	0.468	0	1	0.502	79
Dealer received training on seed handling	0.614	0	1	0.489	114	0.532	0	1	0.502	79
Shop's distance to nearest tarmac road in km	5.595	0	39	8.470	114	6.224	0	40	10.225	78
Distance between shop and farmer in km ¹	5.157	0	13	3.300	114	5.297	0	13	3.087	79
Shop only sells farm inputs	0.702	0	1	0.460	114	0.823	0	1	0.384	79
Number of customers per day	50.850	2	300	55.759	113	35.544	2	150	33.443	79
Number of customers buying maize seed per day	25.327	1	250	31.043	113	19.696	0	100	23.485	79
Number of years since shop's establishment	5.684	0	33	6.252	114	5.063	0	25	5.910	79
Number of maize varieties in stock	3.123	0	10	2.096	114	2.962	0	10	1.720	79
Number of hybrid maize varieties in stock	1.868	0	8	1.594	114	1.709	0	6	1.293	79
Number of maize OPVs in stock	1.351	0	5	0.776	114	1.228	0	3	0.619	79
Sales price of maize seed in UGX/kg	4331.212	2500	10000	1148.673	110	4386.184	2800	12000	1482.227	76
Cost of maize seed for dealer in UGX/kg	3481.250	2000	8500	886.174	108	3585.915	2200	7750	1064.599	71
Revenue from maize seed in million UGX	10.290	0	82	18.041	112	6.722	0	81	14.592	79
Amount of maize seed dealer bought from provider in kg	1344.900	0	52500	5986.667	110	412.662	10	7500	977.195	74
Amount of maize seed dealer sold in kg	2106.874	0	16500	3286.934	111	1352.013	0	16900	2875.766	79
Shop's cleanliness/professionalism rating by enumerator	3.465	1	5	1.191	114	3.557	1	5	1.071	79
Shop received seed-related complaint from customer	0.711	0	1	0.456	114	0.620	0	1	0.488	79
Shop is registered with UNADA	0.486	0	1	0.502	107	0.493	0	1	0.503	75
Shop has trading license from local government	0.830	0	1	0.377	112	0.731	0	1	0.446	78

Note: SD is the standard deviation.

Number of observations: All 193 agro-input dealers are included in this table.

¹The distance between shop and farmer is calculated using the haversine function based on the GPS coordinates obtained during data collection. Farmers were asked to rate agro-input shops they know (one farmer can rate multiple shops) and these shops are not necessarily located in the immediate vicinity.

Table 2.2: Descriptive farmer statistics

	mean	min	max	SD	obs.
Homestead's distance to nearest tarmac road in km	8.850	0	100	9.391	1844
Homestead's distance to village headquarters in km	0.744	0	12	0.899	1914
Homestead's distance to nearest agro-input shop in km ¹	3.826	0	52	4.894	1858
Farmer's age in years	48.513	20	97	13.344	1923
Household head is male	0.789	0	1	0.408	1931
Farmer is married	0.881	0	1	0.323	1931
Farmer finished primary education	0.525	0	1	0.500	1913
Number of people in household (incl. respondent)	8.651	1	25	4.029	1931
Years since farmer started growing maize	22.851	0	82	13.004	1931
Farmer is member of farmer group/association/cooperative	0.132	0	1	0.338	1927
Farmer's land for crop production in acres	3.350	0	80	3.980	1915
Yield in kg/acre	456.998	22	2200	341.842	1881
Farmer used improved maize seed for any field last season	0.499	0	1	0.500	1929
Farmer bought this maize seed at agro-input shop	0.671	0	1	0.470	932
Farmers thinks seed at agro-input shops is counterfeit/adulterated	0.684	0	1	0.465	1512
Farmer is satisfied with maize seed used on plot	0.666	0	1	0.472	1931

Note: SD is the standard deviation.
Number of observations: All 1,931 farmers are included in this table.
¹The homestead's distance to the nearest agro-input shop is the distance reported by the farmer. Respondents could only report one answer for the nearest shop.

eral quality, yield, drought tolerance, pest and disease tolerance, crop duration or maturation period, and germination reliability. We also compute an average of these six seed-level ratings. Here, farmers were asked to rate the product itself. Farmers were also allowed to indicate that they could not rate seed on a particular dimension (e.g., because they never bought seed from the agro-input dealer). Then this dimension was not considered when computing the average.

We asked farmers to rate agro-input dealers twice, a first time in April 2021 and a second time in January and February 2022. The average farmer in our data set provided ratings for about two agro-input dealers, with some farmers rating up to 15 dealers. The average agro-input shop received ratings from almost twelve farmers, while one shop received ratings from almost 50 farmers. Table 2.3 provides descriptive statistics of the ratings used in our study. For example, when assessing the quality of maize seed sold by male agro-input dealers, farmers rate its germination with 3.67 out of 5 on average. Farmers generally rate dimensions related to the dealership better than dimensions related to the product. For instance, the mean location rating is 3.91 out of 5 for female agro-input dealers.

2.4 Empirical strategy

Our empirical strategy exploits the nature of the data, that is, that farmers in our data set rate several agro-input dealers (and dealers are rated by several farmers). A useful starting point is the following specification:

$$y_{f,d} = \alpha + \beta g_d + \varepsilon_{f,d} \quad (2.1)$$

Here, $y_{f,d}$ represents the rating, on a scale of one (poor) to five (excellent), given by farmer f to agro-input dealer d . g_d is the main variable of interest—the gender of dealer d . α and β are parameters to be estimated, and $\varepsilon_{f,d}$ is a residual.

Because the same farmer may rate several agro-input dealers, we cannot assume that the ratings $y_{f,d}$ in equation (2.1) are independent. For example, the ratings that a farmer provides may be affected by (potentially unobservable) characteristics of the farmer (e.g., a poor experience with an agro-input dealer in a previous year), which may affect the ratings of all agro-input dealers this farmer rates. Furthermore, the same agro-input dealer may be rated by several

Table 2.3: Descriptive agro-input dealer ratings by farmers

	Male					Female				
	mean	SD	Q ₁	Q ₃	obs.	mean	SD	Q ₁	Q ₃	obs.
Dealer's maize seed rating on general quality	3.790	0.903	3	4	629	3.815	0.819	3	4	363
Dealer's maize seed rating on yield	3.583	0.942	3	4	616	3.525	0.883	3	4	356
Dealer's maize seed rating on drought tolerance	3.029	0.877	2	4	594	3.000	0.892	2	4	343
Dealer's maize seed rating on pest/disease tolerance	2.467	0.928	2	3	599	2.457	0.974	2	3	350
Dealer's maize seed rating on speed of maturing	3.847	0.750	4	4	601	3.809	0.718	3.5	4	351
Dealer's maize seed rating on germination	3.673	0.913	3	4	608	3.682	0.875	3	4	359
Dealer's rating on general quality	3.773	0.994	3	5	644	3.667	1.025	3	4	378
Dealer's rating on location	3.755	1.260	3	5	644	3.907	1.176	3	5	378
Dealer's rating on price competitiveness	3.280	1.218	3	4	644	3.214	1.121	3	4	378
Dealer's rating on seed quality	3.781	1.096	3	5	644	3.825	1.036	3	5	378
Dealer's rating on seed stock	3.919	1.058	3	5	644	3.894	1.151	3	5	378
Dealer's rating on reputation	4.182	0.940	4	5	644	4.130	0.934	4	5	378

Note: SD is the standard deviation, Q₁ the 1st quartile, and Q₃ the 3rd quartile.
The minimum rating is 1 and the maximum rating is 5.
Number of observations: All ratings by farmers given to dealers during the first round of collection are included in this table, conveying the dyadic nature of the dataset.

farmers, leading to interdependence in the other dimension. For example, the ratings that a dealer receives may be affected by (potentially unobservable) characteristics of the dealer (e.g., dealer friendliness), which may affect the ratings given by all farmers that rated this dealer. To account for this two-way interdependence in equation (2.1), we define a composite error term ($\varepsilon_{f,d}$) that can be decomposed into a farmer specific component (ν_f), an agro-input dealer specific component (ω_d), and a residual ($\epsilon_{f,d}$) that varies at the level of the farmer-dealer interaction.

$$\varepsilon_{f,d} = \nu_f + \omega_d + \epsilon_{f,d} \quad (2.2)$$

Equation (2.2) shows that the dyadic nature of our data leads to two-way clustering in the error term. If the error term is uncorrelated with the explanatory variable(s) included in equation (2.1), Ordinary Least Squares (OLS) remains consistent. However, not considering within-cluster error correlation generally leads to standard errors that are biased downward, leading to under-rejection of the null hypothesis that gender does not affect ratings. In our case, it should be noted that clustering is non-nested. As traditional cluster-robust inference can only deal with clustering in one of the dimensions, our strategy will consist of including sufficient regressors to minimize concerns about error correlation at the agro-input dealer level, and then cluster standard errors at farmer level (Cameron and Miller, 2015).

To test for an agro-input dealer gender effect, we can simply compare average ratings received by male-managed shops and average ratings received by female-managed shops. Equation (2.3) shows how this can be done using a simple OLS regression on dealer-level averages.

$$\frac{1}{F} \sum_f y_{f,d} = \alpha + \beta \frac{1}{F} \sum_f g_d + \frac{1}{F} \sum_f \nu_f + \frac{1}{F} \sum_f \omega_d + \frac{1}{F} \sum_f \epsilon_{f,d} \quad (2.3)$$

$$\bar{y}_d = \mu + \gamma g_d + \bar{\varepsilon}_d \quad (2.4)$$

In equation (2.3), the identification of the gender equity effect (β) relies on differences between agro-input dealers. As a dealer's gender is constant for all farmers that rate this dealer, the average g_d is also a binary indicator of the gender of that particular dealer d . The farmer-specific component ν_f is

absorbed in the intercept term μ , while the dealer-specific component ω_d is now included in the error term $\bar{\varepsilon}$.

It is important to note that in equation (2.4), the dealer-specific error component $\frac{1}{F} \sum_f \omega_d$ in the error term $\bar{\varepsilon}_d$ may be correlated with the independent variable g_d . This would be the case if, for example, female agro-input shop managers are less educated on average than male agro-input shop managers, and less educated dealers get lower ratings by farmers. In this case, differential ratings are not caused by gender, but rather driven by differences in education. Therefore, in all regressions, we control for the education level of the shop manager, and add an additional regressor (x_d) to equation (2.4). A similar argument can be made for the age of the agro-input shop manager, which is a characteristic that is also easily observable and likely to affect perceptions:

$$\bar{y}_d = \mu + \gamma g_d + \varphi x_d + \bar{\varepsilon}_d \quad (2.5)$$

One may wonder if controlling for age and education is sufficient, as causal inference using regressions based on observational data often suffers from unobservable heterogeneity. It is important to note that this is likely to be less of a problem in our setting, because the dependent variable is derived from observations made by farmers while the characteristics included on the right hand side are collected from agro-input shop managers (which is different from the standard case where both dependent and independent variables are obtained from the same actors). For example, it is unlikely that an unobserved characteristic such as the motivation of the agro-input manager directly affects perceptions of farmers, unless this is reflected in the attribute that the farmer is assessing. That is why we also add control variables that differ depending on the attribute that is being rated. For example, when farmers are asked to rate agro-input dealers in terms of price competitiveness, it seems reasonable to include prices charged by these dealers as controls. Similarly, for perceptions related to the quality of seed sold, we are particularly interested in testing if the coefficient on the gender of the agro-input dealer changes after adjusting for various observable dealer characteristics that are directly related to quality, like the storage technology, the infrastructure such as a leak-proof roofing or insulation, and so forth. This way, we attempt to differentiate between situations where farmers perceive female-managed agro-input shops less favorably and situations where differences in ratings reflect real differences between male-

and female-managed shops.²⁰

Farmer-level characteristics could also confound the relationship between an agro-input dealer's gender and the rating that the farmer provides. For example, it may be that farmers who are better educated generally provide higher ratings. At the same time, imagine that better-educated farmers are more inclined to shop at male-managed agro-input dealerships. This would make it difficult to differentiate a gender equity effect from an effect arising from differences in farmer education. Fortunately, we often have instances where the same farmer rates both male- and female-managed agro-input shops. This allows us to exploit within-farmer variation for identification in equation (2.6). While we would be able to control for a farmer's education level by simply including it in an OLS regression, a within-farmer transformation also controls for characteristics that would be difficult or impossible to measure and to control for, like motivation, kindness, locus of control, norms, and values, and so forth. In other words, the within-farmer (fixed-effects) estimator removes all farmer-level heterogeneity.

$$y_{f,d} - \frac{1}{D} \sum_d y_{f,d} = \beta \left(g_{f,d} - \frac{1}{D} \sum_d g_{fd} \right) + \left(\varepsilon_{f,d} - \frac{1}{D} \sum_d \varepsilon_{f,d} \right) \quad (2.6)$$

$$y_{f,d} - \bar{y}_f = \gamma (g_{f,d} - \bar{g}_f) + \varepsilon_{f,d} \quad (2.7)$$

Finally, we also run a fixed-effects model that, in addition to controlling for farmer heterogeneity, also controls for dealer-level observable characteristics. We do so by again including additional regressors (x_d) in equation (2.7), which leads to:

$$y_{f,d} - \bar{y}_f = \gamma (g_{f,d} - \bar{g}_f) + \varphi (x_{f,d} - \bar{x}_f) + \varepsilon_{f,d} \quad (2.8)$$

²⁰Deciding which control variables to include is not always straightforward, especially for seed quality, which is difficult to verify objectively (which is why farmers are likely to rely on perceptions and potential gender differences in perceptions of agro-input shop managers are particularly alarming). The decision on what controls to use is based on quantitative and qualitative data of different stakeholders (including extension staff, agronomists, seed inspectors, and so forth) that was collected prior to the study.

2.5 Results

2.5.1 Between-dealers models

Tables 2.4 and 2.5 report perceived differences between male- and female-managed agro-input shops using an OLS regression based on equation (2.4). The difference between the two tables is that in the first table, farmers rate dealerships on a set of general characteristics like location and pricing, while in the second table, they rate maize seed, a particular product that these dealers sell, on various dimensions like germination and yield.

Looking at general dealership ratings in Table 2.4, we find that on all but one dimension, male-managed agro-input shops are rated higher than female-managed shops, and that the difference in ratings is significant for five out of the seven comparisons. We find a particularly large difference when farmers are asked to rate price competitiveness. Here, female-managed agro-input shops are scored only 3.24 out of 5, while male-managed agro-input shops receive a score of 3.44 out of 5. While these effects may seem small, it should be noted that ratings generally lie within a small range (of approximately 2.5 to 4). Thus, even a small coefficient estimate may reflect a considerable bias. Similar differences exist when farmers are asked to rate an agro-input dealer in terms of stock. Here, male-managed agro-input shops receive an average score of 3.99 out of 5, while female dealers get 3.79. Interestingly, female-managed agro-input shops are not rated significantly worse with regards to the quality of seed sold. They also appear to be equally rated with respect to location.

Table 2.5 repeats the between-dealers analysis but compares ratings for quality attributes of maize seed sold by these dealers. While we still find that on most characteristics, male-managed agro-input shops get a higher rating than female-managed shops, the differences are never significant. Note that these results are consistent with what was found in Table 2.4, where one of the few non-significant differences was related to perceived seed quality. This suggests that when farmers are asked to think about a particular product, they make abstraction of the person selling it, and the gender effect becomes less important.

Results of the between-dealers regressions with added control variables (see equation (2.5)) are presented in Table 2.6 for the more general ratings related to the dealership and in Table 2.7 for the more specific ratings related to seed quality. In all regressions, we add the age and education of the dealer as gen-

Table 2.4: Between-dealers model focusing on dealer ratings (control variables not included)

	<i>Dependent variable: Average rating received by dealer</i>						
	Average dealer rating	Dealer's general quality	Dealer's location	Dealer's price	Dealer's seed quality	Dealer's stock	Dealer's reputation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	3.764 (0.050)	3.738 (0.060)	3.890 (0.086)	3.237 (0.071)	3.878 (0.064)	3.792 (0.077)	4.051 (0.064)
Dealer is male	0.109* (0.063)	0.129* (0.075)	-0.085 (0.109)	0.200** (0.089)	0.052 (0.081)	0.199** (0.097)	0.161** (0.081)
Number of obs.	152	152	152	152	152	152	152

Note: The gender of the dealer is a dummy variable where 1 is male and 0 is female. ***, **, and * denote significance at the 1, 5 and 10% levels. Standard errors are clustered at the farmer level and presented in parentheses. Number of observations: The ratings by farmers given to dealers are averaged at the dealer level while also averaging across rating collection rounds. Max. 152 shops are included in this table as the gender of their manager did not change between rounds.

Table 2.5: Between-dealers model focusing on seed ratings (control variables not included)

<i>Dependent variable: Average rating received by dealer</i>							
	Average seed rating	Seed's general quality	Seed's yield	Seed's drought tolerance	Seed's pest/disease tolerance	Seed's speed of maturing	Seed's germination
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	3.417 (0.044)	3.870 (0.051)	3.592 (0.059)	3.005 (0.061)	2.683 (0.062)	3.716 (0.055)	3.651 (0.066)
Dealer is male	0.053 (0.056)	0.058 (0.064)	0.066 (0.075)	-0.030 (0.077)	0.011 (0.078)	0.081 (0.068)	0.104 (0.083)
Number of obs.	152	152	152	150	152	151	152

Note: The gender of the dealer is a dummy variable where 1 is male and 0 is female. ***, **, and * denote significance at the 1, 5 and 10% levels. Standard errors are clustered at the farmer level and presented in parentheses. Number of observations: The ratings by farmers given to dealers are averaged at the dealer level while also averaging across rating collection rounds. Max. 152 shops are included in this table as the gender of their manager did not change between rounds.

eral control variables as they are considered proxies for quality, and additional controls depending on the attribute being rated.

In column (1) of Table 2.6, we investigate the overall dealer rating, an average of the other attributes. We find that, even after controlling for a range of observable indicators of overall quality in this regression, male-managed shops are rated significantly higher by farmers.

Column (2) of Table 2.6 corresponds to column (2) of Table 2.4, which compares general quality ratings given to male- versus female-managed agro-input shops. We include three relatively objectively observable proxies for general dealership quality. First, we asked enumerators to provide an overall cleanliness and professionalism rating for the agro-input shop for which they collected data. Second, we construct an index that measures dealer effort and a range of services that dealers offer to clients. All indices were constructed by weighing each component by the inverse covariance matrix following Anderson (2008). In particular, this index accounts for whether an agro-input dealer 1) always explains how seed should be used (seed spacing, seed rate, complementary inputs); 2) always recommends complementary inputs such as fertilizers and chemicals; 3) provides extension or training on how to use improved seed varieties to clients; 4) provides discounts to clients who buy large quantities of seed; 5) sells small quantities; 6) provides seed on credit; 7) has received a seed-related complaint from a customer; and 8) accepts mobile money as a payment modality. Descriptive statistics for the variables which constitute this effort and service index are shown in the appendix in Table 2.12. Third, we asked enumerators to carefully observe and note down a range of capital-intensive seed handling and storage practices, which we also summarized in an index. In this index, we account for whether 1) the roof is leak-proof; 2) the roof is insulated to keep the heat out; 3) the walls are insulated to keep the heat out; 4) the area where seed is stored is properly ventilated; 5) any official certificates are on display in the shop (e.g., inspection certificates, training certificates, registration with an association, and so forth). Also for these variables which constitute the capital-intensive practices index, descriptive statistics are shown in the appendix in Table 2.12. We see that after controlling for these three groups of variables, the male premium on general quality ratings increases from 0.13 to 0.16. Note that the index of capital-intensive seed handling and storage practices observed by the enumerator is significant and has the expected sign, as input dealers who score better on this index also receive higher scores

on general dealership quality.²¹

When farmers were asked to assess agro-input dealers in terms of their location, the average distance between dealers and their customers,²² an indication of dealer centrality, provides an obvious candidate as a control variable (column (3) in Table 2.6). We do not find a gender effect on ratings concerning location in Table 2.4, nor do we find a difference after controlling for centrality. It should also be noted that the control variable is significant in the expected direction, as dealers for whom the average distance between dealer and customer is larger (or centrality is lower) also are scored lower in terms of location.

In column (4) of Table 2.6, we look at price competitiveness. To account for the possibility that the difference in price ratings between male- and female-managed agro-input shops is driven by actual price differences, we control for the average price the dealer charges for improved maize seed, as well as for the cost at which the dealer obtains seed, an important determinant of the price. The analysis confirms that there is a difference in perception of male and female dealers, and that this difference cannot be explained by actual price differences. The gender equity effect is larger than the one found without controlling for actual price differences in Table 2.4. Note again that one of the control variables is significant and suggests that dealers who charge higher prices also receive significantly lower price competitiveness ratings, as expected.

When investigating seed quality ratings in column (5) of Table 2.6, we control for another index, one that reflects *all* seed handling and storage practices observed by the enumerator. This index includes the five capital-intensive practices mentioned above, but also accounts for whether the agro-input dealer 1) destroys seed that has exceeded shelf-life; 2) stores seed in a dedicated area, away from other merchandise; 3) has no problem with rats, insects, or other infestations; 4) stores seed in ambient light conditions as recommended; 5) stores seed on pallets or shelves; and 6) does not store seed in open bags or containers. This index also includes the shop's overall cleanliness and professionalism rating provided by the enumerator. Descriptive statistics for the variables which

²¹However, caution should be taken when interpreting control variables, as they do not necessarily have a structural interpretation. For instance, it may be that the relationship between the control variable and the outcome variable is confounded by a third (potentially unobservable) variable (Hünemund and Louw, 2020).

²²The haversine formula calculating the arc distance between two points is used. The latitudes and longitudes are extracted from the GPS coordinates for both farmers and agro-input shops and inserted as paired values in the haversine formula. The formula then calculates the distances between these paired latitudes and longitudes.

constitute this (capital- and labor-intensive) practices index can be found in the appendix in Table 2.12. As in column (5) of Table 2.4, we do not find a gender effect regarding the seed quality rating after controlling for observable quality indicators.

In column (6), we repeat the analysis for perceptions related to a dealers' stock, now controlling for the number of hybrid maize varieties that this dealer has in stock and the quantity bought by the dealer from seed producers or wholesalers, the former being significant and having the expected sign. The male premium on the rating persists, although the effect becomes slightly weaker as compared to a regression without controls (column (6) in Table 2.4).

The analysis regarding the reputation rating is repeated in column (7), now controlling for the number of years the shop has been in business, and whether the shop is registered with the Uganda National Agro-input Dealer Association (UNADA), as we expect both to have an impact on a dealer's reputation. Here we also see that male dealers receive higher scores, but that the effect is slightly weaker after controlling for experience and UNADA registration (as compared to column (7) in Table 2.4).

Table 2.7 repeats the between-dealers analysis for quality attributes of maize seed sold by the agro-input shops as reported in Table 2.5, but controls for practices that are expected to improve seed quality. As all ratings in this table concern quality, we include the same controls in all regressions. We use the most elaborate index of all seed handling and storage practices as observed by the enumerator that was also used in model (5) of Table 2.6. Recall from Table 2.5 that we did not find a gender effect for any seed-quality-related dimension, and adding the index to control for quality does not change this. Note that the index is generally positively correlated with the ratings, but only significantly so when farmers are asked to assess the yield of seed that agro-input shops sell.

Overall, comparing Table 2.6 to Table 2.4 and Table 2.7 to Table 2.5, we notice that results, both in terms of parameter estimates for β and their significance, are very similar. This suggests that differences in ratings between male- and female-managed agro-input shops reflect differences in perception of the two genders, rather than actual differences in the dimension being rated (general quality, price competitiveness, stock, and reputation).

Table 2.6: Between-dealers model focusing on dealer ratings (control variables included)

	<i>Dependent variable: Average rating received by dealer</i>						
	Average dealer rating	Dealer's general quality	Dealer's location	Dealer's price	Dealer's seed quality	Dealer's stock	Dealer's reputation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	3.790 (0.134)	3.770 (0.227)	3.939 (0.185)	3.326 (0.169)	3.822 (0.150)	3.647 (0.204)	3.921 (0.156)
Dealer is male	0.157** (0.063)	0.161** (0.071)	-0.048 (0.100)	0.219** (0.091)	0.057 (0.081)	0.183* (0.097)	0.158* (0.082)
Dealer's age in years	-0.003 (0.003)	-0.003 (0.004)	-0.002 (0.005)	-0.003 (0.004)	-0.001 (0.004)	-0.001 (0.005)	0.000 (0.004)
Dealer finished secondary education	0.110 (0.078)	0.144 (0.088)	-0.001 (0.122)	-0.010 (0.108)	0.198** (0.098)	0.092 (0.122)	0.150 (0.100)
Shop's cleanliness/professionalism rating by enumerator		-0.005 (0.047)					
Index of dealer's efforts and services	0.181* (0.100)	0.152 (0.110)					
Index of capital-intensive seed handling/storage practices observed by enumerator		0.290*** (0.079)					
Standardized distance between farmer and shop	-0.089** (0.040)		-0.302*** (0.051)				
Standardized sales price of maize seed	-0.098* (0.052)			-0.176** (0.074)			
Standardized cost of maize seed for dealer	0.057 (0.052)			0.085 (0.072)			
Index of all seed handling/storage practices observed by enumerator	0.093 (0.103)				0.184 (0.118)		
Number of hybrid maize varieties in stock	-0.020 (0.030)					0.070* (0.038)	
Standardized amount of maize seed dealer bought	-0.013 (0.028)					0.062 (0.041)	
Number of years since shop's establishment	0.000 (0.006)						0.002 (0.007)
Shop's UNADA registration	0.151 (0.097)						0.113 (0.103)
Number of obs.	149	151	152	151	151	152	152

Note: The gender of the dealer is a dummy variable where 1 is male and 0 is female.

***, **, and * denote significance at the 1, 5 and 10% levels.

Standard errors are clustered at the farmer level and presented in parentheses.

Number of observations: The ratings by farmers given to dealers are averaged at the dealer level while also averaging across rating collection rounds. Max. 152 shops are included in this table as the gender of their manager did not change between rounds.

Table 2.7: Between-dealers model focusing on seed ratings (control variables included)

	<i>Dependent variable: Average rating received by dealer</i>						
	Average seed rating (1)	Seed's general quality (2)	Seed's yield (3)	Seed's drought tolerance (4)	Seed's pest/disease tolerance (5)	Seed's speed of maturing (6)	Seed's germination (7)
Constant	3.416 (0.105)	3.768 (0.120)	3.584 (0.139)	3.128 (0.147)	2.779 (0.147)	3.644 (0.131)	3.612 (0.157)
Dealer is male	0.053 (0.057)	0.047 (0.065)	0.066 (0.075)	-0.019 (0.079)	0.023 (0.079)	0.075 (0.070)	0.104 (0.085)
Dealer's age in years	-0.002 (0.003)	0.002 (0.003)	-0.001 (0.004)	-0.005 (0.004)	-0.005 (0.004)	0.002 (0.004)	-0.001 (0.004)
Dealer finished secondary education	0.125* (0.069)	0.122 (0.079)	0.118 (0.091)	0.058 (0.096)	0.188* (0.096)	0.052 (0.085)	0.175* (0.103)
Index of all seed handling/storage practices observed by enumerator	0.061 (0.083)	0.139 (0.095)	0.194* (0.110)	0.064 (0.115)	-0.089 (0.115)	0.083 (0.102)	-0.016 (0.124)
Number of obs.	151	151	151	149	151	150	151

Note: The gender of the dealer is a dummy variable where 1 is male and 0 is female. ***, **, and * denote significance at the 1, 5 and 10% levels. Standard errors are clustered at the farmer level and presented in parentheses. Number of observations: The ratings by farmers given to dealers are averaged at the dealer level while also averaging across rating collection rounds. Max. 152 shops are included in this table as the gender of their manager did not change between rounds.

2.5.2 Farmer fixed-effects models

In order to test whether the male bias persists after accounting for farmer-level heterogeneity, we exploit the fact that farmers generally rated more than one agro-input dealer. If the same farmer rates both male and female-managed agro-input shops, we can exploit this within-farmer variation and control for farmer specific observable and unobservable characteristics by including farmer fixed-effects.

Tables 2.8 and 2.9 show parameter estimates using a model that includes farmer fixed-effects, i.e., the within transformation of equation (2.6). As errors are also correlated within agro-input dealers, we report standard errors that are robust to clustering in this dimension. In Table 2.8, we use the general agro-input dealer ratings as outcome variables, similar to Table 2.4; Table 2.9 estimates the same model, but now for the more specific seed-quality-related ratings, similar to Table 2.5. In the previous subsection, the dependent variable was the *average* rating received by dealers, leading to a sample size of about 150 with one observation per rated dealer. In the following farmer fixed-effects analyses however, the dependent variable is the rating of a particular farmer given to a particular dealer. The number of observations now represents the total number of ratings given by all farmers to all dealers with one observation per farmer-dealer combination, leading to a much larger sample size.

Table 2.8 shows that male-managed agro-input outlets receive significantly higher ratings in the areas of general quality, price competitiveness, and reputation. The average dealer rating also significantly differs between male and female dealers. Comparing Table 2.8 to Table 2.4, the largest difference can still be found for price competitiveness, even though the magnitude of the effect decreased somewhat. The effect of gender on ratings related to stocks reduced sharply after controlling for farmer-level heterogeneity.

For seed-quality-specific ratings, comparing Table 2.9 to Table 2.5, we see that some of the differences between ratings of male and female dealers turn significant after controlling for farmer-level heterogeneity. For perceptions related to seed germination, male-managed agro-input shops receive a score that is on average 0.11 higher than the germination rating female-managed shops receive. The gender equity bias in this dimension is also reflected in a significant difference in the average seed rating between male- and female-managed agro-input shops in column (1).

Table 2.8: Farmer fixed-effects model focusing on dealer ratings (control variables not included)

<i>Dependent variable: Rating of a particular farmer given to a particular dealer</i>						
Average dealer rating	Dealer's general quality	Dealer's location	Dealer's price	Dealer's seed quality	Dealer's stock	Dealer's reputation
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dealer is male	0.116*** (0.041)	0.176** (0.071)	0.023 (0.068)	0.188*** (0.067)	0.049 (0.061)	0.180*** (0.067)
Number of obs.	1781	1781	1781	1781	1781	1781

Note: The gender of the dealer is a dummy variable where 1 is male and 0 is female.
***, **, and * denote significance at the 1, 5 and 10% levels.
Standard errors are clustered at the dealer level and presented in parentheses.
Number of observations: We first stack all ratings by farmers given to dealers of both rounds of collection, then exclude ratings of shops if the gender of their manager changed between rounds. E.g., for Dealer's location, we collected 837 ratings in the first round and 944 in the second round, leading to 1,781 observations included in this table, conveying the dyadic nature of the dataset.

Table 2.9: Farmer fixed-effects model focusing on seed ratings (control variables not included)

<i>Dependent variable: Rating of a particular farmer given to a particular dealer</i>						
Average seed rating	Seed's general quality	Seed's yield	Seed's drought tolerance	Seed's pest/disease tolerance	Seed's speed of maturing	Seed's germination
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dealer is male	0.073* (0.038)	0.075 (0.054)	0.047 (0.061)	0.053 (0.059)	0.041 (0.063)	0.079 (0.055)
Number of obs.	1760	1748	1721	1678	1692	1699
						1714

Note: The gender of the dealer is a dummy variable where 1 is male and 0 is female.
***, **, and * denote significance at the 1, 5 and 10% levels.
Standard errors are clustered at the dealer level and presented in parentheses.
Number of observations: We first stack all ratings by farmers given to dealers of both rounds of collection, then exclude ratings of shops if the gender of their manager changed between rounds. E.g., for Seed's yield, we collected 797 ratings in the first round and 924 in the second round, leading to 1,721 observations included in this table, conveying the dyadic nature of the dataset.

The fact that we do find gender equity bias when farmers are asked to assess seed quality if we control for farmer fixed-effects suggests that, in the between-dealers regressions of Tables 2.5 and 2.7, gender equity bias is obscured by farmer-level confounders. For instance, it could be that farmers that are higher educated also provide higher ratings and that these higher-educated farmers are also more likely to shop at female-managed dealerships. Not controlling for differences in education levels of farmers may then lead to an underestimation of discrimination against female-managed agro-input shops.

Finally, we run a fixed-effects model that, in addition to controlling for farmer heterogeneity, also controls for dealer-level observable characteristics (see equation (2.8)), similar to Tables 2.6 and 2.7. Table 2.10 presents the more general agro-input dealer ratings, and Table 2.11 presents the more specific seed ratings. We find that controlling for observable characteristics at the dealer level does not change the findings for the first set of ratings, which evaluate the dealership. The largest gender equity effects are found when farmers rate price competitiveness in column (4) and agro-input dealer reputation in column (7). In both cases, male-managed agro-input shops are rated about 0.22 points higher. The difference in ratings between male- and female-managed agro-input shops for the stock attribute has become indistinguishable from zero.

Finally, comparing Tables 2.9 and 2.11, the significant difference between male- and female-managed agro-input shops with respect to the average seed rating persists after controlling for observable dealer-level differences in seed quality. The difference in germination ratings in column (7) becomes insignificant but we now find a significant male premium for the general seed quality ratings in column (2).

2.6 Conclusion and policy implications

Using survey data from smallholder maize farmers and agro-input dealers in southeastern Uganda, we test if farmers perceive female-managed shops differently than male-managed shops. To do so, we asked farmers to rate agro-input dealers in their neighborhood on a scale ranging from one (poor) to five (excellent). Farmers rated dealers on a set of general characteristics such as accessibility and price competitiveness. They also rated maize seed, a particular product that these dealers sell, on various dimensions like germination, yield, and so forth.

Table 2.10: Farmer fixed-effects model focusing on dealer ratings (control variables included)

<i>Dependent variable: Rating of a particular farmer given to a particular dealer</i>						
Average dealer rating	Dealer's general quality	Dealer's location	Dealer's price	Dealer's seed quality	Dealer's stock	Dealer's reputation
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dealer is male	0.149*** (0.047)	0.195** (0.082)	0.216*** (0.072)	0.014 (0.067)	0.077 (0.072)	0.217*** (0.069)
Dealer's age in years	0.006*** (0.002)	0.005* (0.003)	0.005** (0.003)	0.007** (0.003)	0.009*** (0.002)	0.000 (0.002)
Dealer finished secondary education	0.013 (0.042)	0.089 (0.063)	-0.014 (0.062)	0.048 (0.063)	-0.059 (0.059)	-0.106* (0.055)
Shop's cleanliness/professionalism rating by enumerator	-0.013 (0.030)					
Index of dealer's efforts and services	0.123** (0.052)	0.143* (0.077)				
Index of capital-intensive seed handling/storage practices observed by enumerator		0.009 (0.058)				
Standardized distance between farmer and shop	-0.063** (0.027)	0.016 (0.048)				
Standardized sales price of maize seed	-0.064** (0.027)		-0.056 (0.041)			
Standardized cost of maize seed for dealer	0.090*** (0.028)		0.048 (0.041)			
Index of all seed handling/storage practices observed by enumerator	-0.070 (0.046)			0.004 (0.075)	0.035 (0.024)	
Number of hybrid maize varieties in stock	0.009 (0.015)				0.001 (0.022)	
Standardized amount of maize seed dealer bought	-0.009 (0.015)					0.017*** (0.005)
Number of years since shop's establishment	0.003 (0.003)					-0.030 (0.055)
Shop's UNADA registration	-0.074 (0.048)					
Number of obs.	1374	1496	1649	1541	1674	1706

Note: The gender of the dealer is a dummy variable where 1 is male and 0 is female.

***, **, and * denote significance at the 1, 5 and 10% levels.

Standard errors are clustered at the dealer level and presented in parentheses.

Number of observations: We first stack all ratings by farmers given to dealers of both rounds of collection, then exclude ratings of shops if the gender of their manager changed between rounds. E.g., for Dealer's location, we collected 837 ratings in the first round and 944 in the second round, but for 75 observations control variables were missing, leading to 1,706 observations being included in this table.

Table 2.11: Farmer fixed-effects model focusing on seed ratings (control variables included)

Dependent variable: Rating of a particular farmer given to a particular dealer							
Average seed rating	Seed's general quality	Seed's yield	Seed's drought tolerance	Seed's pest/disease tolerance	Seed's speed of maturing	Seed's germination	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Dealer is male	0.082* (0.041)	0.099* (0.059)	0.100 (0.065)	0.019 (0.062)	0.060 (0.070)	0.071 (0.058)	0.096 (0.062)
Dealer's age in years	0.001 (0.001)	0.006*** (0.002)	0.003 (0.002)	-0.002 (0.002)	0.004* (0.002)	-0.002 (0.002)	0.003 (0.002)
Dealer finished secondary education	-0.029 (0.035)	-0.045 (0.048)	-0.024 (0.056)	0.038 (0.054)	0.090 (0.058)	-0.124*** (0.047)	-0.025 (0.053)
Index of all seed handling/storage practices observed by enumerator	0.120*** (0.042)	0.139** (0.060)	0.196*** (0.067)	0.093 (0.061)	0.138** (0.065)	0.087 (0.055)	0.061 (0.060)
Number of obs.	1520	1509	1485	1447	1460	1467	1480

Note: The gender of the dealer is a dummy variable where 1 is male and 0 is female.
 ***, **, and * denote significance at the 1, 5 and 10% levels.

Standard errors are clustered at the dealer level and presented in parentheses.

Number of observations: We first stack all ratings by farmers given to dealers of both rounds of collection, then exclude ratings of shops if the gender of their manager changed between rounds. E.g., for Seed's yield, we collected 797 ratings in the first round and 924 in the second round, but for 236 observations control variables were missing, leading to 1,485 observations being included in this table.

Simply comparing average ratings given to male- and female-managed agro-input shops, we find that shops managed by women are generally rated lower than their male-managed competitors. However, when farmers were asked to focus on a specific product, the difference was insignificant. After adding controls for agro-input dealer-level observable characteristics, parameter estimates and significance remain similar, suggesting that differences in ratings between male- and female-managed agro-input shops reflect differences in perceptions rather than actual differences in the attributes being rated.

Furthermore, ratings of agro-input dealers provided by farmers may also be influenced by farmer characteristics. To control for farmer heterogeneity, we exploit the fact that farmers often rated several agro-input dealers of different genders and ran farmer fixed-effects models. Doing so, we confirm the existence of gender equity bias when farmers were asked to rate general characteristics of agro-input dealers, but also find differences in ratings of different dimensions of seed quality sold by dealers of different genders.

Looking into the individual dimensions that were rated, we find particularly strong gender equity bias when farmers rated agro-input dealers in terms of price competitiveness. Furthermore, and especially after controlling for farmer-level heterogeneity, we find that male-managed agro-input shops to have a significantly better reputation than female-managed shops. This contrast in reputation is also reflected in a significant difference between male and female dealers in the general quality rating. On the other hand, we do not find that male- and female-managed agro-input shops were rated differently when farmers were asked to consider location. This may be because location is easier to assess objectively. For attributes related to the quality of seed sold by agro-input dealers, gender bias was only persistently found for the average seed rating after controlling for farmer-level heterogeneity.

Despite this gender bias favoring men, 40% of Ugandan agro-input shops are managed by women, raising the question how this surprisingly large share of female dealers remains in the market. We see that they are on average younger, less educated, and less trained than their male competitors, and that their shops have been established more recently, but perhaps these women have another competitive advantage at their disposal that attracts at least some customers. It cannot be better prices as we saw that farmers discriminate most on this dimension. Seeing that location is the only general dealership characteristic for which female-managed shops receive better ratings than male-managed shops

(see Table 2.4), we investigate whether these shops are located better. However, apparently these shops are further way from roads and customers, implying that location is not the competitive advantage that keeps female dealers in the market. It is however possible that there is less competition in these more remote areas, which could explain why even discriminated dealers are not driven out of the market. Alternatively, shops managed by women may provide better products than their male competitors. We see for example that female-managed shops are more likely to be specialized stores that only sell farm inputs and that they are cleaner, according to enumerators. These shops also have better roofing, walls, and ventilation, and are more likely to store seed on pallets or shelves, instead of in open containers. All this could lead to woman selling better seed, as Barriga and Fiala (2020) document how handling and storage practices affect seed quality. Open air storage of bags can lower the quality of seeds (Bold et al., 2017), temperature control after the seed leaves the breeders is crucial (Barriga and Fiala, 2020), and storage in moist conditions or in direct sunlight further reduce seed quality (Govender, Aveling, and Kritzing, 2008; Curzi, Nota, and Di Falco, 2022). In line with this, we see that shops managed by women receive less seed-related complaints from customers and appear more professional: on average, they are more likely to display official certificates, to be registered with UNADA, and to have a trading license from the local government. If customers realize that these female dealers sell better seed, this will result in a comparative advantage and explain why 40% of Ugandan agro-input shops are managed by women, even though perceptions are stacked against them. An alternative but complementary explanation is that many women simply have no other opportunity to earn money. While men may compare their earnings through agro-input dealing with a lucrative outside option, women may have to compare it with earning nothing. This would explain why they do not leave the market, even though they earn significantly less than their male competitors, pointing to a more structural bias in the economy.

Finding that farmers are discriminating against shops managed by women is troubling for a variety of reasons. Their biased perceptions can influence real purchase decisions which may have long-run implications for agro-input shops. Table 2.1 indicates that an average female-managed agro-input shop in our sample receives only about 36 customers per day while an average male-managed agro-input shop receives about 51 customers per day. The amount

of maize seed sold and the revenue from these sales earned by an average female-managed shop are also lower than the amount sold and revenue of an average male-managed agro-input shop. Farmers' biased perceptions can be particularly damaging in traditional agricultural societies with strong norms and customs. In these societies, women's opportunities are already severely restricted, and gender bias may further restrain women from entering productive activities. This will in turn reinforce gender stereotypes and the view that women are less able to perform particular tasks.

However, gender equity bias does not only directly impact women's capabilities, aspirations, and their empowerment in agri-food systems (Jayachandran, 2021). It is also likely to affect future generations, as women tend to invest more of their income than men in healthcare, nutrition, and education of their children (Thomas, 1990). But there are consequences that go beyond the household. Almost half of the agro-input shops in our sample are managed by women. If farmers do not trust these shops, this may pose challenges for varietal turnover at more aggregate levels: in a village where only women manage shops, farmers may be less likely to buy commercial improved seed from the market, and instead use farmer-saved seed obtained through informal channels, hindering agricultural productivity and rural transformation.

Our finding has important implications. It underscores the importance of customs and norms in rural and more traditional societies. Interventions and initiatives that focus solely on increasing women's empowerment are unlikely to be sufficient and may in some cases even backfire (Ntakyo and Van Den Berg, 2022). It will be important to challenge gender stereotypes and role congruence and such interventions should not focus on only one gender.

Our study serves to draw specific lessons for policy. Over the years, policymakers have encouraged women to enter business domains which were traditionally dominated by men, with women striving towards new opportunities and ways to earn for their livelihoods or families. However, our findings show the need for policies addressing the lack of acceptance or integration that still prevails. If these biased perceptions cannot be corrected, we may see a withdrawal of many women from these sub-sectors in the future as they become increasingly aware about the difficulties and the higher likelihood of restricted growth. We restrict ourselves to three areas where we see scope for policy action.

First, even though we do not find evidence of male-managed agro-input

shops actually providing better quality than female-managed shops, existing training and advisory services for agro-input dealers are also likely to be biased toward men, and this may indirectly influence perceptions related to the abilities of female managers. Ensuring that women entrepreneurs have access to training should be a policy priority. The effectiveness and inclusiveness of training programs depend on many attributes of the program. This includes more obvious aspects such as the training content and who is targeted, but also less obvious attributes such as the gender of who provides the training, the timing of training, and so forth (Lecoutere, Spielman, and Van Campenhout, 2023). At the same time, it is also important to change the perception that female-managed agro-input shops are likely to receive less training. This could be achieved by making training attendance publicly visible, perhaps through a register of trained agro-input dealers, through certificates that are displayed in the shops, and so forth, such that equal capacity between male- and female-managed agro-input shops becomes more apparent to clients.

Second, female role models have been shown effective in increasing female participation in a variety of otherwise male-dominated sectors (Porter and Serra, 2020; Riley, 2022). Considering this, perceptions may evolve in line with the presence of women among agro-input dealers, inspectors, extension providers, and leaders of professional associations such as UNADA. This will not only motivate more women to enter the market but also bring forth wider acceptance across the value chain and in agricultural markets. For public sector positions, quotas may be considered, since research suggests that they can be an effective way to challenge gender stereotypes held by men (Beaman et al., 2009).

Finally, we find that biased perceptions exist especially with respect to prices charged by female agro-input dealers. Simply advertising prices may be sufficient to make them objectively verifiable, and customers will need to depend less on perceptions and the use of mental shortcuts that are prone to gender equity bias.

2.7 Acknowledgments

This research was partly funded by the Consultative Group for International Agricultural Research (CGIAR) Research Program on Policies, Institutions, and Markets, led by the International Food Policy Research Institute (IFPRI) and carried out with support from the CGIAR Fund contributors. The authors also acknowledge financial support from The Netherlands – CGIAR research programme on Seed Systems Development, grant number W08.240.105, funded by the Netherlands Organisation for Scientific Research (NWO-WOTRO), and from the Belgian Public Service for Foreign Affairs, Foreign Trade and Development Cooperation, grant number 302015.002.002 (bridge funding). Acknowledgment is extended to the Fonds Wetenschappelijk Onderzoek – Vlaanderen (FWO), Fonds de la Recherche Scientifique – FNRS (EOS project G0G4318N) and the KU Leuven Methusalem program. We thank Leocardia Nabwire, Marc Charles Wanume, Richard Ariong and Wilberforce Walukano for excellent field support. We also thank Gbenga Akinwehinmi, Gero Laurenz Höhn, Hannah Ameye and Hyejin Lee for their feedback. The analysis is the sole responsibility of the authors and does not reflect the views of any funding agency or organization mentioned here.

2.8 Appendix

Table 2.12: Descriptive agro-input dealer statistics: Variables in indices

	Male					Female				
	mean	min	max	SD	obs.	mean	min	max	SD	obs.
Capital-intensive seed handling/storage practices observed by enumerator										
Shop has leak-proof roof	0.535	0	1	0.501	114	0.544	0	1	0.501	79
Shop has insulated roof	0.579	0	1	0.496	114	0.671	0	1	0.473	79
Shop has insulated walls	0.807	0	1	0.396	114	0.823	0	1	0.384	79
Shop is ventilated	0.833	0	1	0.374	114	0.835	0	1	0.373	79
Shop displays official certificate	0.500	0	1	0.502	114	0.506	0	1	0.503	79
Shop always handles expired seed correctly	0.935	0	1	0.247	108	0.948	0	1	0.223	77
Labor-intensive seed handling/storage practices observed by enumerator										
Shop stores seed away from other products	0.447	0	1	0.499	114	0.342	0	1	0.477	79
Shop has problem with pests	0.623	0	1	0.487	114	0.658	0	1	0.477	79
Shop's light is ambient (not direct sunlight/dark)	0.842	0	1	0.366	114	0.785	0	1	0.414	79
Shop stores seed on pallets/shelves (not directly on wood/floor/cardboard)	0.642	0	1	0.482	109	0.776	0	1	0.419	76
Shop stores maize seed in open containers	0.211	0	1	0.409	114	0.165	0	1	0.373	79
Shop's cleanliness/professionalism rating by enumerator	3.465	1	5	1.191	114	3.557	1	5	1.071	79
Dealer's efforts and services										
Shop always explains to customers how seed should be used	0.561	0	1	0.498	114	0.342	0	1	0.477	79
Shop always recommends complementary inputs to customers	0.614	0	1	0.489	114	0.456	0	1	0.501	79
Shop offers extension/training	0.544	0	1	0.500	114	0.494	0	1	0.503	79
Shop offers discounts for large quantities	0.789	0	1	0.409	114	0.747	0	1	0.438	79
Shop's smallest seed bag is 1 kg (not larger)	0.700	0	1	0.460	110	0.770	0	1	0.424	74
Shop provides seed on credit	0.632	0	1	0.485	114	0.671	0	1	0.473	79
Shop received seed related complaint from customer	0.711	0	1	0.456	114	0.620	0	1	0.488	79
Shop accepts mobile money as payment	0.395	0	1	0.491	114	0.405	0	1	0.494	79

Note: SD is the standard deviation.
Number of observations: All 193 agro-input dealers are included in this table.

Miracle seeds: Biased expectations, complementary input use, and the dynamics of smallholder technology adoption

This chapter is co-authored with Leocardia Nabwire (IFPRI, Uganda), Robert Sparrow (Wageningen University and Erasmus University Rotterdam, Netherlands), David Spielman (IFPRI, United States), and Bjorn Van Campenhout (IFPRI and KU Leuven, Belgium). An earlier version of this chapter is published as IFPRI Discussion Paper 2183.

3.1 Introduction

To feed a growing population in an environmentally sustainable manner and in the midst of a long-term climate crisis, farmers throughout the developing world are expected to grow more food on less land with greater efficiency (Tilman et al., 2011; Garnett et al., 2013). To achieve this goal, much is expected from new technologies, especially from higher-yielding varieties that are resilient to pests, diseases, and other biotic stresses and are tolerant of droughts, floods, heat, and other abiotic stresses (Evenson and Gollin, 2003; Lybbert and Sum-

ner, 2012).

Unfortunately, the adoption of such technologies is lagging in areas where they may have the largest impact. Recent trends in agricultural productivity growth in Africa show that technological progress has largely stagnated on the continent (Suri and Udry, 2022). However, significant heterogeneity underlies this general stagnation. For instance, at the micro level, we often observe dis-adoption patterns and trends, where farmers choose to switch back to technologies and inputs they have been using for decades after trying out a new technology once or twice (Moser and Barrett, 2006; Chen, Hu, and Myers, 2022). In many cases, these patterns and trends cannot be explained by a lack of awareness or information about, for example, improved cultivars or inorganic fertilizers (Sheahan and Barrett, 2017).²³

There are many reasons why farmers may not move into a state of sustained adoption of a given technology. An obvious one is that farmers cannot access the technology through local markets or other means of supply, or may have enjoyed access only for a limited time as part of a promotional campaign or project intervention (Shiferaw et al., 2015). Another reason may be that farmers learn over time that a particular technology is not suitable for them or does not meet their expectations (Custodio et al., 2016). Heterogeneity in the quality of the technology, coupled with the fact that it is often difficult to assess quality prior to purchase or application, may also result in dis-adoption (Bold et al., 2017; Mieke et al., 2023). Farmers that face credit or liquidity constraints, or additional uninsured risk may also reconsider past adoption behavior and tend towards dis-adoption (Karlan et al., 2014). In the longer run, general equilibrium effects that accrue as more farmers adopt a new technology, thereby increasing supply of the commodity and reducing output prices, may also lead farmers with higher marginal costs to exit the market and dis-adopt (Cochrane, 1958).

In this paper, we consider the possibility that farmers hold inflated expectations of new technologies as an explanation for their dis-adoption. These

²³For simplicity, we use the term “technologies” to refer to agricultural technologies such as improved varieties, which are genetic innovations embodied in seed. We use the term “inputs” to refer to organic and inorganic fertilizers and pesticides, and we use the term “practices” to refer to labor and management effort such as precision planting, irrigation, and weeding. Of course, we recognize that these terms can be used interchangeably—seed is also an input, while fertilizers and precision practices can also be technologies—and that each figures differently into our understanding of the conventional agricultural production function.

inflated expectations result from the possibility that farmers may be unaware (or fail to recognize) the need for substantial complementary investment. Indeed, for the new hybrid seeds suitable for East African maize farmers that came on the market a few decades ago, the promise to double or even triple yields could typically only be achieved in favorable climatic conditions and with the addition of fertilizer and other inputs (Quiñones, Borlaug, and Dowsell, 1997). Chen, Hu, and Myers (2022) show that farming with improved maize varieties is far more costly than farming with unimproved maize varieties. The additional production costs include not just the (higher) cost of seed but also higher fertilizer costs required to achieve expected yield improvements, as well as higher costs of labor for farm tasks that are associated with the cultivation of higher-yielding maize.

Inflated expectations about technology performance can have lasting impacts on adoption if farmers attribute poor outcomes to the technology, instead of to insufficient complementary inputs and effort. This learning failure is often understandable: if multiple factors simultaneously affect yields and outputs, then learning about the causal impact of a new technology from a single experience is difficult, especially if the technology performs only under specific or stochastic circumstances such as abiotic stress (Lybbert and Bell, 2010), or if the farmer is unable to learn in a Bayesian manner because it is too cognitively taxing (Gars and Ward, 2019), pays attention to the wrong attributes of the technology (Hanna, Mullainathan, and Schwartzstein, 2014), or is unable to sufficiently complement own experience with social learning (Foster and Rosenzweig, 1995; Conley and Udry, 2010).

This paper was motivated by evidence suggesting that many farmers are unaware that agricultural technologies such as improved varieties require substantial complementary inputs and efforts to reap benefits. Indeed, it is theoretically possible and quite reasonable to believe that farmers overestimate the returns to a technology and are disappointed when they compare realized yields with what they expected at the time of planting. Because it is hard for farmers to learn about the yield response of a single input, farmers may decide that the technology itself is to blame. This is consistent with the observation that farmers think inputs are often counterfeit or of low quality, even when objective assessments of input quality find them to be acceptable (Barriga and Fiala, 2020; Michelson et al., 2021).

Many researchers working in developing-country agriculture will have their

own anecdotal evidence of inflated expectations that illustrate the presence of biased expectations, sub-optimal complementary investments, and subsequent dis-adoption when disappointing outcomes are attributed to the technology itself. For instance, researchers may be familiar with farmers' belief that using inorganic fertilizer for one cropping cycle will lead to long-lasting soil fertility improvements. Others may be familiar with another common belief among farmers—often promoted by extension agents and agro-dealers—that an improved variety is a “miracle seed” that can be planted without additional inputs or management to achieve exceptional harvests. Entire narratives—some with more nuance than others—have been written on the singular power of genetic improvement, from the semi-dwarf “Green Revolution” varieties of wheat and rice to genetically modified crops (Lipton and Longhurst, 1989; Tripp, 2002; Sumberg, Keeney, and Dempsey, 2012). An example of the learning failure and its consequences was provided by extension agents we worked with. Seed of improved maize varieties needs a lot of nutrients, often leaving soil more depleted than when farmer-saved seed is used. In the areas where our research is situated, Striga (*Striga hermonthica*), a parasitic weed that feeds on the roots of maize plants and causes stunted growth, is a serious problem. Unfortunately, Striga proliferates in poor soils and as a result some farmers now believe that improved seed varieties are responsible for increased Striga infestations on their fields.

To develop our theory of inflated expectations, we present a simple model of technology adoption that incorporates the ideas discussed above. In this model, farmers compare the expected returns of an improved technology to their business-as-usual choices. The new technology comes at a cost, while the unimproved technology does not. Both technologies, though, require complementary inputs and efforts that directly affect productivity, with productivity gains from the new technology only materializing when complementary inputs and practices exceed business-as-usual levels for the unimproved technology. Further, recognizing that farmers may be heterogeneous, we define several farmer types and derive predictions about how they might behave if they learn about the true shape of the production function of the new technology.

We test our model's predictions using a field experiment conducted with

almost 3,500 maize farmers in eastern Uganda.²⁴ At the heart of the field experiment is a light-touch information intervention that highlights the importance of complementary investments when using improved maize varieties.²⁵ Specifically, we show all farmers in our sample a short, engaging video about the use of improved inputs and recommended management practices for maize cultivation. In the treatment group, we show the same video, except that in certain points in the narration—for instance, when the use of inorganic fertilizers is demonstrated or when weeding is explained—we highlight the particular importance of using additional inputs and performing certain management practices in conjunction with the improved variety.

We begin by testing whether farmers are able to extract the relevant information from the treatment video. We see that all coefficients move in the expected direction, and find treatment effects that differ significantly from zero for a subset of farmers. Turning to adoption behavior, we find evidence of treated farmers dis-adopting between baseline and follow-up. An exploratory heterogeneity analysis reveals that mainly more remote households dis-adopt due to the treatment. We find no evidence that the intervention affected the use of complementary inputs such as fertilizers and pesticides, or recommended practices for maize management such as row planting and intensive weeding. We also see that among treated farmers, expectations become more in line with realized output.

These findings have implications for our understanding of smallholder technology adoption dynamics. If farmers do not use appropriate complementary inputs and practices when using improved maize varieties because they believe in “miracle seeds”, their yields are likely to be disappointing. Often, disappointment about the performance of a technology is then erroneously attributed to the technology itself, potentially leading to dis-adoption. “Correcting” incorrect beliefs about the needed inputs and efforts may result in farmers dis-adopting technologies in the short run. However, if farmers’ expectations become more

²⁴The overarching study was pre-registered at the AEA RCT registry under RCT ID 0006361. It was primarily designed to examine quality-related constraints to technology adoption with a series of interventions at the agro-input dealer level. This paper makes use of farmer-level interventions that were introduced alongside the main design and described in the pre-analysis plan.

²⁵We use the term “improved variety” throughout this paper to refer to both maize hybrids and open pollinated varieties marketed and sold in our study areas, as opposed to farmer-saved seed or seed obtained through farmer-to-farmer exchanges which, in the specific context of maize, may be less effective due to cross pollination and genetic drift over multiple generations, or due to poor seed storage and handling between seasons.

realistic, the ones that keep adopting (or start adopting in light of the new information) will be less likely to be disappointed in the future, leading to sustained adoption, which in turn could lead to efficiency gains and positive spillover effects. Our findings also imply that public and private actors in the agriculture sector need to promote new technologies as highly site- and context-specific combinations of technologies, inputs, practices, and efforts instead of single “miracle seeds.”

3.2 Related literature

Agricultural technology adoption is at the heart of a rich body of research on food security, poverty reduction, economic development, and structural transformation. Studies on the economics of technical change in agriculture go back to at least Griliches (1957) and are reviewed in widely cited articles such as Feder, Just, and Zilberman (1985) and Sunding and Zilberman (2001). More recently and with the proliferation of field experiments and randomized controlled trials, economic theories that explore alternative drivers of technology adoption have received greater empirical attention.

Most of these studies (implicitly) assume that some kind of graduation model underlies the technical change process, wherein farmers switch from a low-level equilibrium to a high-level equilibrium in which technology use is sustained once initial conditions—typically, access to information or finance—are satisfied or binding constraints removed (Karlan et al., 2014; Shiferaw et al., 2015; Abate et al., 2016). Yet most of these studies follow farmers across a limited number of agricultural seasons, and are unable to fully appreciate the dynamics of technology adoption over time. Only a few studies offer a long-term perspective, with several documenting significant levels of dis-adoption (for example, Ainembabazi and Mugisha, 2014), or transient technology use over time (Moser and Barrett, 2006; Chen, Hu, and Myers, 2022).

At the core of our theoretical framework described in Section 3.3 is a model of learning failures where farmers have inflated expectations about the returns to a new technology, but fail to uncover the true form of the production function through experience, leading to disappointment and subsequent dis-adoption. Indeed, heterogeneity in farmer characteristics implies that farmers need to learn whether using a new technology is optimal for their specific context given costs and benefits (Suri, 2011). Farmers learn through a combination of own

experiences and observing others (Foster and Rosenzweig, 1995; Conley and Udry, 2010). However, learning about a new technology is often difficult for reasons related to the technology’s complexity and the observability of its quality or performance (i.e., its experience good nature) (Lybbert and Bell, 2010; Bold et al., 2017; Ashour et al., 2019), or the social, psychological, and behavioral attributes of the farmer and her learning process (Foster and Rosenzweig, 1995; Hanna, Mullainathan, and Schwartzstein, 2014).

One strand of the literature argues that sequential adoption leads to experiential learning by farmers. In cases where technologies are bundled in packages, it is often observed that farmers sequentially adopt components of the package, rather than adopting the entire package at once (for example, Byerlee and De Polanco, 1986). Leathers and Smale (1991) argue that this occurs when farmers employ a Bayesian approach to learning in which they try to isolate the impact of one component of the package at a time. However, there are circumstances under which this strategy is not optimal because it can prevent farmers from identifying potential synergies between technologies, inputs, and practices. Indeed, the reason why many interventions are presented as a package is because these interaction effects are not trivial. For instance, Kabunga, Dubois, and Qaim (2012) find that banana tissue culture, a technology to ensure that banana plantlets are free from pests and diseases, leads to a 7% yield gain in Kenya. However, they also find that improving access to irrigation can lift yield gains above 20%. If many complementarities like this exist, it seems unlikely that farmers are in a position to follow a sequential learning path that allows for all possible interactions between the different technologies within a reasonable time frame. Furthermore, as mentioned above, farmers may face certain behavioral constraints that inhibit their ability to learn about interaction effects if, for example, they pay attention to minor or tangential attributes of the package and miss the more important attributes (Hanna, Mullainathan, and Schwartzstein, 2014). Our study contributes to this literature by providing additional evidence on the limits of Bayesian learning in the context of agricultural technology adoption.

Another strand of the literature addresses the technology learning process in terms of how farmers compare realized yields against expected yields to inform their subsequent, longer-term adoption decisions. The effect of incorrect expectations about future returns on decision-making has been studied most in the context of education, but is readily applicable to learning in agriculture.

For example, both Nguyen (2008) and Jensen (2010) find that providing accurate information about the returns to education significantly increases investment in schooling (in Madagascar and the Dominican Republic respectively). Van Campenhout (2021) finds that a video intervention that informs Ugandan farmers about the returns on intensification investments in rice growing improved practices and increased input use and production. Note that across these studies, it is assumed that the individual underestimates the returns in question. In our study, as a result of under-investment in complementary inputs, farmers are in a sense overestimating returns to a new technology, leading to over-investment in technologies.

Finally, the intervention we use to test our hypothesis builds on a strand of the literature that focuses on the role of video-mediated messaging to convey salient information to farmers. This literature explores the ways in which informational videos can change behavior in a variety of settings and through a range of mechanisms. Ferrara, Chong, and Duryea (2012) show how telenovelas have an impact on fertility in Brazil. Riley (2022) finds that in Uganda, students that watched a Disney feel-good movie called “Queen of Katwe” about a chess prodigy growing up in the slums of Kampala do better on their exams, particularly in math. In the context of agricultural technology adoption, Van Campenhout, Spielman, and Lecoutere (2021) show that farmers that were exposed to videos similar to those we use in the present study perform significantly better on a knowledge test, and are more likely to apply recommended practices and fertilizers than households that did not view the video. These same households also report maize yields 10.5% higher than the control group. In Ethiopia, Abate et al. (2023) assess the impacts of video-mediated agricultural extension service provision on farmers’ adoption of improved agricultural technologies and practices using data from a two-year randomized experiment. Our study uses a light touch intervention where treatment and control videos are very similar, except for one small piece of information. Our study thus contributes to this literature by testing if videos are also effective in conveying subtle messages.

3.3 Theoretical framework

In our theoretical framework, we describe farmers as solving an intertemporal problem in which they allocate resources at t in order to maximize profits at

$t+1$.²⁶ In line with Suri (2011), we assume that farmers (indexed i in the model below) are risk-neutral and choose to plant seed which is either of a Variety H, a new variety that is stochastically dominant in yield and other attributes in all states, or of a Variety L, an old variety that is stochastically inferior in yield and other attributes in all states,²⁷ to maximize their profits per area of land. In doing so, they compare the expected profit functions of Variety H π_{it}^{*H} and Variety L π_{it}^{*L} which are defined as:

$$E(\pi_{it+1}^H) = E(p_{t+1}Y_{it+1}^H) - b_t s_{it} - \sum w_t X_{it}^H \quad (3.1)$$

$$E(\pi_{it+1}^L) = E(p_{t+1}Y_{it+1}^L) - \sum w_t X_{it}^L \quad (3.2)$$

where E is an expectations operator and $E(p_{t+1})$ is the expected price at which output is valued, assuming that the end commodity, maize grain, is indistinguishable to consumers by variety.²⁸ $E(Y_{it+1}^H)$ and $E(Y_{it+1}^L)$ reflect the expected yield for seed of Variety H and L respectively. Seed of Variety L is assumed to be free, while for seed of Variety H, s_{it} is procured at a cost $b_t > 0$.²⁹ In both profit functions, the cost of a range of complementary inputs and management practices, referred to as inputs, are deducted and summarized by the vector X_{it} with corresponding factor prices w_t .

Farmers adopt the stochastically dominant Variety H if they expect it to be more profitable than using the stochastically inferior Variety L, that is, if $E(\pi_{it+1}^H) > E(\pi_{it+1}^L)$ or:

²⁶For simplicity, we assume a discount factor of one, but another discount factor will not alter the results.

²⁷The model is applicable to a variety of cases as Variety H and Variety L can be interpreted as improved and unimproved, farmer-saved and commercially-purchased, modern and traditional, newer and older, hybrid and open pollinated varieties, and so forth.

²⁸In a country like Uganda, where most grain is combined, milled, and sold without varietal denomination, this is a reasonable assumption. In other countries such as Malawi or Mexico, where consumers have distinct varietal preferences related to taste, texture, and color, this assumption might not always hold.

²⁹Seed of Variety L may not be free but have a shadow price of at least the grain price, which could be subtracted from the expected revenue in equation (3.2), so that the adoption decision in equation (3.3) would not only depend on yield comparisons but also on cost comparisons. Suri (2011) takes this into account but also notes that the cost of, in her case, farmer-saved seed is likely to be low, if not zero. Rather than complicating the model by explicitly modeling the price of the stochastically inferior variety, we decide to set it to zero. Setting it to a small positive value would not change the predictions derived from the model.

$$\left(E(Y_{it+1}^H) - \sum \frac{w_t}{E(p_{t+1})} X_{it}^H \right) - \left(E(Y_{it+1}^L) - \sum \frac{w_t}{E(p_{t+1})} X_{it}^L \right) > \frac{b_t}{E(p_{t+1})} s_{it}^* \quad (3.3)$$

where we normalize by output price.³⁰

Equation (3.3) shows that adoption decisions fundamentally depend on yield comparisons. We assume that yield for Variety L is a function of inputs used:

$$Y_{it+1}^L = Y_{it}(X_{it}^L) \quad (3.4)$$

and that this relationship is assumed to be positive with decreasing returns to scale: $\frac{dY_{it}}{dX_{it}} > 0$ and $\frac{d^2Y_{it}}{dX_{it}^2} < 0$.

Yield for Variety H follows the same function, but adds a positive and constant adoption premium ($A > 0$). However, the adoption premium only applies when the farmer uses at least the same amount of complementary inputs as they would when using Variety L ($X_{it}^H \geq X_{it}^L$):

$$Y_{it+1}^H = A(X_{it}^H \geq X_{it}^L) + Y_{it}(X_{it}^H) \quad (3.5)$$

If farmers are able to predict yields—at least on average—in $t+1$, such that $E(Y_{it+1}) = Y_{it+1} + \varepsilon$ and $\varepsilon \sim N(0, \sigma)$, their decision to adopt would depend on the difference in yield between Variety H and L, on the relative prices of inputs, and on the yield responses to the inputs.

Next, we introduce farmer heterogeneity into the model by assuming that at least some farmers are not aware of the true relationship between Y_{it}^H and X_{it} , but instead believe that the adoption premium is always present, that is $E(Y_{it+1}^H) = A + Y_{it}(X_{it}^H)$. As a result, some farmer will use Variety H but not enough complementary inputs, leading to disappointing outcomes.

This variation in the model leads to different farmer-types based on their dynamic profile and knowledge, as summarized in Table 3.1. Some farmers are knowledgeable about the true relationship between Y_{it}^H and X_{it} in equation (3.5), and as a result make correct investment choices. For at least some of these farmers, referred to as Type 1 farmers in Table 3.1, the marginal cost of adoption will be lower than the expected marginal return in equation (3.3),

³⁰For simplicity, we assume that farmers have only one plot and model the decision to adopt as a binary process, instead of expressing s_{it} in kilograms of seed used. As such, b_t refers to the cost of planting an entire plot with seed of Variety H.

and as a result they adopt (and will continue to do so in the future unless there is a change in fundamentals such as the cost of seed). For another subset of these farmers, referred to as Type 2 farmers in Table 3.1, the marginal cost of adoption will be higher than the expected marginal return, so they will not adopt (and are unlikely to adopt in the near future).

Another group of farmers is not knowledgeable about the true relationship between Y_{it}^H and X_{it} and believes there is always an adoption premium. A subset of these farmers may adopt because their marginal cost of adoption is lower than their expected marginal return. We refer to these farmers in Table 3.1 as Type 3 farmers. Another subset of this group of farmers that is not knowledgeable about the true relationship between Y_{it}^H and X_{it} , referred to as Type 4 farmers in Table 3.1, does not adopt at baseline because, even though they have inflated expectations of Variety H's yield, the marginal cost of adoption still exceeds the expected marginal return.

Another group of farmers is also not knowledgeable about the true relationship between Y_{it}^H and X_{it} . But unlike Type 3 and 4 farmers, they underestimate the adoption premium (much like the rice farmers underestimate the returns to intensification investments in Van Campenhout (2021)), perhaps due to a disappointing experience in the past. Some of these farmers, Type 5 in Table 3.1, adopt at baseline as the expected marginal return may still be larger than the marginal cost of adoption, even if they underestimate the return. For another fraction of farmers that underestimate the adoption premium, referred to as Type 6 in Table 3.1, the expected marginal return will be less than the marginal cost of adoption, such that they do not adopt.

Heterogeneity in terms of prior experiences, expectations, and adoption behavior will lead to different effects of an intervention aimed at “correcting” incorrect beliefs about the relationship between the returns to Variety H and investments in inputs and practices (described in detail in the next section). In some cases, such as for adoption, effects for different farmers may go in opposite directions, potentially canceling out an overall average treatment effect. In other cases, such as for knowledge, some farmers may not be affected, diluting the overall effect. The model and the different farmer types summarized in Table 3.1 allow us to make predictions on the impact of an intervention designed to increase knowledge about the true relationship between the performance of a stochastically dominant variety and complementary efforts on four key outcome areas:

Table 3.1: Farmer types and model predictions

farmer type	baseline expectations	baseline adoption	effect on knowledge	effect on adoption	effect on expectations	effect on inputs
1	correct exp. of adoption premium	yes	none	none (always adopt)	none (correct at baseline)	none
2	correct exp. of adoption premium	no	none	none (never adopt)	none (correct at baseline)	none
3	inflated exp. of adoption premium	yes	yes ++	dis-adopt due to decr. exp. marg. return	more realistic	none
4	inflated exp. of adoption premium	no	yes +	none (never adopt)	none (correct at baseline)	none
5	reduced exp. of adoption premium	yes	yes ++	none (always adopt)	more realistic	increase +
6	reduced exp. of adoption premium	no	yes +	adopt due to incr. exp. marg. return	none (correct at baseline)	increase ++

++ indicates a large predicted effect, + indicates a small predicted effect.

1. Effect on knowledge: As Type 1 and 2 farmers are assumed to be already knowledgeable about the true relationship between Y_{it}^H and X_{it} , the intervention will have little effect on them.³¹ Types 3 to 6 are assumed to be unaware of the true relationship between Variety H and complementary inputs and practices; the intervention will thus increase knowledge. The knowledge effect will be larger for farmers that adopt at baseline (Types 1, 3, 5) since this removes “Never Adopters” who are likely to be less interested in the information (Types 2 and 4) from the sample.
2. Effect on adoption: We predict opposing effects on adoption behavior for farmer Types 3 and 6. Providing Type 3 farmers with information may cause them to dis-adopt if the new information reduces their expected marginal return below their marginal cost. For Type 6 farmers, the intervention may increase expectations of the return, and they may start adopting in response to the treatment if the increase is sufficiently high. Reducing expectations of farmers that do not adopt at baseline even though they have inflated expectations will not change their mind as this will reduce their expected returns even more (Type 4). Similarly, we do not expect that the intervention will change the adoption behavior of farmers who already adopt even though they underestimate potential yield effects (Type 5): these farmers will keep adopting as the intervention increases their expected returns to the stochastically dominant variety. Finally, as for knowledge, farmers that are aware of the correct relationship between inputs and Variety H (Types 1 and 2) are not expected to change adoption behavior in response to the intervention. The direction of the intervention’s effect on adoption will thus depend on the share of Type 3 and 6 farmers respectively. Note that if we only consider farmers that adopt at baseline, the expected effect on adoption will be negative as this excludes Type 6 farmers from the analysis.
3. Effect on expectations: We predict that the intervention results in expectations that are more in line with realized outcomes. This will likely only be the case for farmers that are unaware of the true relationship, and so we again do not expect an effect for Types 1 and 2. Furthermore, since our intervention aims to “correct” perceptions only for Variety H,

³¹Note that we do not know which farmers are knowledgeable and which farmers are not as we only measure knowledge at endline to avoid priming effects.

expectations of farmers that use Variety L at baseline are unlikely to be affected (as it is assumed that the production function of stochastically inferior varieties is common knowledge). Thus, we only expect an impact on farmers that plant the stochastically dominant variety at baseline and also have incorrect expectations (Types 3 and 5).

4. Effect on use of inputs and practices: Some farmers that were unaware of the true relationship between Variety H and complementary inputs and practices and receive new information about the importance of these complements may start investing more. This will be especially the case for Type 6 farmers who adopt due to the intervention (potentially after previously dis-adopting due to disappointing outcomes in the past) and will put the new knowledge into practice. To a lesser extent, farmers that consider adoption to be profitable despite low yield expectations may try to further increase yields by increasing effort (Type 5). Hence, for inputs and efforts, we expect a positive effect that becomes less strong if we restrict ourselves to farmers that adopt at baseline.

3.4 Intervention

The model predictions were tested using a field experiment conducted with almost 3,500 maize farmers in eastern Uganda. The intervention consisted of screening short, engaging videos about best practices in maize cultivation. The videos were shown individually to participating farmers on tablet computers by specially trained field enumerators. The content of the video scripts was developed following extensive interviews with experts, including agricultural extension agents, plant breeders, seed producers, government officials, and farmers themselves.

The video opens with a woman and a man standing in a well-kept maize plot inspecting their crop. The couple explains that they have been farmers for more than ten years but that their fields have not always been productive. They recount how they used to struggle to feed their children, but that over time, they learned how to grow more maize on less land. The secret of their success, the couple continues, lies in the adoption of improved technologies and best practices, such as the use of organic fertilizer, optimal plant spacing, and reduced seed rates. Furthermore, they explain that the use of an improved va-

riety and inorganic fertilizer contributed significantly to increased production. They conclude this introduction by stating that they are proud to be successful farmers who can feed their families and even produce surpluses that they can sell in the market. The viewer is then invited to become equally successful in farming by paying close attention as the featured (role model) farmers explain in detail the most important technologies, inputs, and practices that transformed their lives.

The treatment was implemented in the form of two variations of this video. The control group viewed the video as described above. The treatment group viewed a similar video that differed slightly in terms of content. Specifically, we added subtle recommendations for inputs and practices that are particularly important when cultivating improved maize varieties. The only difference between the control and treatment videos is that the latter makes explicit the fact that significant complementarities exist between improved varieties and recommended inputs and practices such as inorganic fertilizers and row planting. In effect, the treatment and control videos are identical, except that, after each practice or input that is shown, the treatment video explicitly mentions that the practice or input is “[...] particularly or even more important when you are using seed of an improved variety”.³² Generally, the treatment video focuses on the *benefits* of complements instead of the *need* for complements. The control video is about eight minutes long and can be found at <https://vimeo.com/781882803>. The treatment video is about twelve minutes long and can be found at <https://vimeo.com/781882930>, indicating four extra minutes of material. The other eight minutes are equal to the control video, no scenes are replaced or modified.

By randomizing which video is viewed by our sampled farmers, we can isolate the causal effect of making salient the fact that improved varieties do not substitute for inputs and effort, but in fact require more investment. The use of a control video has an additional advantage: since it is not clear to farmers or enumerators which video is the treatment and which is the control, we reduce the likelihood that results are driven by experimenter demand effects

³²For example, in the control video, the farmer explains that: “At planting time, I paid attention to recommended spacing, carefully measuring one foot between plants and 2.5 feet between rows. I first dug a four inch deep hole and added one water bottle cap of Di-Ammonium Phosphate (DAP). Then I added some soil. Afterwards, I put one maize seed in and covered it with soil.” In the treatment video, the farmer narrates the same scene but adds a pointed comment at the end of the exposition, stating: “Did you know that recommended spacing and using DAP is even more important when using improved seeds?”

(Bulte et al., 2014). Furthermore, to reduce the likelihood that treated households could provide information to households in the control group—a common problem in video-mediated information treatments (Van Campenhout, 2021)—randomization was conducted at the village level in a manner that ensured reasonable geographic and social distance between villages.

The experiment targeted the second agricultural season of 2021, where maize is sown in August and September and harvested in November and December. We implemented the treatment in April 2021, well before the start of the season, to ensure that farmers had the necessary information before making decisions on seed and input use. At this point in time, we also collected baseline data on our sample households.

The intervention was repeated just before planting in August 2021, and post-treatment data was collected in January and February 2022. The intervention was again repeated in the first season of 2022, with a final round of data collection conducted in July and August 2022. Note that this paper focuses on outcomes following the 2021 agricultural season since we do not expect significant results from continuing the intervention (i.e., providing farmers with the same information) in 2022. However, we do explore descriptive results from 2022 to provide insight into patterns of sustained adoption among treated farmers.

3.5 Data and empirical strategy

3.5.1 Sample

The field experiment was conducted in southeastern Uganda, an area known for its maize production by smallholder farmers, and where maize is considered both a food and cash crop. Because it was conducted as part of a larger study on maize seed supply chains, farmers were drawn from the catchment areas (market-sheds) of agro-input shops. The sampling frame was developed as follows: first, we listed all agro-input shops in eleven districts in southeastern Uganda, resulting in the identification of 347 agro-input dealers. We then asked these dealers to identify the villages where most of their customers come from. This sampling frame allows us to assume that sampled farmers have both reasonable and similar access to improved maize varieties if they choose to adopt as a result of our intervention, and that other constraining factors such as seed

quality, credit access, or individual preferences were similarly distributed across our treatment and control groups.

Next, field supervisors compiled household lists for each village and randomly sampled ten maize-cultivating households per village using systematic sampling (nth name selection technique). The enumerators interviewed 3,470 farmers using a household survey instrument that contained a wide range of questions about the individual, their household, and their farm. From an initial sample of 3,470 farmers who were interviewed in the baseline survey round, only 63 farmers dropped out in the subsequent survey round. We did not find that attrition differed significantly between treatment and control group, and thus proceed with the analysis on a balanced panel of 3,407 farmers.

3.5.2 Adoption

In this section, we explore the dynamics of improved variety adoption by smallholder farmers in our study area. We define adoption of improved maize varieties as follows. First, we asked farmers on how many plots they cultivated maize during the preceding season. From these plots, we randomly selected one plot and asked detailed questions about seed and varietal use, input use, and management practices. Based on the information collected, we then defined a farmer as an “Adopter” if they used either non-recycled (newly purchased, not saved) seed of (a) a hybrid or (b) an open pollinated variety. All others were defined as “Non-adopters.”³³

Figure 3.1 illustrates the evolution of varietal adoption among farmers over different survey rounds using this definition. We see that the share of adopters slowly increases over time: at Survey 1 (baseline), about 43% of farmers report to have sown an improved maize variety on the randomly selected plot. At the end of the first season, at the time of Survey 2 in January 2022, this figure increased to about 49% and, by Survey 3 in July and August 2022, to about 52%.

³³We acknowledge that this definition of adoption is not perfect. Seed of an open pollinated variety that has been recycled (saved) up to four times could still be considered as improved, and farmers using this seed could still be counted as adopters. However, we expect the problem of incorrect beliefs about “miracle seeds” and biased expectations to be most pronounced when smallholders do not have experience with the seed. If farmers recycle and use seed several times, their beliefs about the relationship between efforts and the returns to improved varieties will become closer to reality as they learn from season to season. That is why this stricter definition is useful to answer the questions raised in this paper. Also, most of our results remain robust to different definitions of adoption.

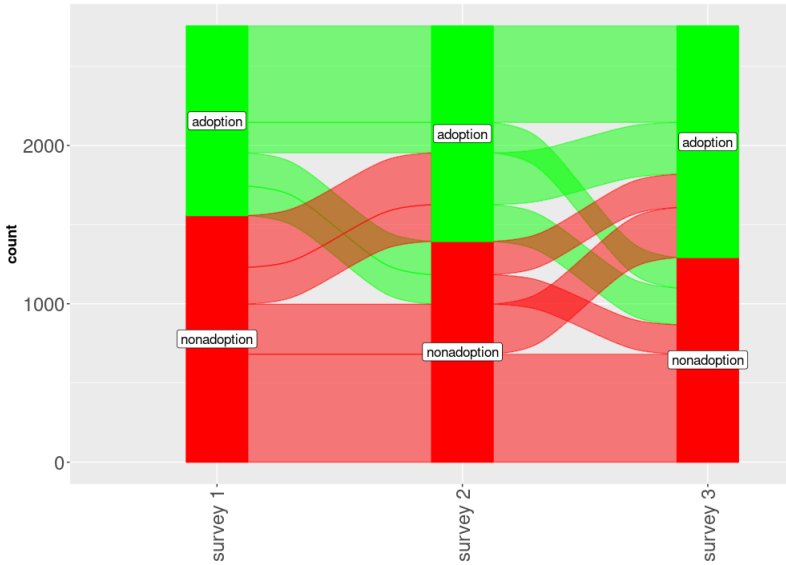


Figure 3.1: Dynamics of varietal adoption

Figure 3.1 also illustrates the dynamics of adoption in our sample. At the top, we see a substantial share of households (19%) that adopted in all three survey rounds. These could be considered “Always Adopters” or Type 1 and Type 5 farmers, as described earlier. At the bottom of the chart, we find an equally substantial share (22%) that can similarly be categorized as “Never Adopters” or Type 2 and Type 4 farmers. However, we also see that a large group of farmers that adopts during Survey 1 reverts to farmer-saved seed at the time of the second survey (13%) or still adopts at the time of Survey 2 but eventually dis-adopts at the time of Survey 3 (6%). During this same period, large numbers of households also start adopting. We see that 19% of non-adopting households adopt at the time of Survey 2 and 10% of households do not adopt in both Surveys 1 and 2 but do adopt by Survey 3. Finally, we find that the some households seem to be moving in and out of adoption (7%) or moving out and back into adoption (6%). Similar figures distinguished by treatment group can be found in Appendix 3.9.

Another indication of the dynamic nature of adoption is the fact that a substantial share of farmers that adopted at the time of the first survey seemed to be disappointed. Baseline data shows that 30% of farmers indicated that

were not satisfied with the quality of the planting material that they used; one in four indicated that they will not use the seed again in the future.

3.5.3 Empirical strategy

Due to the random assignment of participants to treatment and control groups, comparing outcome variable averages of treated and control participants provides unbiased estimates of the average treatment effects. Using an Analysis of Covariance (ANCOVA) regression framework, we regress outcomes of interest (knowledge, adoption, input use and effort, and expectations) on an indicator variable that takes the value of one if the household was in the treatment group and 0 otherwise, and include baseline values of the outcome variables as controls. Furthermore, as this study was part of a larger project with additional cross-randomized treatments, controls are included for the orthogonal treatments (demeaned and interacted with the main (video) treatment (Lin, 2013; Muralidharan, Romero, and Wüthrich, 2023)).

Since we have almost 3,500 observations in about 350 clusters, the original form of the sandwich estimator (Liang and Zeger, 1986) is used, with standard errors clustered at the village level, the level of randomization. For each of the four outcome families (knowledge, adoption, input use and effort, and expectations), we compute outcome indices, which is a common way to account for multiple hypothesis testing. To do so, we follow Anderson (2008), where each index is computed as a weighted mean of the standardized values of the outcome variables. The weights are derived from the (inverse) covariance matrix, such that less weight is given to outcomes that are highly correlated with each other. For these indices, signs of outcomes were switched where necessary so that the positive direction always indicates a “better” outcome.

3.6 Results

We look at impact on knowledge, adoption, expectations and harvest, and inputs and practices in separate subsections.

3.6.1 Impact on knowledge

First, we examine whether the treated participants are able to pick up the subtle messages in the treatment video. According to Prediction 1, we expect

a positive effect of the treatment on farmers' knowledge, and a larger effect for farmers that adopt at baseline. We test farmer knowledge by means of a short quiz where a number of questions were asked and enumerators read a set of alternative answers to farmers who then selected the response that they felt most appropriate.

The quiz begins with a general question asking farmers whether they think recommended cultivation practices like weeding and fertilizer application are less, equally, or more important when using improved varieties. This is followed by a more specific multiple-choice question on a particular practice—weeding—when cultivating an improved variety. Response options are: (1) you do not need to weed and remove Striga because seed of improved varieties is treated to resist weed infestation; (2) you do not need to weed and remove Striga in the first four weeks because seed of improved varieties is better at competing for sun, nutrients, and water than normal seed; and (3) you need to weed and remove Striga just as you would with unimproved varieties because maize seed does not compete well for sunlight, water, and nutrients. The quiz contains a similar question about a key input as well—fertilizer—when cultivating an improved variety. The options here are: (1) you do not need to use inorganic fertilizer because you already purchased seed; (2) you can use less fertilizer than you normally would since seed of an improved maize variety grows faster; (3) you need to use the amount of fertilizer that you would with unimproved varieties because also seed of an improved variety needs nutrition; and (4) you should use more fertilizer than you would normally use.

The quiz also contains a question that checks if farmers use sub-optimal plots to cultivate improved varieties by asking which plots are best suited. Response options are: (1) that it is best to save seed of an improved variety for poor plots, as it needs less nutrients; (2) that is best to use seed of an improved variety for plots that are furthest away from the home, as it needs less care; and (3) that the decision on what plot to plant seed should not be affected by the seed type. Another question explores how farmers think about the optimal investment in agriculture, i.e., whether to invest their resources in a single input or in a combination of inputs. The options are: (1) it is best to invest all your money in seed, because poor seed quality is the main cause of low yields; (2) it is best to invest all your money in fertilizer, because poor soil is the main cause of low yields; and (3) it is best to buy both fertilizer and seed, because good seed without fertilizer does not give good results.

Finally, the quiz includes a control question, answers to which are not expected to differ between treatment and control groups because they are featured in both versions of the video. Specifically, the question asks about the optimal spacing and seed rate for maize, with response options as: (1) one foot between plants and two and a half feet between rows with one seed per hill; (2) one foot between plants and two and a half feet between rows with two seeds per hill; and (3) two feet between plants and two and a half feet between rows with two seeds per hill. The four outcomes (excluding the control question) are also combined in an index following Anderson (2008).

Estimates of the average treatment effects on knowledge can be found in Table 3.2. The first column (1) provides the mean in the control group (with standard deviations in brackets below), mainly to get an idea of effect sizes. We see that knowledge is already high: 87% of farmers in the control group know that recommended inputs and cultivation practices like weeding or applying fertilizer are also important when using improved varieties.

Column (2) shows the estimated difference between the treatment and control groups for outcomes after the intervention, while column (3) also reports this difference, but only for the subset of farmers that adopted an improved variety at baseline. The rationale for restricting our sample is alluded to in Section 3.3: because the restricted sample retains farmers for whom the treatment effect is likely to be largest, we expect larger estimates in column (3) than in column (2).

We find that knowledge, as measured by the quiz questions, increases for all variables, and generally more so for the subset of farmers that used an improved variety at baseline. For instance, the share of farmers that knows complementary inputs and practices are at least as important when using improved varieties increases from 87.1 to 89.3%. Furthermore, the share of farmers that recommends investing in different inputs (as opposed to investing all money in only one input), increases from 73.5 to 75.7%. If we only consider farmers that adopted at baseline, the increase over the control amounts to almost 5 percentage points.

After adjusting standard errors for clustering at the village level, none of the differences for the entire sample is statistically significant at conventional levels. However, if we only consider the subset of farmers that adopted at baseline, we see that the intervention increased knowledge, as measured by the index, significantly, probably because these farmers were more interested in this

Table 3.2: Average treatment effects on knowledge

	(1)	(2)	(3)
Farmer knows inputs and practices are important when using an improved variety	0.871 (0.336)	0.022 (0.015)	0.026 (0.019)
Farmer knows weeding is important when using an improved variety	0.790 (0.407)	0.025 (0.022)	0.028 (0.026)
Farmer knows applying fertilizer is important when using an improved variety	0.835 (0.371)	0.009 (0.016)	0.011 (0.021)
Farmer knows plot selection should be independent of using an improved variety	0.792 (0.406)	0.007 (0.025)	0.020 (0.031)
Farmer knows it is best to invest in different inputs instead of putting all eggs in one basket	0.735 (0.441)	0.022 (0.023)	0.060* (0.028)
Farmer knows recommended seed spacing and rate	0.687 (0.464)	0.029 (0.024)	0.017 (0.030)
Knowledge index	0.015 (0.580)	0.046 (0.036)	0.083+ (0.042)
Observations	1707	3407	1435

Note: Column (1) reports control group means post-intervention (and standard deviations below); column (2) reports differences between treatment and control post-intervention; column (3) reports differences between treatment and control post-intervention for farmers that adopt at baseline; **, * and + denote significance at the 1, 5 and 10% levels; standard errors are clustered at the village level.

information. The overall effect is driven by increased knowledge about optimal agricultural investments among treated farmers. Even though we cannot detect a treatment effect for the entire sample, we note that all coefficient estimates are moving in the same direction. This may be due to the fact that, ex-post, it turns out that many of the farmers were already able to indicate the correct response, and hence there is little scope for further improvement. The significant results for the baseline adopters are in line with Prediction 1 in Section 3.3.

3.6.2 Impact on adoption

We now test the main hypothesis of this paper: whether farmers who were informed with subtle messages that improved varieties need substantial investment in complementary inputs and management practices behave differently in

terms of seed use in subsequent seasons than farmers that were not similarly informed. To this end, we asked farmers which maize variety they planted on the randomly selected maize plot in the season prior to the survey. We again define adoption as described earlier and used in Figure 3.1. In addition, we investigate other outcomes that are related or even partly overlapping. For instance, we test if there are differences in the use of recycled seed between the treatment and control group, where we define recycled seed as seed that a farmer has saved themselves or obtained from another farmer who saved it (e.g., a neighbor or relative). Another related outcome is the share of farmers that report having purchased seed from an agro-input shop. The three outcomes are also combined in an index following Anderson (2008).

Results are summarized in Table 3.3 and show that the intervention decreases adoption. Column (1) shows sample means of the four outcomes at baseline with standard deviations in the brackets below. We find that 44% of farmers use fresh seed of improved varieties and that one-third of farmers reports that the seed that they planted on the randomly selected plot was obtained from an agro-input dealer. Column (2) shows pre-treatment balance between treatment and control groups. We see that the randomization was successful, as there is no significant difference in varietal adoption behavior between farmers that will be exposed to the treatment and those that will not.

Column (3) shows the difference between treatment and control groups for outcomes after the intervention. Our theory suggests that in response to being sensitized about the importance of using complementary inputs and management practices when using an improved variety, some farmers (Types 3 and 6) will change their adoption behavior (Prediction 2 in Section 3.3). A share of farmers that initially underestimated the returns to improved varieties (Type 6) will start adopting as their expected marginal return is increased by the treatment. Another share of farmers that initially overestimated the probability of an adoption premium (Type 3) will dis-adopt as their expected marginal return is reduced by the treatment. We find that adoption, as measured by the index, significantly decreases for the entire sample. Furthermore, all coefficients move in the direction of dis-adoption. This implies that farmers are less likely to use improved seed and seed bought at an agro-input shop but more likely to use farmer-saved seed in accordance with our earlier definitions of adoption. Since the two opposing effects for farmer Types 3 and 6 partly cancel each other out, the dis-adoption effect is not pronounced.

Table 3.3: Average treatment effects on adoption

	(1)	(2)	(3)	(4)
Farmer planted seed of an improved variety	0.435 (0.496)	-0.002 (0.022)	-0.042* (0.021)	-0.077** (0.029)
Farmer planted seed from agro-input shop	0.328 (0.469)	-0.004 (0.020)	-0.022 (0.020)	-0.056* (0.028)
Farmer planted seed that was recycled	0.569 (0.495)	0.020 (0.022)	0.032 (0.021)	0.076** (0.028)
Adoption index ¹	0.009 (0.942)	-0.004 (0.042)	-0.068+ (0.041)	-0.121* (0.055)
Observations	3470	3470	3407	1435

Note: Column (1) reports means at baseline (and standard deviations below); column (2) reports differences between treatment and control at baseline; column (3) reports differences between treatment and control post-intervention; column (4) reports differences between treatment and control post-intervention for farmers that adopt at baseline; **, * and + denote significance at the 1, 5 and 10% levels; standard errors are clustered at the village level. ¹For this index, signs of outcomes were switched where necessary so that the positive direction always indicates adoption of improved varieties.

To separate the two opposing effects, we restrict the sample to farmers that adopted at baseline in column (4). We see that the estimated effects become stronger when we restrict attention to this subgroup (and exclude Type 6 farmers from the analysis). Farmers who were exposed to the treatment are almost 8 percentage points less likely to adopt fresh seed of an improved variety. We see another particularly strong increase in the share of farmers that uses seed recycled from the previous harvest in the treatment group and a somewhat lower but still significant reduction in farmers who bought seed from an agro-input dealer. The treatment also has a significant and more pronounced negative effect on the adoption index for this subgroup of farmers that adopted at baseline.

3.6.3 Impact on expectations and harvest

Since the intervention is designed to affect farmer behavior by “correcting” their expectations, we explore the plausibility of this impact pathway by testing if post-intervention farmers feel their yield expectations were met. As mentioned in Prediction 3 in Section 3.3, we think this will particularly be the case if

we restrict the sample to farmers that adopt at baseline. We also measure harvest-related outcomes (production and yield) on a randomly selected maize plot. The three outcomes are also combined in an index following Anderson (2008).

The results in Table 3.4 show that yield expectations have been significantly affected. We again report baseline means and balance in columns (1) and (2). However, we did not ask if expectations were met at baseline, and so we report the control group average post-intervention and do not test for baseline balance for the expectations variable. Note that a large majority of farmers indicated that they harvested less than expected.

Column (3) shows that, in line with our prediction, a significantly higher share of farmers in the treatment group state that they produced what they expected. The effect is larger for the subset of farmers that adopted at baseline, see column (4). This suggests that a subset of farmers indeed started out with inflated expectations, which were “corrected” after they learned that improved varieties are not “miracle seeds.”

Finally, the table shows that the average farmer produces about 460 kg of maize on the randomly selected plot. The average size of these plots is slightly larger than one acre, such that yields are about 440 kg per acre. The intervention does not seem to have any impact on maize production or yield.

3.6.4 Impact on use of inputs and practices

Finally, we investigate how the intervention affects the use of inputs (other than seed) and practices. For inputs and practices, the effect is expected to be positive, but weak (see Prediction 4 in Section 3.3).

We examine a range of cultivation practices and complementary inputs in line with what is featured in both treatment and control videos. The first outcome is an indicator for single-stand row planting. Row planting is an important management practice that can lead to significant yield gains. Under row planting, space is used optimally such that plants have sufficient nutrients, sunlight, and room to grow. However, row planting increases workload, meaning farmers often engage in broadcast planting, which is less demanding on their labor.

Reducing the seeding rate (i.e., the number of seeds sown) is the second outcome of interest. Farmers often plant more seed than necessary because

Table 3.4: Average treatment effects on expectations and harvest

	(1)	(2)	(3)	(4)
Yield as expected	0.15 (0.36)		0.029 ⁺ (0.017)	0.052* (0.024)
Production in kg	463.702 (399.319)	16.444 (18.004)	2.562 (12.713)	-4.289 (19.308)
Yield in kg/acre	436.332 (280.790)	9.559 (12.128)	6.790 (12.129)	23.875 (16.447)
Harvest index	-0.004 (0.755)	0.006 (0.038)	0.026 (0.035)	0.051 (0.049)
Observations	3470	3470	3407	1435

Note: Column (1) reports means at baseline (and standard deviations below); column (2) reports differences between treatment and control at baseline; column (3) reports differences between treatment and control post-intervention; column (4) reports differences between treatment and control post-intervention for farmers that adopt at baseline; **, * and + denote significance at the 1, 5 and 10% levels; standard errors are clustered at the village level.

they fear that it may not germinate. However, using more than two seeds per hill leads to stunted maize growth due to competition for light and nutrients. At the same time, just as for row planting, a lower seed rate may increase the workload, as farmers need to engage in gap filling after one week if seeds fail to germinate.

The next three outcomes relate to fertilizer use. The application of organic fertilizer is important for soil structure, while inorganic fertilizers such as diammonium phosphate (DAP) or nitrogen, phosphorus, and potassium (NPK) and urea (nitrogen) are used to provide essential nutrients at particular points in time. The cost of organic fertilizer is mainly in terms of labor, whereas both DAP and urea need to be bought from an agro-input shop and applied during planting (DAP) and at early stages of growth (urea).

Farmers should weed within the first week after planting and as often as possible. Official recommendations are to weed at least three times per seasons. Furthermore, invasive insects such as the fall armyworm (*Spodoptera frugiperda*) or maize stalk borer (*Busseola fusca*) can severely reduce yields. Pesticides, herbicides, fungicides, and insecticides are widely available in agro-input shops under commercial names such as Rocket, Lalafos and Dudu ace-lamectin. While weeding requires labor, pesticides come at a pecuniary cost.

Finally, we look at differences in re-sowing or gap-filling. This involves revisiting the plot after planting and inspecting the hills for seed germination. If a seed does not germinate, a new seed is planted in that location. Re-sowing, reduced seed rate, and row-planting are thus likely to be correlated. We also combine all outcomes in an overall index following Anderson (2008).

Results are reported in Table 3.5 and show no impact of the intervention on the use of inputs and practices. As in previous tables, columns (1) and (2) report means and orthogonality for outcomes before the treatment. We find an imbalance for the number of times that a farmer reports to have weeded and the likelihood that farmers re-sow after one week. Note that the imbalance goes in different directions, which makes it less likely that it is caused by a structural difference between treatment and control group such as consistently lower efforts in one group, and more likely to be the result of chance.

Column (3) shows that farmers do not invest more effort in response to the intervention. On the contrary (and especially if we only consider the subset of farmers that adopted at baseline, see column (4)), farmers appear to be less likely to plant in rows. The negative effect on some practices may be due to the fact that farmers may adopt some complementary inputs or practices, but in sufficient quantities and/or suboptimal combinations for the adoption premium to realize. When farmers subsequently dis-adopt, they may also stop using these inputs or practices, leading to a negative expected effect on practices for Type 3 farmers in Table 3.1. In particular, for the case of row planting, it may be that farmers that adopt at baseline simply follow planting instructions that are typically printed on the seed bags. However, only row-planting without additional inputs such as fertilizer or pesticide use may not lead to expected yields, and disappointed farmers may dis-adopt in the next season which case they may also revert to plating methods used before adoption.

3.6.5 Heterogeneity analysis: remote households

To better understand the impact of the video on adoption, we investigate heterogeneous treatment effects on remote households. Note that this analysis is of exploratory nature and was not registered in the pre-analysis plan at the AEA RCT registry. Instead, we conducted a Lasso regression analysis to find variables that predict dis-adoption due to the treatment, which yielded remoteness as an interesting variable to explore. After further investigation and finding

Table 3.5: Average treatment effects on use of inputs and practices

	(1)	(2)	(3)	(4)
Row-planting	0.243 (0.429)	0.025 (0.022)	-0.070* (0.027)	-0.093** (0.033)
Reduced seed rate	0.237 (0.425)	0.010 (0.021)	0.009 (0.019)	-0.007 (0.028)
Organic fertilizer use	0.075 (0.263)	-0.009 (0.011)	-0.013 (0.017)	-0.013 (0.023)
DAP/NPK use	0.251 (0.434)	-0.020 (0.024)	-0.029 (0.019)	-0.045 (0.028)
Urea use	0.076 (0.265)	0.001 (0.013)	0.002 (0.015)	0.013 (0.024)
Weeding frequency	2.561 (0.650)	0.084** (0.026)	-0.021 (0.027)	-0.001 (0.037)
Pesticide etc. use	0.412 (0.492)	0.031 (0.024)	0.003 (0.023)	0.004 (0.032)
Re-sowing	0.482 (0.500)	-0.046* (0.023)	0.013 (0.022)	0.033 (0.029)
Early planting	0.699 (0.459)	-0.018 (0.024)	0.012 (0.025)	0.021 (0.031)
Early weeding	0.606 (0.489)	0.032 (0.020)	0.026 (0.021)	0.040 (0.028)
Inputs index	0.008 (0.400)	0.009 (0.020)	-0.008 (0.019)	0.005 (0.025)
Observations	3470	3470	3407	1435

Note: Column (1) reports means at baseline (and standard deviations below); column (2) reports differences between treatment and control at baseline; column (3) reports differences between treatment and control post-intervention; column (4) reports differences between treatment and control post-intervention for farmers that adopt at baseline; **, * and + denote significance at the 1, 5 and 10% levels; standard errors are clustered at the village level.

that the interaction of the treatment indicator and the remoteness indicator is indeed significant at the 1% level in the full ANCOVA model, we decided to add this heterogeneity analysis.

As a measure of remoteness, we use a homestead's distance to the village headquarters in kilometers, a variable with a minimum of zero and a maximum of 15, a median of 0.5, and a mean of 0.75. A household is defined as more remote or less connected if its distance to the village headquarters is larger than the average distance. A household that is less than average far away from the village headquarters is defined as less remote or more connected. Tables 3.6 and 3.7 show that the less connected households clearly drive the results. Smallholders who live further away from the city center are the ones that dis-adopt when learning about the complementary between improved maize varieties and other investments.³⁴

Suri (2011) provides one possible explanation for this result. It is more difficult for less connected farmers to access improved varieties but also the necessary complementary inputs such as DAP/NPK, Urea, and pesticide. Generally, agro-input shops are located in the centers of towns and villages, in markets, trading centers, and other key market sheds. For more remote farmers, this implies higher costs of acquiring the agricultural inputs and technologies and lower net returns due to often poor infrastructure. As the information intervention highlights the importance of complementary investments when using improved maize varieties and further reduces the perceived net benefits of adoption for farmers that initially had inflated expectations of the adoption premium, the remote farmers with higher costs and lower net returns may be the most sensitive to this information and the first ones to dis-adopt. This is only one possible reason for this exploratory result, there are many alternative explanations like the larger insecurity and variability in output prices in more isolated areas (Stifel and Minten, 2008).

3.7 Conclusion

This paper was motivated by evidence suggesting that farmers are often unaware that many agricultural technologies such as improved seed varieties require substantial complementary inputs, better management practices, and

³⁴Note that this result must be interpreted with care as sampling sizes are smaller when we split the sample, having implications for statistical power.

Table 3.6: Average treatment effects on adoption - Less connected households

	(1)	(2)	(3)	(4)
Farmer planted seed of an improved variety	0.423 (0.494)	0.032 (0.033)	-0.065* (0.031)	-0.130** (0.046)
Farmer planted seed from agro-input shop	0.320 (0.467)	0.003 (0.030)	-0.038 (0.029)	-0.119* (0.046)
Farmer planted seed that was recycled	0.583 (0.493)	0.001 (0.031)	0.053+ (0.031)	0.116* (0.046)
Adoption index ¹	-0.011 (0.936)	0.033 (0.061)	-0.114+ (0.059)	-0.226* (0.090)
Observations	1258	1258	1131	458

Note: Column (1) reports means at baseline (and standard deviations below); column (2) reports differences between treatment and control at baseline; column (3) reports differences between treatment and control post-intervention; column (4) reports differences between treatment and control post-intervention for farmers that adopt at baseline; **, * and + denote significance at the 1, 5 and 10% levels; standard errors are clustered at the village level. ¹For this index, signs of outcomes were switched where necessary so that the positive direction always indicates adoption of improved varieties.

Table 3.7: Average treatment effects on adoption - More connected households

	(1)	(2)	(3)	(4)
Farmer planted seed of an improved variety	0.443 (0.497)	-0.018 (0.026)	-0.029 (0.025)	-0.041 (0.036)
Farmer planted seed from agro-input shop	0.333 (0.471)	-0.007 (0.025)	-0.011 (0.025)	-0.011 (0.036)
Farmer planted seed that was recycled	0.560 (0.497)	0.032 (0.026)	0.021 (0.025)	0.046 (0.036)
Adoption index ¹	0.023 (0.946)	-0.023 (0.051)	-0.041 (0.049)	-0.047 (0.071)
Observations	2097	2097	1897	807

Note: Column (1) reports means at baseline (and standard deviations below); column (2) reports differences between treatment and control at baseline; column (3) reports differences between treatment and control post-intervention; column (4) reports differences between treatment and control post-intervention for farmers that adopt at baseline; **, * and + denote significance at the 1, 5 and 10%F levels; standard errors are clustered at the village level. ¹For this index, signs of outcomes were switched where necessary so that the positive direction always indicates adoption of improved varieties.

greater effort for their benefits to realize. In a sense, farmers overestimate the returns to a technology and are disappointed when they compare expectations to realized yields. As learning about a new technology is hard, farmers may attribute the disappointing results to the technology itself and dis-adopt. This is consistent with findings suggesting that farmers blame poor returns on inputs they believe to be counterfeit or of low quality even when objective quality assessments show otherwise (Barriga and Fiala, 2020; Michelson et al., 2021).

To credibly test this hypothesis—specifically, that farmers think of improved varieties as “miracle seed”—we conducted a field experiment built around a short, engaging video on recommended input use and management practices for maize cultivation in eastern Uganda. We produced two versions of the video that differ only in terms of the presence (absence) of subtle messaging about the salience of recommended inputs and practices for the treatment (control) group. Screenings of the two versions were randomly assigned to villages in our study area, and then to maize farmers in those villages, resulting in a sample of almost 3,500 farmers who were interviewed at regular intervals to uncover any changes in their knowledge about best practices in maize cultivation as well as their seed or variety choices, their expectations of yield and output, and their use of complementary inputs and management practices.

While we do not find treatment effects that differ significantly from zero for knowledge outcomes for the entire sample, we do observe that all coefficients move in the expected direction and suspect that the lack of statistical significance may be caused by low power given an already high level of knowledge among our sampled farmers. We do find, consistent with our theory, that the intervention significantly improved knowledge for farmers that adopted at baseline, probably because they were more interested in the information.

For the main outcome of interest—behavior related to seed choices—we find that treated farmers were less likely to use improved varieties, and generally more likely to dis-adopt. An exploratory heterogeneity analysis reveals that mainly more remote households dis-adopt due to the treatment. We also find that farmers that received the treatment were more likely to report that their harvest was in line with what they expected. Both findings are again consistent with our theory. We found no overall effect of the treatment on input use and management practices, although there is some indication that especially costly inputs and practices were reduced. Taken together, we conclude that there

are indeed indications that farmers consider improved maize seed varieties as “miracle seed” and that it is challenging to learn from own experience.

The treatment effect on adoption could imply that treated farmers dis-adopt permanently if the video makes the hidden price of adopting more salient to them and reduces their expected profits. Alternatively, the observed dis-adoption in the short run could be accelerated dis-adoption that would have happened anyway if the video reduces a bias and accelerates the process of expectations becoming more realistic. If it is the latter, then the effect is likely to be short-lived and there will be lower dis-adoption in subsequent seasons for two reasons. Firstly, if the treatment only accelerated dis-adoption that would have happened anyway, part of this dis-adoption process will have now also started in the control group: a share of control farmers with biased expectations that initially adopted will have experienced a disappointing harvest, adjusted their expectations, and dis-adopted. Secondly, a share of treated farmers that initially dis-adopted will re-adopt: information about complementarities and changes in perceptions may reduce their risk premia, particularly in the longer run. In this scenario, farmers try improved seed with complementary inputs and learn that outcomes are consistently in line with expectations. They may experience that combining improved technologies with complementary inputs is consistently more profitable. And even if margins are lower, they may trade off the reduction in risk with increased investment. This is in line with what we find looking at data from subsequent seasons: when we revisit the farmers after one additional agricultural season, we no longer find any differences between treatment and control groups.³⁵ At least some treated farmers who had dis-adopted in response to new information and more realistic expectations are, in fact, cultivating the new variety again; these may now be in for the long run. This indicates that the video did not increase dis-adoption permanently but only accelerated it.

Our findings have implications for the study of technology adoption dynamics. We have seen that disappointment about the performance of a technology that is erroneously attributed to the technology itself may lead to dis-adoption. As long as this learning failure is not corrected—for instance, by pointing out that the seed is good; the problem is with complementary input use—farmers will not adopt anew. Worse, as we learned from extension workers who complain farmers blame improved seed varieties for the proliferation of the parasitic

³⁵Data is available from the authors upon request.

Striga weed, “fake news” may travel faster than correct information (Ledgerwood and Boydstun, 2014; Hornik et al., 2015) leading to dis-adoption at more aggregate levels, further complicating (social) learning.

Our study also casts some doubt on the suggestion that Bayesian learning via sequential adoption can be a successful strategy for smallholder farmers in the long run (Leathers and Smale, 1991; Ma and Shi, 2015). If there are important interaction effects between technologies, inputs and practices, it seems unreasonable to assume that farmers can try out all possible combinations of inputs to learn about these interactions in a Bayesian fashion, at least in a reasonable time frame.

Our results differ from other studies that find that improved technologies increase agricultural productivity by crowding in modern inputs and cultivation practices (Emerick et al., 2016; Bulte et al., 2023). A possible explanation for our opposing results may be that Emerick et al. (2016) and Bulte et al. (2023) provided the improved technology (also an improved seed variety) for free as part of the experiment, potentially resulting in an income effect, i.e., the money that treated farmers did *not* use to purchase seed was instead allocated to the purchase of complementary inputs.³⁶ In our experiment, no free seed was provided, so when adoption decisions were made, farmers had to take the combined cost of seed *and* cost of complementary inputs into account, further eroding the expected profitability of the improved technology.

Finally, our findings have implications for how public and private actors in the agriculture sector should promote new technologies. If smallholders’ information sources such as private input dealers and public extension agents are not sufficiently able to communicate the importance of complementary inputs and practices, then lower likelihoods of sustained adoption may result. Worse, if smallholders have incorrect perceptions about poor quality caused by misattribution, the persistence of these perceptions may crowd out the market for quality inputs (Bold et al., 2017). And while the distribution of free or subsidized technologies and inputs may go some way in encouraging farmers’ learning processes and “correcting” their perceptions (e.g., with unique standalone technologies (Omotilewa, Ricker-Gilbert, and Ainembabazi, 2019)), this

³⁶Emerick et al. (2016) do discuss the possibility that their effects are driven by an income effect. However, in the presence of an income effect, they understand the effect of the additional income resulting from the adoption of the technology (a flood-tolerant rice variety). The income effect we are concerned about is one that results from farmers receiving seed for free, potentially freeing up money for other investments.

approach can break down when complementary inputs and practices are not part of the package, which may again lead to disappointment among farmers.

Our findings suggest that agricultural development programs, extension providers, and agri-input companies need to focus less on marketing single “miracle” technologies for smallholders, and more on the design and communication of comprehensive packages that include both agronomic and economic information on topics such as expected variation in yield and output, sensitivity of timing for specific farming tasks, magnitude and costs of family and hired labor, and the relative drudgery of effort, among many others. We conclude that the design and communication of comprehensive packages requires greater investment in the form and content of rural education, extension and advisory services, and agri-input marketing strategies.

3.8 Acknowledgments

This study was undertaken with generous financial support from the Netherlands–CGIAR research programme on Seed Systems Development (grant number W08.240.105), which is funded by the Netherlands Organisation for Scientific Research (NWO-WOTRO). Additional support was provided by the CGIAR Research Program on Policies, Institutions, and Markets (PIM), the CGIAR Seed Equal Research Initiative, and the CGIAR Market Intelligence Research Initiative which are funded by contributors to the CGIAR Fund. Furthermore, Caroline Mieke benefited from support by the Fonds Wetenschappelijk Onderzoek – Vlaanderen (FWO) and the Fonds de la Recherche Scientifique (FNRS) under EOS project G0G4318N.

3.9 Appendix

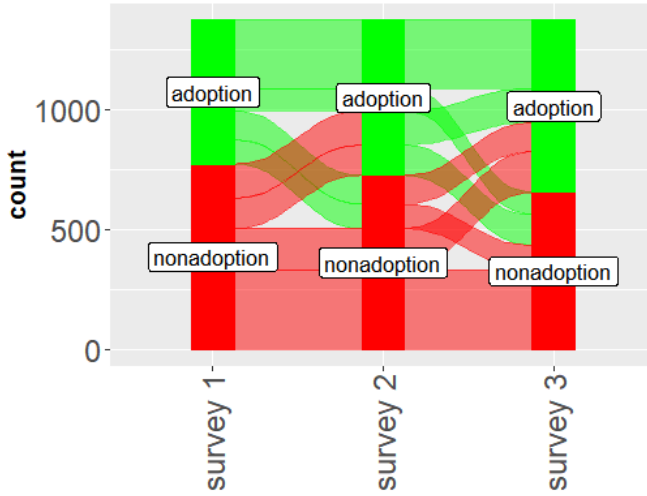


Figure 3.2: Dynamics of varietal adoption - Treated farmers

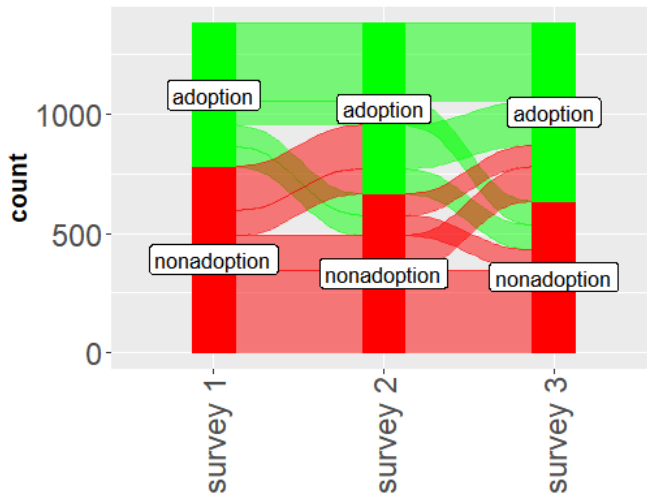


Figure 3.3: Dynamics of varietal adoption - Control farmers

The (perceived) quality of agricultural technology and its adoption: Experimental evidence from Uganda

This chapter is co-authored with Robert Sparrow (Wageningen University and Erasmus University Rotterdam, Netherlands), David Spielman (IFPRI, United States), and Bjorn Van Campenhout (IFPRI and KU Leuven, Belgium).

4.1 Introduction

Over the next decades, farmers in sub-Saharan Africa will need to produce more food on less land under increasingly difficult climatic conditions (Tilman et al., 2011). The use of climate-smart agricultural practices and improved inputs such as higher-yielding and drought-tolerant crop varieties is thought to be at least part of the solution (Evenson and Gollin, 2003). Unfortunately, the adoption of improved agricultural inputs and technologies across the region seems to be stagnating, or at least advancing at a slower pace than required (Suri and Udry, 2022). As a result, differences in agricultural yields between sub-Saharan Africa and countries in Latin America and Asia have almost doubled since 1961 (Magruder, 2018).

Several key constraints to agricultural technology adoption have been tested in recent years. These include poor access to information (Ashraf, Giné, and

Karlan, 2009), procrastination and time-inconsistent preferences (Duflo, Kremer, and Robinson, 2011), heterogeneity in the net benefits derived from the technology due to differences in infrastructure and transaction costs (Suri, 2011), missing markets for risk and credit (Karlan et al., 2014), and challenges related to learning about new technologies (Hanna, Mullainathan, and Schwartzstein, 2014).

More recently, issues related to the quality of inputs such as improved seed varieties, inorganic fertilizers, and pesticides have emerged as a potential constraint to their adoption by smallholder farmers. Bold et al. (2017) argue that farmers can hardly assess quality from simple visual inspection at the time of purchase, so information asymmetries between sellers and buyers characterize the markets for seed and fertilizer, in turn crowding out the market for quality inputs in Uganda, similar to what happens in Akerlof’s seminal “Market for Lemons” study (1970). However, subsequent research argues that it is not clear if these quality issues are due to agro-input dealers intentionally adulterating inputs, or if this is simply because they lack knowledge and skills to preserve quality (Barriga and Fiala, 2020). Furthermore, it is not even clear if these quality issues are real: while some studies argue that input quality is indeed lacking (Ashour et al., 2019), others argue that input quality is sufficient but farmers’ perceptions are to blame (Michelson et al., 2021; Wossen, Abay, and Abdoulaye, 2022).

We attempt to answer some of these questions through a field experiment targeting agro-input dealers and smallholder farmers in their catchment areas in the nascent market for improved maize varieties (high-yielding cultivars like open-pollinated and hybrid varieties) in eastern Uganda. Agro-input dealers are essential for agricultural technology adoption in countries with large farmer populations living in remote areas with poor infrastructure. A reasonably dense network of semi-formal agro-input shops provides access to technologies to rural farmers. Often, these dealers also provide services like agricultural advice or even credit to smallholders.

At the same time, the informal nature of many agro-input shops may imply that they are a weak link in the supply chain for quality inputs, a risk that is likely to be smaller upstream where larger producers and importers face more scrutiny from the government. Agricultural inputs such as seeds and fertilizers are sometimes stored in sub-optimal conditions (e.g., in direct sunlight or in moist environments) or handled in harmful ways (e.g., stored beyond the expiry

date or repackaged). There is some evidence of this kind of quality reduction. In a comprehensive study of the Ugandan seed supply chain, Barriga and Fiala (2020) document various issues related to handling and storage that may reduce input quality. For example, dealers often repack seed from larger bags packed by seed companies into smaller bags in order to offer quantities which are convenient and affordable to smallholder farmers. Important information including the expiry date and variety name can be lost during repackaging. Furthermore, seed is often repackaged in air tight polyethylene bags, which affect aeration and seed viability. Open air storage of bags can also lower the quality of seeds (Bold et al., 2017). Temperature control after the seed leaves the breeders is crucial, too (Barriga and Fiala, 2020). Inventory carryover, poor rotation of seed stock and storage in moist conditions or in direct sunlight further reduce seed quality. That is because the bio-deterioration of maize is sensitive to temperature and humidity (Curzi, Nota, and Di Falco, 2022), seed moisture affects the occurrence of storage fungi (Govender, Aveling, and Kritzinger, 2008), and many quality attributes of seed tend to degrade with storage duration time and shelf life (Hoffmann et al., 2021).

In a first hypothesis, we expect that a lack of dealer knowledge leads to deterioration in maize seed quality. Training agro-input dealers and providing them with information on proper seed handling and storage could increase quality and subsequent adoption. Lack of information is pervasive in developing countries and often leads to sub-optimal outcomes for the rural poor. As a result, a small piece of information can make a large difference (Duflo and Banerjee, 2011). Also in the context of smallholder technology adoption, knowledge gaps have been identified as a key constraint, and governments around the world invest in public agricultural advisory services (Anderson and Feder, 2004). While the need for policies and interventions that strengthen input marketing capacity and infrastructure has been acknowledged decades ago (Tripp and Rohrbach, 2001), most studies target smallholder farmers with information, and we are unaware of studies that tackle knowledge gaps among (small) agro-input dealers.

The fact that seed quality cannot easily be observed by farmers may also result in a lack of incentives for agro-input dealers to invest in quality preservation. In a context similar to ours, Hoffmann et al. (2021) look at maize grain as an output in rural Kenya and find that there is no incentive for sellers to address food safety issues because they are not observable for buyers. Worse,

agro-input dealers may intentionally sacrifice quality to cut costs and increase profits, for example, by mixing improved or fresh seed with local or old seed. There is some evidence of this kind of adulteration and counterfeiting in the Ugandan agricultural input supply chain. Bold et al. (2017) find that hybrid maize seed contains less than 50% authentic seeds and that 30% of nutrient is missing in fertilizer. Ashour et al. (2019) find that the average bottle of herbicide is missing 15% of the active ingredient and nearly one in three bottles contains less than 75% of the ingredient advertised.

That is why, in a second hypothesis, we expect that asymmetric information reduces dealers' incentives to provide quality seed. Addressing these asymmetries may lead to better quality, in turn increasing adoption. Uganda regulates seed quality by means of certifications and standards, but they provide farmers with a relatively weak and unreliable indication of quality. Alternatives such as electronic verification systems have also been experimented with, but the cost of implementation has proven challenging, and they depend on the reliability of the underlying seed certification system.

In addition to these problems caused by the lack of incentives for agro-input dealers, asymmetric information may also lead to situations where farmers fail to adopt because they misperceive the quality of the inputs in the market. Michelson et al. (2021) establish that the nutrient content of fertilizer in Tanzania meets industry standards but that farmers believe that it is adulterated. Wossen, Abay, and Abdoulaye (2022) show that farmers in developing countries routinely misperceive input quality and that rectifying this misperception may improve farmers' investment choices and productivity outcomes. Note that also here, a vicious cycle emerges, where farmers continue to perceive quality to be poor even though it may have improved, which in turn reduces incentives for agro-input dealers to maintain quality.

To address issues caused by asymmetric information, we implement a novel intervention: a decentralized information clearinghouse that is based on crowd-sourced information and works through reputational mechanisms, much like [yelp.com](https://www.yelp.com) or [tripadvisor.com](https://www.tripadvisor.com). We ask smallholder farmers to rate agro-input dealers in their neighborhood on a number of seed quality attributes. We use this information to score and rank agro-input dealers, and disseminate these ratings back to both, farmers and agro-input dealers. Ratings were collected and disseminated twice to increase the likelihood that dealers change their behavior and that farmers know and trust the scores.

The objective of the clearinghouse is thus to make maize seed quality observable. However, some may argue that farmers can assess seed quality well after one agricultural season: shortly after planting, farmers can observe germination rates, i.e., the proportion of seeds that germinate, and later how fast the seed matures. Some seed may also be more susceptible to pests and diseases, while other seed may be particularly tolerant in terms of drought. After harvest, the farmer can observe the yield. In the limit, farmers can perfectly observe seed quality, and there is no need for a clearinghouse. However, others may argue that farmers cannot assess seed quality even after using it, because there are so many factors at play in agricultural production: if farmers experience a disappointing harvest, they cannot safely conclude that the seed material was poor because it could have also been poor soil, insufficient, late or too much rain, or own mismanagement like late planting or insufficient weeding. Misattribution occurs when farmers mistakenly ascribe bad outcomes to bad inputs, rather than to other possible causes. This would imply that the clearinghouse will not work, as farmers cannot assess seed quality at all. In the limit, improved maize seed would resemble a credence good and the clearinghouse ratings of farmers would be mainly noise. Even though farmers and dealers might still change their behavior in the short run because they expect the clearinghouse to work, this effect would fade out as soon as both actors learn that the ratings are as good as random. While there is considerable evidence that farmers cannot perfectly assess seed quality (e.g., Bold et al., 2017) and authors like Tjernström et al. (2021) argue that sub-Saharan Africa's soil heterogeneity further hampers farmer learning about the returns to inputs, we argue that it seems unlikely that farmers cannot learn anything from their own experience. Research has shown that farmers do experiment with new technologies, but that (Bayesian) learning takes time. Therefore, farmers also learn about new technologies through peer networks (Conley and Udry, 2010; Foster and Rosenzweig, 1995). The ability to combine own experience with the experience of farmers in a similar location is therefore likely to provide a good signal about the quality of seed.³⁷

Seeing that farmers cannot perfectly assess seed quality *ex ante* or *ex post*, one could argue that a better remedy against information asymmetries would be to objectively measure seed quality (e.g., by sending mystery shoppers, fol-

³⁷To support our claim, we show that the ratings are correlated with objective indications of seed quality in Appendix 4.11.1.

lowed by DNA fingerprinting) or agro-input dealer practices (e.g., by sending objective inspectors incognito). However, these strategies are often not practical, not cost effective, or introduce other challenges (e.g., they may be prone to corruption or may not reach the most remote areas). Furthermore, peer ratings measure the dimensions of seed quality that matter most to smallholders. It is for example plausible that farmers do not mind about varietal purity (indicating whether seed is a particular variety, only detectable by DNA fingerprinting) but care a lot about seed performance (e.g., germination rate, vigor, and yield). Their preferences may be reflected in their ratings, whereas DNA fingerprinting may measure a dimension that matters less to farmers. The opinion of peers, who are familiar with the heterogeneous conditions farmers face, may be more useful and trustworthy for smallholders than the judgment of an inspector or DNA test. That is why we believe that our crowd-sourced clearinghouse strategy is an alternative worth exploring.

A training is expected to work mainly through increasing agro-input dealer knowledge, which when applied will lead to improved seed quality. An information clearinghouse is expected to work through various impact pathways. Firstly, farmers may switch from lower rated shops to higher rated shops after ratings are revealed. Secondly, dealers could anticipate this and increase their efforts to outperform their competitors. This in turn may improve quality and agro-input dealers may want to signal this to farmers. To achieve good ratings, dealers could also start offering credit, advice, or other services that may increase customer satisfaction but will not directly affect seed quality. Finally, farmers who did not buy improved maize seed before (because they believed agro-input dealer sell poor quality seed) could start adopting when they learn that agro-input dealers generally receive good ratings. Ultimately, all this is expected to increase business at the agro-input dealer level and adoption and yields at the farmer level. Note that both interventions can only work if the perceived quality issues arise at the agro-input dealer level, they will fail if poor seed quality is caused further up the seed supply chain, for example, by breeding results that are not in line with farmers' expectations.

We test the training and the clearinghouse in a randomized control trial (RCT) among 350 agro-input dealers and an associated 3,500 smallholder maize farmers in their catchment areas in eastern Uganda over the course of two agricultural seasons. We find that the information clearinghouse improves outcomes for both, dealers and farmers. Shops in areas exposed to the clearinghouse in-

intervention receive more customers, sell more, and have higher revenues from maize seed than shops in control areas, and these effects become stronger with time. Clearinghouse treated farmers are significantly more likely to use improved maize varieties from agro-input dealers, and consequently have higher yields than control farmers after two seasons. We find indications that farmers move from lower rated agro-input dealers to higher rated ones. Impact also seems to come from treated agro-input dealers who increase their efforts and expand the services that they provide to farmers. Treated shops are also more likely to be registered with the Uganda National Agro-input Dealers Association (UNADA), perhaps to signal quality. Finally, we find that farmers in the treatment group rate maize seed of shops in their neighborhood better, suggesting that the clearinghouse treatment is also effective in changing perceptions.

The agro-input dealer training does not have a clear impact on dealers nor their customers. We find no effect on knowledge as measured by a simple quiz, even though all coefficients point in the expected direction. Interestingly, we also find that the information clearinghouse does increase knowledge about proper seed storage and handling. This suggests that agro-input dealers search and find information that can help them to improve if they are incentivized. This further suggests that providing information is most effective when combined with incentives. Exploiting the factorial nature of our experimental design and zooming in on the interaction effect, we find suggestive evidence that the training does increase outcomes for the subgroup of farmers that is also exposed to the clearinghouse intervention.

Our study contributes to the literature on effective ways to reduce information asymmetries. For instance, Lane, Schonholzer, and Kelley (2022) show how providing information about bus safety to passengers affects the demand and supply of safer public transport, but only if there is a public signal (i.e., when bus drivers know that they are tracked and that this information is revealed to passengers). The article further fits into an emerging literature that tests how crowd-sourced information and reputational mechanisms can reduce information asymmetries and effectively change behavior. Even though advances in Information and Communications Technology and the rise of e-commerce has led to numerous platforms that allow for consumer feedback and a variety of websites that aggregate crowd-sourced reviews, there is surprisingly little evidence on the effects of these developments. The few rigorous studies that are

available report impressive impact. Reimers and Waldfogel (2021) compare the effects of professional critics and Amazon star ratings of books on consumer welfare and find the effect of star ratings on consumer surplus to be more than ten times the effect of traditional review outlets.

More specifically, we advance the literature on information clearinghouse mechanisms in developing countries, which have been studied to some extent, but mostly address market price information asymmetries between smallholders and middlemen. Assuming that middlemen are better informed about prevailing prices in the market than farmers, theory suggests that providing farmers with price information improves the functioning of rural agricultural markets. However, evidence is mixed: while Goyal (2010) finds that internet kiosks that provide wholesale price information significantly increase soy prices farmers receive in India, Fafchamps and Minten (2012) do not find a statistically significant effect of price information delivered to farmers' mobile phones in a neighboring state. However, a clearinghouse that relies on crowd-sourced ratings may be more effective in increasing the (perceived) quality of agricultural inputs in the market: while prices can generally be observed quite easily, assessing the performance of inputs such as seed or fertilizer is more difficult. Aggregating the experiences of many users may thus be a particularly powerful way to reveal their quality. For example, Hasanain, Khan, and Rezaee (2023) implement a crowd-sourced information clearinghouse in the market for artificial insemination of livestock in Punjab, Pakistan, where individual signals of quality are noisy. Using an RCT, they find that farmers who receive information enjoy 25% higher insemination success and no higher prices. The existing veterinarians seem to increase effort, while farmers are not likely to switch to better providers. This result illustrates how information clearinghouses can successfully aggregate information in low-capacity markets.

Our study also contributes to a large literature on the effectiveness of providing training to small businesses in developing countries. Helping entrepreneurs to grow small firms by teaching them business skills has yielded mixed results when subjected to rigorous impact evaluation methods (e.g., Karlan and Valdivia, 2011; Drexler, Fischer, and Schoar, 2014; Giné and Mansuri, 2021). While these studies often suffer from methodological issues such as lacking statistical power, it has also been argued that simply providing knowledge may be insufficient to move the needle (McKenzie and Woodruff, 2013). More promising results have emerged recently when the focus shifts from traditional trainings

to trainings designed to instill personal initiative (Campos et al., 2017).³⁸ Our study similarly shows the importance of (external) motivation in making trainings reach their objective.

4.2 Experimental design

We designed an experiment with two interventions (described in detail in the next section), and evaluate their impact using an RCT. The interventions are randomized at the catchment area level. Generally, agro-input shops are clustered in towns, villages, markets, trading centers, and other key market sheds, so that a single catchment area may be served by several dealers. If the catchment areas of two or more shops overlap because these dealers operate in the same town, street or right next to each other, they are assigned to the same catchment area and treatment. Clustering agro-input dealers into catchment areas is done on the basis of geographical location.³⁹

We randomize at catchment area level for three reasons. Firstly, randomizing at the level of the individual agro-input shop prompted ethical concerns. In cases where two or more agro-input dealers operate very close to each other, treating only one of them may lead to a competitive (dis-)advantage. Randomizing at catchment area level substantially reduces the risk of (dis-)advantaging shops. Secondly, it reduces the likelihood of spillovers from treated to control agro-input dealers. Thirdly, randomizing at catchment area level allows us to measure the effect of the treatments on farmers, as all farmers in the catchment area are now exposed to agro-input dealers who all received the same treatment.

We used simulations to determine the sample sizes for this experiment. Simulating provides a flexible and intuitive way to analyze statistical power. Furthermore, instead of relying on theoretical distributions for the outcome variables that make assumptions and return analytic solutions, we run simula-

³⁸Personal initiative is defined as a self-starting, future-oriented, and persistent proactive mindset.

³⁹We use the haversine function to construct an adjacency matrix based on GPS coordinates, and shops that are less than five kilometer apart are assigned to the same catchment area. The five kilometer threshold was selected based on a visual inspection of the map, the size of an average village and the reported distance between farmers and dealers in survey data from a previous study of the maize value chain that can be found here.

Table 4.1: Factorial design

		dealer training	
		1	0
		28 areas	28 areas
clearinghouse	1	28 areas	28 areas
	0	28 areas	28 areas

tions that (re-)sample from real data that was collected in previous surveys.⁴⁰ Power simulations show that if the number of catchment areas is larger than 112, our experiments will return statistically significant results 80% of the time on a selection of primary outcomes. This corresponds to about 318 agro-input dealers. Based on further simulations to study impacts at the farmer-household level, we decide to collect information on ten farmers per dealer, leading to a sample size of 3,180 households.⁴¹

The two interventions are combined in a field experiment which takes the form of a 2^2 factorial design. The resulting layout is illustrated in Table 4.1. We measure impact on both, agro-input dealers and farmers.

4.3 Interventions

4.3.1 Agro-input dealer training

Training content and material

To determine the content of the training and to make sure it is locally anchored, we consulted experts from different Ugandan institutions and organizations like the ministry of agriculture, the seed sector, and agro-input dealer associations. A series of semi-structured interviews and a workshop were organized. The experts identified common problems and malpractices, and then determined effective and realistic solutions and best practices in seed storage and handling. We then developed a training manual to ensure standardization and a simple but visually appealing poster illustrating the most important best practices.

⁴⁰We use data from 78 agro-input dealers and 1,529 smallholder farmers in the catchment areas of these dealers that were collected in three districts in eastern Uganda in July 2019. These surveys were part of another study of the maize value chain and can be found here.

⁴¹More detailed information can be found in the pre-analysis plan which was pre-registered at the AEA RCT registry under RCT ID 0006361.

Training

In each treated catchment area, all shops were selected to receive a training. Of each treated shop, both the owner and the shop manager who is in charge of day-to-day activities were invited. The owner was invited because some of the recommended techniques and practices require investments. The shop manager was invited because many of the recommendations are hands-on practices. We handed out one free portable seed moisture meter per shop as an incentive. All attendants were compensated for transport, lunch and drinks were provided. The training took place at a time of the year when dealers were not too busy. Trainings were organized in small groups, with on average about 10-15 agro-input dealers present. The trainings took place in locations that were easily reachable for the participants.

The trainers explained the correct handling and storage practices for improved maize seed and used the poster and an example seed bag for illustration. Afterwards the dealers rehearsed the more challenging practices like measuring moisture using a moisture meter. At the end of the training, they were asked to answer a couple of multiple choice questions. The dealers were told at the beginning of the training that receiving a moisture meter was conditional on passing this test, which might have motivated them to pay closer attention. They also received the poster as a handout which could be hung in their store.

The trainings took one day and were organized in May 2021, late enough so that dealers were not busy with selling for the first agricultural season but early enough so that they could use the newly learned practices on the seed of the second agricultural season. A timeline is illustrated in Figure 4.1. The trainings were organized together with UNADA, the national association for agro-input dealers in Uganda.

4.3.2 Information clearinghouse

Rating collection and computation

At the time of baseline data collection from smallholder farmers, we asked them to rate agro-input dealers in their proximity on a number of characteristics. Enumerators were guided by an application on a tablet computer that iterated through all agro-input dealers in the catchment area. For each dealer, we provided the common names that are used to refer to the shop, a description of where the store is located, and a picture of the store front (obtained during the

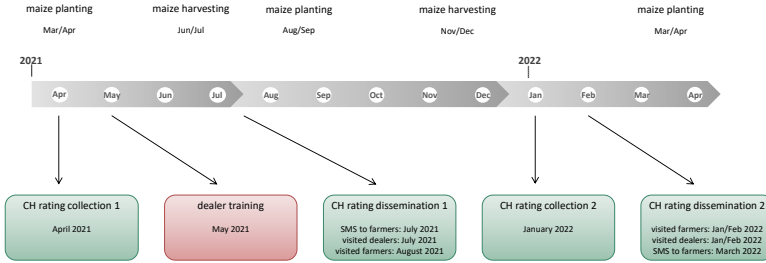


Figure 4.1: Timeline

agro-input dealer census—see Subsection 4.5.1). If farmers knew the dealer, we asked them to provide ratings using the questions which are outlined in Table 4.2. For example, we asked farmers to rate the maize seed that an agro-input shop sells on a scale of one to five stars on germination. Ratings were always collected after harvest, when smallholders were able to assess seed quality based on observing germination and yield, the resistance against droughts, pests and diseases, and how fast the seed matures, see Figure 4.1 for a timeline of the interventions.

Some may argue that by asking farmers to rate dealers, one also makes farmers aware of the existence of all dealers in the area, and that this awareness effect potentially confounds the clearinghouse effect. In the control group, we thus also iterated through dealers in the catchment areas, to make control farmers similarly aware of the existence of dealers in their vicinity. However, control farmers were not asked to rate dealers as the process of rating a dealer’s seed could make quality more salient, which we consider to be an important aspect of the treatment.

Based on the answers of all farmers about all dealers in a catchment area, we computed the ratings for each agro-input shop. These ratings were translated into words and stars for dissemination, such that they are comprehensible for farmers and dealers who are not used to interpreting numbers. As illustrated in

Table 4.2: Questions for farmers to rate dealers

	min	max
Do you know this <i>shop name</i> or <i>dealer name</i> , sometimes called <i>nickname</i> , located in <i>market name</i> ?	no	yes
The place can be described as <i>description</i> .		
Please rate this agro-input shop on:		
Quality and authenticity of seed	1 star	5 stars
Please rate the maize seed that this agro-input shop sells on:		
General quality	1 star	5 stars
Yield as advertised	1 star	5 stars
Drought tolerance as advertised	1 star	5 stars
Pest/disease tolerance as advertised	1 star	5 stars
Speed of maturing as advertised	1 star	5 stars
Germination	1 star	5 stars

Figure 4.1, there were two rounds of rating collections. However, these ratings were not pooled, meaning that the second score is independent of the first score. More details about the rating computations can be found in Appendix 4.11.2.

Rating dissemination to farmers

For the success of the clearinghouse it is crucial to disseminate the agro-input dealer ratings before farmers start buying seed for the next agricultural season, such that they can use this information when choosing whether and where to purchase inputs, see Figure 4.1. Ratings were disseminated to farmers by means of text messages and in person.

Text messages We sent farmers one text message per dealer in their proximity by Short Message Service (SMS). This message was translated into three local languages - Lusoga, Lugwere, Samia - chosen at the sub-county level to increase specificity. Table 4.19 in Appendix 4.11.3 provides more details about the messages. Also in control catchment areas, farmers received text messages with the names of dealers in their proximity, so that they were aware of the presence of these dealers. Dillon, Aker, and Blumenstock (2020) demonstrate the importance of these control messages. They introduced a “Yellow Pages” phone directory with contact information for local enterprises in central Tanza-

nia. They find that enterprises randomly assigned to be listed in the directory receive more business calls, make greater use of mobile money, and are more likely to employ workers. To separate this knowledge effect from the effect emanating from the information clearinghouse, we also disseminate control dealer information including their names but excluding the ratings. An additional advantage is that it is harder for farmers to identify if they are being treated or not, reducing the likelihood of experimenter demand effects.

In person The enumerators also re-visited the farmers in our sample. For this purpose, we designed a visually appealing dissemination application (shown to farmers on tablet computers) which cycles through all dealers in the catchment area of each farmer and states: “We wanted to let you know that customers from [name of the shop] rate the quality of maize seed sold there as okay/good/very good/excellent! The quality of the maize seed that this agro-input shop sells got a score of [score] out of 5!” in treated catchment areas. The application also showed the stars associated with the score. Again, we also cycled through dealers in control areas without providing ratings to inform farmers that these agro-dealers are operating in their neighborhood.

As it was the case for the collection of ratings, the application provides different names under which the shop is known, a description of where the store is located, and a picture of the store front to make sure farmers associate the rating with the correct shop.

Rating dissemination to dealers

Agro-input dealers received their ratings by means of a report on laminated paper which was delivered to their shops. The front shows a visually appealing certificate with a logo and the own general rating, see Figure 4.2. We encouraged agro-input dealers to display the ratings in the shop, similar to a “certificate of excellence” from TripAdvisor.

The back of the report shows more detailed information. In addition to the dealer’s general rating, it shows the separate ratings that the seed of the agro-input dealer received on the different attributes (overall quality, yield, drought and disease resistance, speed of maturing, and germination) and the average ratings of other agro-input dealers in the same catchment area in a table, visualized by stars. This shows dealers their relative position in the area and could provide an important incentive to improve.



Figure 4.2: SeedAdvisor certificate

The intervention was repeated in the course of 2022, see Figure 4.1. We expect dealers to be more likely to change their behavior if they know that the clearinghouse will remain in place for some time, so that they will be scored again. It could also increase farmers' trust in the ratings. Again, the second rating does not depend on the first rating.

4.4 Empirical strategy

Due to the random assignment to treatment and control groups, comparing outcome variable means of treated and control participants provides unbiased estimates of the effects of the interventions. Note that impact will be judged by looking at outcomes at the agro-input dealer level as well as at the farmer level. To increase power, we condition the estimates on baseline values of the outcome variables. We estimate the following specification using Ordinary Least-Squares to get the average treatment effects of both interventions:

$$Y_{ij} = \alpha + \beta T_j + \gamma' X_{ij} + \delta Y_{0ij} + \varepsilon_{ij} \quad (4.1)$$

where for dealer level outcomes, Y_{ij} is the outcome variable for dealer i in catchment area j at mid- or endline, Y_{0ij} is the corresponding outcome at baseline, T_j is a dummy for the treatment status of catchment area j , X_{ij} is a vector of controls for the orthogonal treatments in the factorial design (de-meanned and interacted with the main treatment effect, see Lin, 2013; Muralidharan, Romero, and Wüthrich, 2023), and ε_{ij} an error term that is potentially correlated withing catchment areas. The coefficient β is the estimated average treatment effect. For farmer level outcomes a similar equation is estimated, where Y_{ij} is now the outcome variable for farmer i in catchment area j at mid- or endline, and all other terms are defined as above.

Because we randomize at catchment area level, we use cluster-robust variance-covariance matrices that cluster standard errors at this level. For outcomes at the farmer level where we have almost 3,500 observations in 130 clusters, the original form of the sandwich estimator (Liang and Zeger, 1986) which does not make any small-sample correction, is used. For outcomes at the agro-input dealer level where we have almost 350 observations in 130 clusters, we approximate the leave-one-cluster-out jackknife variance estimator (Bell and McCaffrey, 2002).

In terms of variable construction, we follow some pre-registered principles. For continuous variables, trimmed values are used to reduce the influence of outliers. In particular, we trim 1% of each side of the distribution for agro-input dealer level outcomes and 2.5% of each side of the distribution for farmer level outcomes. Inverse hyperbolic sine transforms are used if variables are skewed, with skewness being defined as the adjusted Fisher-Pearson coefficient of skewness exceeding 1.96. Outcomes for which 95% of observations have the same value within the relevant sample will be omitted from the analysis.

We account for multiple hypothesis testing by aggregating different outcomes within each domain into summary indices, following Anderson (2008).⁴² However, interpreting these overall impacts can be difficult while effects on individual outcomes show us which variables drive the results and inform us about the different impact channels. That is why we also report the treatment effects on individual variables even though they have to be interpreted with care.

⁴²Each index is computed as a weighted mean of the standardized values of the outcome variables. The weights of this efficient generalized least squares estimator are calculated to maximize the amount of information captured in the index by giving less weight to outcomes that are highly correlated with each other.

4.5 Data

4.5.1 Sample

Our sample consists of agro-input dealers, and smallholder maize farmers who live in the catchment areas of these dealers. The dealer sample was obtained by listing all input shops in eleven districts in southeastern Uganda. We found 348 dealers, sufficient to detect treatment effects according to our power simulations, see Section 4.2.

After the census, these agro-input shops were assigned to 130 catchment areas (for details, see Section 4.2, Footnote 39 in particular). We find that one to 18 dealers operate in an area, with a mean of 2.7. To connect shops to customers, we asked dealers for the names of the villages where most of their customers come from. Then enumerators were instructed to randomly sample ten households that grow maize in these villages. Consequently, about 3,500 smallholder maize farmers were sampled, sufficient to detect treatment effects for the outcomes used in our power simulations (see Section 4.2).

Baseline data was collected from dealers in September and October 2020 and from farmers in April 2021. Midline data from both farmers and dealers was collected in January and February 2022, and endline data from farmers and dealers was collected in July and August 2022.

4.5.2 Descriptive statistics

This subsection describes the baseline sample. Information about the average agro-input shop can be found in Table 4.3. When enumerators approached a shop, they tried to interview the person who is most knowledgeable about the day to day operations, which was usually the shop manager. The average respondent is 32 years old. 60% are male and more than 90% finished primary education. In 55% of the cases, the respondent is also the owner of the shop.

We see substantial heterogeneity among agro-input shops. Some are small informal stores which are located in rural areas and sell maize seed only during the planting season. Others have many customers, are located in towns and specialize in inputs and equipment used in agricultural production. The average shop was established five years ago and is located seven kilometers from the nearest tarmac road. It has 41 customers per day. 74% are specialized shops which only sell farm inputs.

We also collected information that would allow us to assess the quality of maize seed sold at these agro-input shops. Enumerators asked if they could inspect the area where seed is stored and noted that there is quite some room for improvement. 65% of shops have problems with pests like rats or insects and 16% store maize seed in open containers. When we asked dealers about the services that agro-input dealers provide, around half reported to offer credit and extension or training. On the other hand, two in three shops received a complaint about seed they sold from a customer over the course of the last season.

We also purchased a bag of maize seed. However, only 232 of the 348 shops in our sample had seed in stock at the time of the baseline interview. We measured the moisture in the bag and found that it was 13.6% on average, with a minimum of 10.3 and a maximum of 17.4. Note that seed moisture content determines whether molds and storage pests thrive. It is recommended to keep moisture below 13%. While 68% of seed bags show a packaging date, only 18% show an expiry date, and 8% show a certification sticker.

Table 4.4 reports means in the farmer sample. When approaching a household, enumerators were instructed to interview the person who is most knowledgeable about maize farming. However, a set of questions deals with the household head, who could be or could not be the respondent. 78% of household heads in our sample are male, 51% have finished primary education. The average household head is 49 years old.

On average, nine people belong to one household and share three rooms. The homestead is located four kilometers from the nearest agro-input shop and nine kilometers from the nearest tarmac road. The average farmer started growing maize 23 years ago and has three acres of land for crop production.

Half of the farmers in our sample adopted improved maize seed on at least one of their plots last season. One out of three bought this seed at an agro-input shop. Only 25% applied chemical fertilizers like Di-Ammonium Phosphate (DAP) or Nitrogen, Phosphorus, and Potassium (NPK) on a randomly selected maize plot. Productivity is low at about 440 kilograms per acre.

4.5.3 Orthogonality tests of randomization balance

To test if treatment and control groups are comparable in terms of a set of baseline characteristics we include standard orthogonality tables with pre-registered

variables for both dealers and farmers (Tables 4.3 and 4.4 respectively). Some of these characteristics are unlikely to be affected by the intervention, while others are picked from the outcome variables we will use to measure the impact of our interventions and explore impact pathways in the next sections.

For outcomes at the agro-input dealer level reported in Table 4.3, we find that from a total of 16 comparisons, only one is significant at the 5% level for the agro-input dealer training. For the clearinghouse treatment, we find two significant differences, both at the 10% level. This is consistent with a balanced sample. For outcomes at the farmer level, out of 32 comparisons, one is significant at the 10% level.

4.6 Results

This section presents results on the impact of the agro-input dealer training and the information clearinghouse. We report effects at the agro-input dealer level as well as at the level of the farmers that reside in catchment areas of the dealers. Furthermore, we separately present impact one agricultural season after the intervention (referred to as impact at midline) and two seasons after the intervention (referred to as impact at endline).

We take transparency and replicability seriously. All outcome variables have been registered in a pre-analysis plan which can be found in the American Economic Association (AEA)’s registry for RCTs. In addition to the pre-analysis plan, we completed the entire econometric analysis on simulated data in a mock report and added it to the AEA’s registry before midline data was collected.⁴³ Mock reports are dynamic documents that integrate all code.⁴⁴ As such, when midline and endline data became available, we simply replaced the simulated data with the real data. All documents, code, and data are under revision control and publicly accessible via GitHub which provides time-stamped records

⁴³Mock reports serve to further tie the hands of researchers, reducing their freedom in choosing which specifications and variables to select when testing hypotheses. Humphreys, De la Sierra, and Van der Windt (2013) argue that mock reports can reduce intentional and unintentional fishing, and make published research more reliable.

⁴⁴We use the knitr engine to integrate R code in L^AT_EX (Xie, 2017).

Table 4.3: Descriptive statistics and orthogonality tests - Agro-input dealer

	mean	training	CH
Respondent's age in years	32.43 (11.49)	0.56 (1.19)	-2.24 ⁺ (1.21)
Respondent is male	0.59 (0.49)	0.02 (0.06)	-0.01 (0.06)
Respondent finished primary education	0.92 (0.27)	0.01 (0.03)	-0.01 (0.03)
Respondent owns shop	0.55 (0.50)	0.03 (0.06)	0.02 (0.06)
Respondent received training on maize seed handling	0.53 (0.50)	0.05 (0.07)	0.12 ⁺ (0.07)
Respondent knows how to store seed after repackaging	0.27 (0.44)	0.07 (0.06)	0.08 (0.06)
Shop's distance to nearest tarmac road in km	6.56 (10.39)	-0.92 (2.21)	-1.58 (2.24)
Shop only sells farm inputs	0.74 (0.44)	-0.09 (0.07)	0.03 (0.06)
Years since shop establishment	5.34 (6.30)	-0.09 (0.77)	0.21 (0.78)
Number of customers per day	41.49 (46.49)	11.35 (7.16)	6.43 (6.72)
Quantity of maize seed sold in kg	695.50 (1497.18)	201.06 (252.97)	176.31 (235.92)
Amount of maize seed lost/wasted last season in kg	3.50 (18.65)	1.99 (2.47)	2.40 (2.30)
Shop has problem with pests	0.65 (0.48)	-0.01 (0.06)	-0.03 (0.06)
Shop stores maize seed in open containers	0.16 (0.36)	0.00 (0.05)	0.08 (0.05)
Shop received seed related complaint from customer	0.64 (0.48)	-0.11* (0.05)	0.07 (0.05)
Moisture in bag of maize seed in %	13.56 (1.44)	0.25 (0.25)	-0.18 (0.26)

Note: Column (1) reports sample means at baseline and standard deviations below; columns (2)-(3) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; **, *, and + denote significance at the 1, 5 and 10% levels.

Table 4.4: Descriptive statistics and orthogonality tests - Farmer

	mean	training	CH
Household head's age in years	48.62 (13.38)	-0.08 (0.56)	-0.24 (0.56)
Household head is male	0.78 (0.42)	-0.02 (0.03)	0.03 (0.03)
Household head finished primary education	0.51 (0.50)	0.00 (0.03)	0.04 (0.03)
Homestead's distance to nearest tarmac road in km	9.39 (10.81)	0.33 (1.69)	-1.23 (1.71)
Homestead's distance to nearest agro-input shop in km	3.78 (4.79)	-0.11 (0.37)	0.11 (0.37)
Number of people in household (incl. respondent)	8.70 (3.98)	-0.16 (0.18)	-0.09 (0.18)
Number of rooms in house	3.49 (1.45)	-0.01 (0.09)	0.02 (0.09)
Farmer's land for crop production in acres	3.35 (4.32)	0.07 (0.21)	0.00 (0.22)
Years since farmer started growing maize	23.09 (13.14)	0.61 (0.55)	-0.55 (0.58)
Yield in kg/acre	443.01 (304.99)	27.15 ⁺ (13.71)	-6.14 (13.52)
Farmer used quality maize seed on any plot	0.49 (0.50)	0.02 (0.02)	0.01 (0.02)
Farmer bought this seed at agro-input shop	0.32 (0.47)	-0.01 (0.02)	0.01 (0.02)
Amount of this seed farmer bought at agro-input shop in kg	9.52 (6.92)	0.16 (0.53)	-0.34 (0.53)
Farmer thinks maize seed at agro-input shops is adulterated	0.68 (0.46)	0.01 (0.03)	0.00 (0.03)
Farmer used DAP/NPK	0.25 (0.43)	0.04 (0.03)	0.02 (0.04)
Farmer used organic manure	0.07 (0.26)	-0.01 (0.01)	0.01 (0.01)

Note: Column (1) reports sample means at baseline and standard deviations below; columns (2)-(3) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; **, *, and + denote significance at the 1, 5 and 10% levels.

of all changes made over the course of the project.⁴⁵

The results Tables 4.5 to 4.15 have a common layout. Column (1) provides baseline sample means with standard deviations in parentheses below, mainly to get an idea of effect sizes. That is why we always provide these averages in levels, even though we may report the difference between treatment and control group after using inverse hyperbolic sine transforms. In column (2), we provide the average treatment effect of the agro-input dealer training at midline, while column (3) reports the average treatment effect of the information clearinghouse treatment at midline. We also report the standard errors below in parentheses and the number of observations that were used for the regressions at midline in column (4). Column (5) and (6) report average treatment effects of the training and the clearinghouse treatment respectively, but now at endline, after two seasons. Also here, we report the number of observations that was used in the estimation in column (7). As mentioned in Section 4.4, we account for multiple hypothesis testing by aggregating different outcomes within families into overall summary indices, following Anderson (2008). Results for these indices are reported at the bottom of the tables.⁴⁶

4.6.1 Impact on agro-input dealers

We start by testing if the interventions have an impact on general business operations of agro-input dealers in Table 4.5. Sales volume and price, revenue, and number of customers and maize varieties in stock are included as outcomes

⁴⁵The presentation of results in this paper differs somewhat from the way it was pre-registered and presented in the mock report (and midline report, endline report, and previous versions of this manuscript). The pre-registered reports mainly serve to tie our hands, commit to decisions, and reduce our freedom in choosing which specifications and variables to select when testing hypotheses, to avoid fishing and make this study more reliable (Humphreys, De la Sierra, and Van der Windt, 2013). Hence the presentation of results in the pre-registered reports mainly determines if the interventions worked or not. As such, we included tables of key outcomes along the entire causal chain (and combined them in an index) to assess overall treatment impact, see Appendix 4.11.4. Other tables then went into detail (e.g., focusing on a particular seed type sold by an agro-input dealer or on a particular plot cultivated by a farmer). In this paper however, we reorganize the reporting to match a structure where we first look at impact on outcomes at the end of the causal chain and then look at impact on intermediate outcomes to explore potential mechanisms. While this change does affect the construction of some of the indices, overall conclusions remain the same. Some of the pre-registered tables can be found in Appendix 4.11.4. The entire analysis that follows the pre-registered structure can be found through the project history in GitHub, for instance here.

⁴⁶In the regressions with these overall indices, we do not control for the baseline values because this would imply having the result only for dealers and farmers who have no missing values for any of the variables constituting these indices at mid-/endline and at baseline, severely reducing statistical power.

under this heading. A measure of sales volume was constructed by asking how much of a specific maize seed variety dealers sold in the previous season. We restrict attention to the four most popular improved varieties, two of which are hybrid varieties (Longe 7H and Longe 10H) and two of which are Open Pollinated Varieties (OPVs) (Longe 4 and Longe 5). Total quantity sold is the sum of quantities sold of these four varieties. We also asked dealers about the sales price of the four seed types at the start of the season and then calculated the simple average at the dealer level. We then calculate the revenue (expressed in million UGX) by first multiplying prices with quantities sold and then summing over the four seed types. We also include the number of customers that buy maize seed on an average day at the start of the season, as well as the number of maize varieties that the agro-input dealer has in stock.

Table 4.5 shows that we do not find an impact of training agro-input dealers on their business operations. Both at midline and at endline, the index is not significantly different from zero. Looking at the individual outcomes, there is no consistent pattern. At midline, we find a negative impact of the training on the average sales price. At endline, the training seems to have reduced amounts sold, which is also reflected in a lower revenue.

However, we do find a positive impact of the information clearinghouse intervention on agro-input dealer operations. At midline, the overall index is significantly higher among dealers in the clearinghouse treatment group. Looking at individual outcomes, we see that treated dealers sold more maize seed at a higher price, albeit not significantly so. However, in combination, this leads to revenues that are almost 20% higher (and this difference is significant at the 10% level).⁴⁷ At endline, the positive effect of the clearinghouse intervention seems to become stronger, with the overall index now being significant at the 1% level. The effect is driven by a 31% increase in the number of customers that a treated shop attracts, which translates into six additional customers.

While Table 4.5 looks at the impact on overall business operations, Tables 4.6 and 4.7 focus on the effect of the interventions on operations related to one particular seed type. We look at the most recently released hybrid variety (Longe 10H) in Table 4.6 and the most recent OPV (Longe 5) in Table 4.7.

In line with Table 4.5, we start by looking at sales volumes, prices, and revenues. We also focus on outcomes related to stock management, as seed

⁴⁷For reasonably large values, coefficients of regressions that involve a dependent variable that has been transformed using the inverse hyperbolic sine can be interpreted as elasticities (Bellemare and Wichman, 2020).

Table 4.5: Effects on agro-input dealer outcomes: Operations

	<i>baseline</i>		<i>midline</i>		<i>endline</i>	
	mean		training	CH	training	CH
Quantity of maize seed sold in kg ^{§†}	695.503 (1497.183)		-0.092 (0.220)	0.284 (0.227)	-0.499 ⁺ (0.250)	0.239 (0.253)
Sales price of maize seed in UGX/kg [†]	4273.897 (955.073)		-192.784 ⁺ (114.934)	99.272 (113.292)	-33.867 (143.152)	145.861 (138.816)
Revenue from maize seed in mln UGX ^{§†}	2.890 (6.286)		-0.069 (0.104)	0.185 ⁺ (0.108)	-0.227 ⁺ (0.118)	0.143 (0.118)
Number of maize seed customers per day ^{§†}	19.764 (20.689)		-0.056 (0.098)	0.127 (0.101)	-0.190 (0.116)	0.310 ^{**} (0.112)
Number of maize varieties in stock [†]	2.834 (1.589)		0.042 (0.266)	0.245 (0.245)	-0.216 (0.234)	0.221 (0.220)
Overall index	0.031 (0.610)		-0.130 (0.095)	0.197 [*] (0.092)	-0.131 (0.086)	0.238 ^{**} (0.082)
Max. number of obs.				306		297

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
§Due to the skewness of this variable, the regression was run after an Inverse Hyperbolic Sine transformation. Coefficient estimates can therefore be interpreted as percentage changes. The baseline mean column shows the untransformed variable.

quality decreases with shelf-life. We asked agro-input dealers how much of the particular seed was carried over from the previous season. Many dealers reported that they did not carry over any seed, leading to low baseline means. Furthermore, we asked them to estimate how much they bought from any provider during the same season. For both varieties, this is slightly more than what dealers reported to have sold. We expect our treatments to decrease the amount of seed carried forward and increase the amount of fresh seed procured from providers. We also asked agro-input dealers to provide an estimate of how much of the seed stock they lost or wasted during the season, and how often they ran out of stock. We expect the interventions to reduce both, losses and stock-outs.

For the variety Longe 10H, we do not find significant effects of the training nor the clearinghouse treatment at midline. At endline however, all individual coefficient estimates go in the expected direction for the information clearinghouse, and when outcomes are combined in an index, the effect is positive and significant at the 1% level. Results are similar for Longe 5.

One step further up the impact chain, increased numbers of customers, sales, and revenues are likely to be driven by an increase in the (perceived) quality of improved maize seed that these agro-input dealers sell. Unfortunately, the quality of seed is hard to assess, which is one of the key reasons why information asymmetries exist and the clearinghouse intervention was implemented. Nevertheless, to get an idea of the quality of seed sold, we bought a bag of seed at each dealer and inspected it on a number of attributes. First, we measured the moisture content of the seed. In Table 4.8, we see that the clearinghouse treatment reduced moisture as expected, but the parameter is estimated imprecisely, perhaps due to the smaller data set as we were not able to source seed from all dealers and the comparisons were only made for shops in which the enumerator was able to buy a bag of maize seed at mid- or endline. We further look at the integrity of the package and whether it shows important information such as the packaging date and the lot number, but also at the shelf-life and whether seed is in the original bag without any signs of damage. We do not find that the clearinghouse treatment nor the agro-input dealer training affected quality proxies of the seed that agro-input dealers sell. However, because our proxies of

Table 4.6: Effects on agro-input dealer outcomes: Operations - Longe 10H

	<i>baseline</i>		<i>midline</i>		<i>endline</i>		
	mean	training	CH	obs.	training	CH	obs.
Quantity sold in kg ^{‡†}	288.384 (727.049)	0.050 (0.206)	0.236 (0.204)	256	-0.205 (0.231)	0.352 (0.239)	242
Sales price in UGX/kg ^{‡†}	9.417 (0.145)	-0.025 (0.026)	-0.013 (0.026)	194	-0.019 (0.030)	0.010 (0.029)	187
Revenue in mln UGX ^{‡†}	1.625 (3.839)	0.008 (0.119)	0.130 (0.123)	255	-0.106 (0.130)	0.173 (0.136)	241
Amount carried over in kg ^{‡†}	2.679 (12.137)	-0.186 (0.212)	0.090 (0.215)	262	-0.012 (0.138)	-0.034 (0.134)	250
Amount shop bought from provider in kg ^{‡†}	294.672 (741.810)	0.118 (0.218)	0.206 (0.213)	257	-0.022 (0.250)	0.283 (0.253)	243
Amount lost/wasted in kg ^{‡†}	0.036 (0.405)	-0.001 (0.093)	0.019 (0.097)	257	-0.058 (0.037)	-0.038 (0.041)	243
Number of times per month shop ran out ^{‡†}	1.039 (1.575)	-0.236 ⁺ (0.129)	-0.045 (0.133)	192	-0.180 (0.128)	-0.205 (0.136)	185
Overall index	0.080 (0.437)	0.030 (0.067)	0.029 (0.070)	244	0.021 (0.052)	0.217 ^{**} (0.057)	233
Max. number of obs. ¹							254

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

[‡]Due to the skewness of this variable, the regression was run after an Inverse Hyperbolic Sine transformation. Coefficient estimates can therefore be interpreted as percentage changes. The baseline mean column shows the untransformed variable.

¹The comparisons were only made for shops which had Longe 10H in stock at mid- or endline.

Table 4.7: Effects on agro-input dealer outcomes: Operations - Longe 5

	<i>baseline</i>		<i>midline</i>		<i>endline</i>	
	mean		training	CH	training	CH
Quantity sold in kg ^{§†}	389.492 (716.556)		-0.040 (0.222)	0.304 (0.216)	-0.215 (0.234)	0.316 (0.230)
Sales price in UGX/kg ^{§†}	8.730 (0.110)		0.017 (0.016)	-0.015 (0.016)	-0.002 (0.022)	0.013 (0.022)
Revenue in mln UGX ^{§†}	1.193 (2.175)		0.019 (0.099)	0.111 (0.096)	-0.080 (0.100)	0.114 (0.105)
Amount carried over in kg ^{§†}	4.312 (19.088)		0.247 (0.324)	-0.092 (0.306)	-0.095 (0.148)	-0.004 (0.155)
Amount shop bought from provider in kg ^{§†}	431.451 (803.696)		-0.005 (0.221)	0.253 (0.215)	-0.179 (0.232)	0.289 (0.235)
Amount lost/wasted in kg ^{§†}	1.756 (10.173)		-0.150 (0.128)	0.031 (0.128)	-0.055 (0.055)	-0.033 (0.058)
Number of times per month shop ran out ^{§†}	0.839 (1.509)		0.053 (0.100)	0.086 (0.101)	0.094 (0.120)	-0.054 (0.126)
Overall index	0.039 (0.401)		0.037 (0.068)	0.012 (0.062)	-0.038 (0.058)	0.152* (0.058)
Max. number of obs. ¹						269

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
§Due to the skewness of this variable, the regression was run after an Inverse Hyperbolic Sine transformation. Coefficient estimates can therefore be interpreted as percentage changes. The baseline mean column shows the untransformed variable.
1The comparisons were only made for shops which had Longe 5 in stock at mid- or endline.

seed quality are far from perfect⁴⁸ and we rely on a smaller sample, we cannot safely conclude that the treatments did not affect seed quality.

4.6.2 Impact on smallholder farmers

We first look at harvest related outcomes for farmers that live in the catchment areas of agro-input dealers involved in our study and report the results in Table 4.9. We start by looking at overall production, the area of a specific maize plot, and the production scaled by plot size, i.e., yield. We also look at market participation (amount sold, sales price, and revenue from maize sales) and how much grain farmers save to use as seed in the next season. All these outcome variables deal with one randomly selected plot. While we expect positive effects on harvest and sales, the amount kept as seed enters the index negatively.

The coefficient estimates for the overall index show no effect of the agro-input dealer training, and a positive effect of the information clearinghouse, but only after two seasons of implementation. Farmers that live in areas where the clearinghouse was implemented report higher production and productivity at endline. Yield differences are significant at the 1% level and amount to 10% compared to the baseline means. Finally, we look at the amount of maize that farmers keep to use as seed in the next season. At midline, we see that, in line with expectations, clearinghouse treated farmers save less grain for seed.

To explore the large effect of the clearinghouse on yields, we investigate which subgroup of farmers enjoyed most yield gains. Rerunning the regression only for farmers who did not adopt at baseline led to a coefficient of 56.44 with a standard error of 17.38 (hence, significance at the 1% level). For farmers that did adopt at baseline, we find a coefficient of 30.79 with a standard error of 20.38 (hence, no significance). This indicates that the effect is driven by farmers who did not adopt at baseline, started adopting due to the clearinghouse, and in turn enjoy higher yields. Assuming that non-adopting farmers are the

⁴⁸As elaborated in a recent World Bank Blog, assessing seed quality is not only challenging for farmers, but also for researchers (Beegle et al., 2021). The most important seed quality dimensions are analytical purity (indicating whether seed is the correct species), germination rate (indicating whether seed germinates) and varietal purity (indicating whether seed is the correct variety, for example, a particular variety with specific traits, only detectable by DNA fingerprinting). For example, Barriga and Fiala (2020) use laboratory tests to investigate the DNA variation (indicating how genetically similar a sample of seed is to itself), analytical purity, and performance of seeds (germination rate, moisture, and vigor) as their measure of seed quality. Even though investigating moisture is an attempt to test the quality of seeds in a quantitative and objective way, this variable is one-dimensional and only a weak indication of seed quality.

Table 4.8: Effects on agro-input dealer outcomes: Bag of maize seed

	<i>baseline</i>		<i>midline</i>		<i>endline</i>	
	mean	training	CH	obs.	training	CH
Moisture in % [†]	13.564 (1.482)	0.017 (0.142)	-0.122 (0.144)	175	-0.041 (0.198)	-0.220 (0.197)
Bag shows packaging date [†]	0.689 (0.464)	0.053 (0.069)	0.050 (0.072)	179	-0.091 (0.063)	0.035 (0.064)
Shelf-life in days ^{1†}	60.951 (40.960)	-18.930 (22.091)	-8.272 (20.869)	164	13.091 (8.243)	6.352 (8.289)
Seed is in original undamaged bag [†]	0.940 (0.238)	0.025 (0.044)	0.002 (0.046)	179	0.006 (0.053)	0.051 (0.055)
Bag shows lot number [†]	0.508 (0.501)	0.025 (0.106)	-0.001 (0.107)	179	-0.138* (0.062)	0.027 (0.064)
Overall index	0.065 (0.364)	0.083 (0.103)	0.108 (0.103)	160	-0.067 (0.094)	0.108 (0.090)
Max. number of obs. ²				179		265

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

¹Days since the packaging date or, if the bag does not show the packaging date, days since the expiry date minus 6 months.

²The comparisons were only made for shops in which the enumerator was able to buy a bag of maize seed at mid- or endline. Also, we do not control for the baseline values of the outcome variables in the entire table because only 144 of the 179 dealers who had seed at midline also had seed at baseline and only 183 of the 265 dealers who had seed at endline also had seed at baseline, so that controlling for baseline values would reduce the sample sizes drastically.

less advantaged ones, the impact on their adoption and yield is particularly desirable.

To further test this, we look at the adoption of improved maize seed varieties as a second important family of outcomes at the smallholder level. For the agro-input dealer training, we do not find any effect at midline, nor at endline. The effect of the information clearinghouse treatment on overall adoption as measured by the index is positive and significant at the 5% level at mid- and endline.

Zooming in on individual outcomes, a first key question simply asks farmers if they used quality maize seed for any plot in the previous season. Here, the definition of “quality maize seed” is up to the farmer. We see that at midline, farmers that were subjected to the clearinghouse treatment were 3.5 percentage points more likely to answer this question with “yes” than control farmers. After two agricultural seasons, the difference between treatment and control farmers increases to 4.2 percentage points. Related, we ask if farmers bought maize seed at an agro-input shop for any plot. At midline, we find a difference between the clearinghouse treatment and control groups of about 6 percentage points, and this amounts to an almost 20% increase relative to the baseline mean. At endline, the difference is about 3 percentage points, but not significant anymore. We do not find an impact of the clearinghouse on the amount that farmers bought at agro-input shops. However, one should note that the sample size is smaller because we only ask this question to farmers who bought from agro-input dealers.

We then ask questions about adoption on a randomly selected plot. For the adoption of hybrid or open-pollinated maize varieties, we estimate positive treatment effects of the clearinghouse, but the coefficients are not significantly different from zero.⁴⁹ As for the more general questions above, we also ask if the seed that was used on the random plot was obtained from an agro-input dealer. We find an almost 5 percentage point treatment effect for the clearinghouse at midline and a 3.6 percentage point effect at endline. A related question asks if farmers used farmer-saved seed on the randomly selected plot. Again in line with expectations, we find that farmers that were exposed to the clearinghouse treatment reduced the use of farmer-saved seed, albeit only significantly so at

⁴⁹Here, we asked farmers which variety they planted in the previous season. If a farmer used Longe 10H, Longe 7H, Longe 7R/Kayongo-go, Bazooka, Longe 6H, Longe 5/Nalongo, Longe 4, Panner, Wema, KH series, or other hybrid/OPV, and this seed was not recycled or farmer-saved but newly purchased, it counted as hybrid/open-pollinated maize seed.

Table 4.9: Effects on farmer outcomes: Harvest on specific maize plot

	<i>baseline</i>		<i>midline</i>		<i>endline</i>	
	mean	training	CH	obs.	training	CH
Production in kg [†]	463.203 (399.595)	-0.806 (14.050)	-20.372 (14.529)	2884	16.959 (17.957)	43.937* (17.922)
Area in acres	1.094 (0.655)	-0.013 (0.029)	-0.003 (0.029)	3004	0.000 (0.032)	0.006 (0.038)
Yield in kg/acre [†]	443.222 (304.964)	-12.216 (16.234)	-23.006 (16.964)	2878	5.118 (15.596)	56.436** (17.382)
Amount sold in kg ^{§†}	195.295 (297.545)	-0.046 (0.126)	-0.201 (0.124)	3063	-0.147 (0.159)	0.173 (0.173)
Sales price in UGX/kg	506.954 (139.389)	-7.787 (14.395)	33.027* (14.244)	610	-47.215 (30.547)	12.614 (41.238)
Revenue in UGX ^{§†}	97.783 (156.538)	-0.141 (0.260)	-0.393 (0.257)	3058	-0.354 (0.341)	0.355 (0.363)
Amount kept as seed in kg [§]	14.506 (18.530)	-0.098 (0.092)	-0.188* (0.092)	2931	-0.043 (0.108)	0.036 (0.104)
Overall index	-0.020 (0.784)	-0.015 (0.039)	-0.061 (0.039)	2932	0.018 (0.041)	0.097* (0.041)
Max. number of obs.				3407		
				3441		

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
§Due to the skewness of this variable, the regression was run after an Inverse Hyperbolic Sine transformation. Coefficient estimates can therefore be interpreted as percentage changes. The baseline mean column shows the untransformed variable.

midline. Finally, we look at the product of the amount and the price of maize seed, i.e., the total cost of seed on that plot. We see that in areas where the clearinghouse was implemented, farmers invest significantly more in seed.

4.7 Causal chain and mechanisms

The information clearinghouse is a unique intervention because it solves a variety of potentially interlinked problems simultaneously. If the quality of maize seed at agro-input shops is sufficient but some farmers think that dealers provide sub-standard quality, a clearinghouse may correct their perceptions. If the quality of seed differs between agro-input dealers, it provides farmers with information that may help them to switch to dealers that provide better products. Furthermore, the rating system is expected to provide a direct incentive to agro-input dealers to stay ahead of immediate competitors. For the agro-input dealer training, the underlying mechanism is increased knowledge. In this section, we investigate the relative importance of these different impact pathways.

4.7.1 Dealer knowledge

The primary mechanism underlying the agro-input dealer training is learning, which is in turn expected to increase knowledge of treated dealers. To test if the interventions affect knowledge at the agro-input dealer level, we construct two indices that summarize different measures of knowledge. The first index aims to measure knowledge about seed storage and handling and tests if dealers retain the information that was provided during the training, using a short multiple choice quiz of five questions. The questions test knowledge related to seed carryover between agricultural seasons, how seed should be stored after repackaging, how seed should be stored in the storeroom, and whether seed should be repackaged. The questions, the options presented to farmers, and the correct answers are outlined in Appendix 4.11.5 and all variables constituting this index are self-reported.

The second knowledge index does not focus on seed handling recommendations, but aims to capture knowledge about seed more broadly. We again use multiple choice questions to test if dealers know which seed variety to recommend if a farmer complains about poor soil or lack of rain, if a farmer is late

Table 4.10: Effects on farmer outcomes: Adoption

	baseline		midline		endline	
	mean	training	CH	obs.	training	CH
Farmer used quality maize seed on any plot [†]	0.492 (0.500)	-0.021 (0.020)	0.035 ⁺ (0.020)	3206	-0.009 (0.020)	0.042* (0.020)
Farmer bought maize seed at agro-input shop for any plot [†]	0.325 (0.468)	-0.014 (0.021)	0.059** (0.021)	3145	0.004 (0.019)	0.031 (0.020)
Amount of this maize seed farmer bought at agro-input shop in kg	9.519 (6.920)	0.512 (0.348)	-0.105 (0.358)	599	0.457 (0.419)	0.378 (0.431)
Farmer used hybrid/open-pollinated maize seed on specific plot ^{1†}	0.432 (0.495)	-0.019 (0.023)	0.035 (0.023)	2954	0.009 (0.023)	0.030 (0.023)
Farmer bought maize seed at agro-input shop for specific plot [†]	0.330 (0.470)	-0.010 (0.022)	0.047* (0.022)	3153	0.012 (0.019)	0.036 ⁺ (0.019)
Farmer used farmer-saved maize seed on specific plot	0.579 (0.494)	0.020 (0.022)	-0.042 ⁺ (0.022)	3153	-0.009 (0.020)	-0.016 (0.020)
Cost of maize seed used on specific plot in UGX ^{‡†}	14078.272 (24654.685)	-0.181 (0.235)	0.499* (0.235)	2848	0.283 (0.208)	0.350 ⁺ (0.209)
Overall index	-0.013 (0.899)	-0.030 (0.043)	0.087* (0.042)	2854	0.015 (0.039)	0.086* (0.039)
Max. number of obs.				3407		3441

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

¹For this variable, only non-recycled (newly purchased, not farmer-saved) seed counted hybrid/open-pollinated seed.

[‡]Due to the skewness of this variable, the regression was run after an Inverse Hyperbolic Sine transformation. Coefficient estimates can therefore be interpreted as percentage changes. The baseline mean column shows the untransformed variable.

for planting, and whether they know what to tell clients who inquire about the yield benefits of hybrid seed. Again, the questions and (correct) answer options are explained in Appendix 4.11.5 and all variables in this index are self-reported.

Table 4.11 suggests a positive impact of the agro-input dealer training on knowledge at midline, but the coefficient is just not significant at the 10% level. The (insignificant) effect of the training is strongest at midline, which seems reasonable as the training was organized only once at the start of the study, see Figure 4.1. Interestingly, we find knowledge effects from the clearinghouse treatment, particularly when it comes to dealer knowledge related to seed storage seed and handling. This effect becomes stronger over time, which again seems reasonable as this treatment is repeated over several agricultural seasons. As agro-input dealers become aware of the recurrent nature of the ratings, they may try to improve the quality of their products by searching for information on better ways to store and handle seed.

The above suggests that providing only knowledge through trainings is unlikely to improve outcomes. However, if dealers have incentives to improve quality, they may respond by actively seeking out information. Together, this suggests that extra knowledge is only useful if agro-input dealers are also motivated to put it into practice. To further look into this, we exploit the factorial design, where a random subset of agro-input dealers was assigned to both, the training and the clearinghouse treatment group. We indeed find a significant positive interaction effect on key outcomes at the agro-input dealer level at endline (most notably on the overall operations index, as in Table 4.5, and the overall index of primary dealer outcomes, as in Table 4.20). The positive interaction effect seems to be driven by significant improvements in efforts and practices (as in Table 4.12). This confirms that if a clearinghouse encourages dealers to excel, they are more likely to put the seed handling practices that they were taught during the training into practice. If we look at knowledge in particular, we also find an interesting interaction effect on the index of dealer knowledge about seed at midline, but when we correctly cluster at catchment area level, the effect is just not significant ($p\text{-value} = 0.14$). However, these findings should be interpreted with care as we are likely to be under-powered to estimate interaction effects.

In addition to the hypothesis that trainings are less effective if dealers do not have incentives to learn, social desirability bias could be another expla-

Table 4.11: Effects on agro-input dealer outcomes: Knowledge

	<i>baseline</i>		<i>midline</i>		<i>endline</i>	
	mean		training	CH	training	CH
Index of dealer knowledge about seed storage ^{1†}	0.000 (0.482)		0.091 (0.076)	0.115 (0.075)	0.030 (0.053)	0.124* (0.055)
Index of dealer knowledge about seed ^{2†}	0.000 (0.533)		0.102 (0.072)	0.065 (0.070)	-0.009 (0.080)	-0.007 (0.078)
Overall index	0.000 (0.729)		0.208 (0.125)	0.211 ⁺ (0.119)	0.038 (0.107)	0.142 (0.102)
Max. number of obs.				306		297

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
¹The index of dealer knowledge about seed storage contains 5 variables: whether dealer knows how long seed can be carried over, how seed should be stored after repackaging, what the min. distance between floor and seed is, how seed should be stored in storeroom, whether seed should be repackaged.
²The index of dealer knowledge about seed contains 4 variables: whether dealer knows which seed variety to recommend if farmer complains about poor soil, if farmer complains about little rain, if farmer is late for planting, what to tell clients about yield benefits of hybrid seed.

nation why we do not find stronger effects of the training on knowledge, as all variables which constitute the knowledge indices are self-reported. If dealers strategically do something incorrectly, they might report that they do not know how to do it, as being uninformed is more socially desirable than making mistakes intentionally. For example, if dealers often repack seed even though they know that they should not, and the enumerator is aware of that because repacked seed is laying on the counter for sale, they may report that they agree most with the statement “You should repackage all your seed to visually verify that you are selling good quality seed” instead of “You should avoid repackaging your seed as much as possible.” However, we do not expect this to be a problem for the second Index of dealer knowledge about seed, as it mainly includes questions related to which maize varieties dealers would recommend under specific circumstances and what they tell clients about the yield benefits of hybrid seeds. We do not see why dealers would answer these questions incorrectly if they know the correct answer, as neither enumerators nor anyone else would notice if there is a difference between the answer they provide to us and the information they give clients, for example, about the yield benefits of hybrid seeds. As we do not have any measures of knowledge that are not self-reported, we cannot provide an analysis that is robust to social desirability bias.

4.7.2 Dealer efforts, services, and practices

The information clearinghouse provides agro-input dealers with an incentive to become better than their direct competitors. To do so, they may increase effort and start providing more or better services to get higher ratings, hoping that they will retain or even increase their number of customers.

In Table 4.12, we provide evidence that agro-input dealers that are exposed to the clearinghouse indeed invest more effort than dealers in the control group. The table shows results for one overall index and four individual indices that each capture different dimensions of efforts, services, and practices. A first index focuses on effort and service provision as reported by agro-input dealers themselves. It is composed of seven different variables: whether dealers offer explanations on how to use improved seed, recommend complementary inputs to get optimal results from improved varieties, provide extension or training, offer discounts for larger quantities, offer credit, received a seed related cus-

tomer complaint since last season, and accept mobile money. A second index summarizes the perceptions of farmers that are customers at these agro-input dealers. This index is also constructed from seven variables: whether a shop offers refunds or insurance, provides credit, offers training or advice to customers, delivers to the farm-gate, provides after-sales service, accepts different payment methods, and sells small quantities. The answers of farmers are aggregated at the dealer level before the index is computed.

To handle and store seed correctly, a combination of investments and labor-intensive practices is necessary. Also during the agro-input dealer training, we recommended a mix of practices that are in reach of different types of dealers, some of which may have excess labor while others may have access to money to invest. A third index groups a set of labor-intensive seed handling and storage practices. It contains six variables: whether seed is stored in a dedicated area, in correct lighting, on appropriate surface, and not in open containers, whether the shop has no pest problem, and a cleanness and professionalism rating provided by the enumerator. All of these variables were collected or at least confirmed by enumerators through visual inspection, none of them is self-reported. A fourth index summarizes some capital-intensive practices, including these six variables: whether the roof is leak-proof, and insulated, whether the walls are insulated, whether the shop is ventilated, displays any official certificate, and whether expired seed is handled correctly. Most of these variables were collected or at least confirmed by enumerators through visual inspection, only one of them (whether expired seed is handled correctly) is self-reported.⁵⁰

We find that the clearinghouse intervention increases dealer efforts and services, especially at midline, where the coefficient of the overall index is significant at the 1% level. This effect is driven by treated agro-input dealers who significantly raised their efforts and services, according to farmers. We see that impact persists until endline, where the significant effect on the overall index seems to be driven by the self-reported measure of effort. We do not find that the agro-input dealer training improved services or practices.

Furthermore, we asked enumerators if they saw the SeedAdvisor certificate in the shop at endline. The answer to this question does not contribute to any of the outcome indices because it only showed up in interviews with clearinghouse

⁵⁰To check whether social desirability affects this result, we exclude the variable from the index and rerun the analysis, as dealers might report that they handle expired seed correctly but strategically not do so, for example, to cut costs. Doing this does not change the coefficients for the Index of capital-intensive seed handling practices in any notable way.

treated dealers. For this subgroup, 55% of enumerators reported that they saw it.

As it is difficult to assess seed quality via visual inspection, agro-input dealers may use various strategies to signal to customers that their products are of good quality. Becoming a member of professional organizations is one way to do so, as these memberships signal professionalism. Dealers who try to signal quality will also not shy away from inspections. On the contrary, they may actively seek inspection so that they can advertise the result in their shops. Alternatively, the SeedAdvisor certificates could attract inspectors if they increase the visibility of the shop.

Table 4.13 collects a set of variables related to signaling quality, including memberships in UNADA and other professional associations, trading licenses, the number of inspections in the last season, and warnings or confiscations of seed after inspection. We find that at endline, judged by the overall index, the clearinghouse treatment led to a significant increase in quality assurance measures. Looking at the individual outcomes, the overall effect seems to be driven by an increase in registrations with UNADA. We also see that shops were inspected significantly more often.

4.7.3 Farmer switching

An important potential mechanism underlying the effect of the information clearinghouse is the possibility that farmers switch from lower rated agro-input dealers to shops that have better ratings. To explore this impact pathway, we asked farmers if they switched agro-input dealers since the previous season and report the results in Table 4.14. We see that only 17% of farmers reported switching at baseline. However, at midline, a significantly higher share of farmers in the clearinghouse treatment group reported switching dealers. Also at endline, we find a higher propensity to switch dealers among clearinghouse treated farmers.

The above reveals increased mobility in the treatment group, but it does not establish that farmers move from lower rated agro-input dealers to higher rated ones. To further investigate this, we calculate the difference between the rating of shop the farmer is switching to and the rating of the shop the farmer is switching from. If farmers move to better rated shops, this difference would be positive. We find that this is indeed the case, and more so during the second

Table 4.12: Effects on agro-input dealer outcomes: Efforts and practices

	<i>baseline</i>		<i>midline</i>		<i>endline</i>	
	mean		training	CH	training	CH
Index of dealer efforts and services, self-reported ^{1†}	0.000 (0.454)		-0.063 (0.062)	0.066 (0.060)	-0.031 (0.051)	0.086 ⁺ (0.048)
Index of dealer efforts and services, according to farmers ^{2†}	-0.027 (0.583)		-0.151* (0.074)	0.301** (0.069)	0.006 (0.092)	0.086 (0.084)
Index of labor-intensive seed handling practices ^{3†}	0.010 (0.484)		0.058 (0.070)	0.099 (0.065)	0.083 (0.067)	0.074 (0.068)
Index of capital-intensive seed handling practices ^{4†}	0.000 (0.508)		-0.019 (0.063)	0.000 (0.072)	-0.087 (0.092)	0.070 (0.081)
Overall index	0.032 (0.540)		-0.029 (0.121)	0.359** (0.113)	0.006 (0.099)	0.165 ⁺ (0.091)
Max. number of obs.				306		297

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

¹The index of dealer efforts and services, self-reported contains 7 variables: whether shop offers explanations, complementary input recommendations, extension/training, discounts for larger quantities, credit, did not receive seed related customer complaint, accepts mobile money.

²The index of dealer efforts and services, according to farmers contains 7 variables: whether shop offers refund/insurance, credit, training/advice, delivery, after-sales service, accepts different payment methods, sells small quantities. The answers are aggregated at dealer level, then the index is computed.

³The index of labor-intensive seed handling practices contains 6 variables: whether seed is stored in dedicated area, in correct lighting, on correct surface, not in open containers, whether shop has no pest problem, cleanliness and professionalism rating by enumerator.

⁴The index of capital-intensive seed handling practices contains 6 variables: whether roof is leak-proof, roof is insulated, walls are insulated, shop is ventilated, shop displays official certificate, expired seed is handled correctly.

Table 4.13: Effects on agro-input dealer outcomes: Memberships, licenses, inspections

	<i>baseline</i>		<i>midline</i>		<i>endline</i>	
	mean	training	CH	obs.	training	CH
Shop is registered with UNADA [†]	0.442	0.040	0.066	252	-0.050	0.118 ⁺
	(0.497)	(0.072)	(0.068)		(0.072)	(0.070)
Shop is member of other professional association [†]	0.345	-0.035	0.058	268	0.001	0.069
	(0.476)	(0.051)	(0.052)		(0.073)	(0.066)
Shop has trading license issued by local government [†]	0.749	-0.042	0.021	288	-0.033	0.008
	(0.435)	(0.053)	(0.054)		(0.056)	(0.057)
Number of shop inspections ^{§†}	1.532	0.037	-0.097	293	0.038	0.292 [*]
	(1.859)	(0.247)	(0.259)		(0.109)	(0.111)
Shop received warning after inspection [†]	0.317	0.045	0.005	291	0.013	-0.009
	(0.466)	(0.072)	(0.073)		(0.062)	(0.063)
Shop's products were confiscated after inspection [†]	0.145	0.021	-0.027	293	0.014	-0.025
	(0.353)	(0.046)	(0.046)		(0.033)	(0.036)
Overall index	-0.004	-0.005	0.047	266	-0.006	0.203 ^{**}
	(0.433)	(0.056)	(0.055)		(0.078)	(0.074)
Max. number of obs.				306		297

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; [†] indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
[§]Due to the skewness of this variable, the regression was run after an Inverse Hyperbolic Sine transformation. Coefficient estimates can therefore be interpreted as percentage changes. The baseline mean column shows the untransformed variable.

Table 4.14: Effects on farmer outcomes: Switching behavior

	<i>midline</i>		<i>midline</i>		<i>endline</i>	
	mean		training	CH	training	CH
Farmer switched to different agro-input shop ¹	0.168		-0.013	0.042**	-0.024	0.026 ⁺
	(0.374)		(0.014)	(0.014)	(0.015)	(0.015)
Max. number of obs.				3407		3441

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
¹We report the mean and standard deviation at midline because this variable was not collected at baseline.

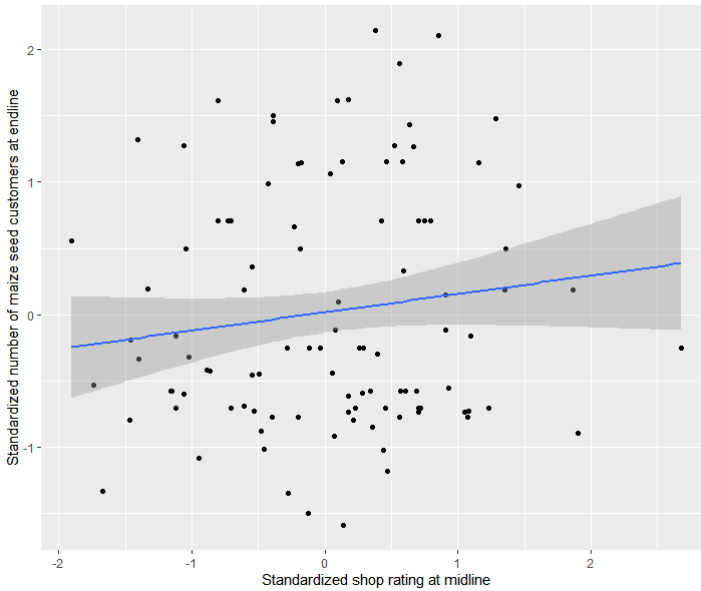


Figure 4.3: Relationship between midline rating and endline number of customers

season, but the difference is not significantly different from zero at conventional levels ($p\text{-value} = 0.166$).

We also explore switching from the agro-input dealer perspective. Here we look at the relationship between the rating a shop received and its number of customers (standardized within the catchment area). If farmers switch from poorly rated dealers to better rated ones, we would expect to see a positive correlation in areas where the clearinghouse treatment was implemented. Figure 4.3 shows that shops with higher ratings at midline receive more customers at endline. However, this evidence is not very strong and at most suggestive.

4.7.4 Farmer perceptions

Finally, the information clearinghouse may change farmer perceptions of the quality of seed sold by agro-input dealers. Table 4.15 provides an analysis of this impact pathway. As a first measure of their perceptions, we asked farmers if they think that maize seed that they can buy at agro-input shops is counterfeit or adulterated. At baseline, two in three farmers responded affirmative to this question. Columns (2) to (5) show the impact of the clearinghouse for the

full sample. The treatment does not significantly affect farmer perceptions as measured by this variable at mid- or endline. However, we expect the effect of the clearinghouse on perceptions to be strongest for farmers who did not adopt improved maize varieties at baseline. Therefore, we repeat the analysis for this subgroup of farmers in columns (6) to (9). At midline, farmers that did not adopt at baseline and live in areas exposed to the clearinghouse are 12.5 percentage points less likely to think that agro-input dealers sell adulterated seed than similar farmers in areas not assigned to the treatment. The effect disappears at endline.

Another important outcome is related to farmer perceptions of product quality: the index of maize seed ratings contains the general quality, yield, drought tolerance, pest and disease tolerance, time of maturity, and germination rating. These ratings are aggregated at the farmer level (as one farmer rates multiple shops), then the index is calculated. To compute this index at smallholder level, a farmer needs to have rated at least one dealer in the catchment area on all components of the index, which leads to a sample size reduction, which in turn may affect statistical power. Nonetheless, we see that the index is positively and significantly affected by the clearinghouse treatment, even though the effect is only significant at the 10% level. If we restrict the sample to farmers that did not adopt improved maize varieties at baseline, the treatment effect on the ratings is significant at the 5% level. Also the impact on the overall combining index is significant for this sub-sample.

It would be straightforward to assume that treated farmers perceive seed to be better because the clearinghouse improved seed quality. Unfortunately, we cannot verify this because objectively measuring seed quality is challenging (see Footnote 48). That is why we cannot confidently conclude that farmers' perceptions of seed improve because treated dealers improve seed quality by handling seed better. Instead, there could be other variables at play, farmers could, for example, rate seed at treated dealers better because these dealers decide to provide advice and recommendations in response to the clearinghouse, which is surely useful but has no impact on the actual quality of maize seed. Even though we cannot conclude that the clearinghouse improved objective maize seed quality, it did so in the eyes of smallholder farmers, which is what counts after all.

Alternatively, the clearinghouse may have mainly improved perceptions, and there are also arguments in favor of this hypothesis. We demonstrate that

Table 4.15: Effects of the clearinghouse on farmer outcomes: Perceptions

	baseline mean	full sample		sub-sample	
		midline CH	endline CH	midline CH	endline CH
		obs.	obs.	obs.	obs.
Farmer thinks maize seed at agro-input shops is adulterated [†]	0.685 (0.465)	-0.041 (0.027)	0.020 (0.028)	-0.125** (0.036)	0.010 (0.035)
Index of farmer's maize seed ratings of shops within area ^{1†}	0.000 (0.637)		0.092+ (0.054)	1664	0.141* (0.063)
Overall index	0.019 (0.770)		0.104 (0.071)	1462	0.160* (0.074)
Max. number of obs.		3407	3441	1719	1741

Note: Column (1) reports baseline means and standard deviations below; columns (2) and (4) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (3) and (5) report number of observations; columns (6) to (9) mirror this structure for the sub-sample of farmers that did not adopt at baseline, **, *, and + denote significance at the 1, 5 and 10% levels; [†] indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

¹The index of farmer's maize seed ratings contains 6 ratings: general quality, yield, drought tolerance, pest/disease tolerance, time of maturity, germination. The ratings are aggregated at farmer level (one farmer rates multiple shops), then the index is computed. Note that treatment and control groups can only be compared at endline. At base- and midline, only clearinghouse treated farmers rated dealers in their proximity because being confronted with these questions is part of the treatment. Hence control dealers were not rated and this line is left blank at midline. At endline, all farmers rated all shops, so that this variable can be investigated.

the intervention affects several measures of adoption already at midline. If we assume that changing dealer behavior and farmers noticing this change takes some time, rectifying incorrect perceptions of smallholders must have played an important role in increasing their adoption.

Also note that the average agro-input shop was rated 3.4 out of 5 at baseline which indicates that seed quality was not so poor to begin with. One could argue that this statement contradicts the one above stating that two in three farmers thought that maize seed at agro-input shops is counterfeit or adulterated at baseline. However, we do not necessarily agree: when farmers are asked about seed quality at agro-input shops *in general*, their perceptions are negative, but when they rate seed *at a particular shop*, they are not. This indicates that smallholders' general opinions about the market are pessimistic but when they focus on their own or someone else's individual experiences, they realize that they are not too dissatisfied after all. Hence the two statements do not necessarily contradict each other, they are answers to differently asked questions. This illustrates that farmers' perceptions of the quality of the inputs in the market are not robust and that misperceptions are quite likely. Furthermore, note that the entire farmer sample provided answers to the counterfeiting/adulteration question, whereas only farmers who bought seed at a particular shop or know someone who did were able to provide ratings, so that the share of adopters is likely to be larger among farmers who rated at baseline than in the full sample. This could indicate that farmers who do not adopt have worse perceptions of seed quality than farmers who do.

We interpret this as support for the hypothesis that the quality of maize seed at some shops in our sample is sufficient but non-adopting farmers misperceive it. This is in line with Michelson et al. (2021) and Wossen, Abay, and Abdoulaye (2022) who establish that input quality is good but that farmers' beliefs are often incorrect, so that one simply needs to rectify this misperception to increase adoption. The information clearinghouse provides an innovative way to do so.

4.8 Attrition

Table 4.16 reports attrition levels in the treatment and comparison groups. We failed to collect data from 12% of dealers and 2% of farmers at midline, and from 14% of dealers and 1% of farmers at endline. To test if non-response is

related to one of the treatments, we regress the likelihood of leaving the sample on the treatment indicators. We find that clearinghouse treated dealers are significantly less likely to leave the sample.

Whether our estimates are biased or not depends on whether this attrition is random or not. It is for instance plausible that the worst performing shops in the clearinghouse control group went out of business. Our clearinghouse treatment might have prevented bankruptcy and helped dealers to stay in the market because it served as some kind of advertisement if the rating was good.

On the other hand, it is plausible that enumerators invested less effort when searching control dealers because they did not have to deliver their SeedAdvisor certificates. Carrying this certificate might have made them more persistent when looking for a shop because they did not want to return to their supervisor without having delivered that paper. Moreover, the certificate might have helped enumerators to find the treated dealers because they were able to show the names to neighbors and so forth (instead of just asking) who in turn helped finding them. In that case, a larger number of random dealers left the control sample, meaning that the dealers who were not found are not different from the ones that were found. The sub-sample of dealers that remained in the control group would then be representative for the entire control group, hence our estimates would be unbiased. Attrition would only reduce power.

We noticed the attrition problem after midline data collection and instructed our enumerators to be more thorough at endline. Consequently, seven of 28 clearinghouse control dealers who were not found for the midline interview, were found for the endline interview later that year. This supports our claim that at least a share control dealer attrition can be explained by a lack of enumerator effort instead of bankruptcy. Furthermore, even if attrition is non-random, the bias is likely to be negative and treatment effects are expected to be positive. As such, the unadjusted selection-contaminated estimates provide lower bounds for the true treatment effect (Angrist, Bettinger, and Kremer, 2006; Duflo, Glennerster, and Kremer, 2007).

4.9 Conclusion

Even though agricultural technologies like high yielding seed varieties and inorganic fertilizers are considered to be key in increasing agricultural productivity and accelerating rural transformation, the adoption by smallholders remains

Table 4.16: Attrition

	mean	training	CH
		<i>midline</i>	
Agro-input dealer left the sample	0.121 (0.326)	-0.007 (0.034)	-0.108** (0.035)
Farmer left the sample	0.018 (0.134)	-0.005 (0.005)	0.001 (0.005)
		<i>endline</i>	
Agro-input dealer left the sample	0.144 (0.351)	0.017 (0.040)	-0.079+ (0.042)
Farmer left the sample	0.008 (0.091)	-0.003 (0.003)	-0.001 (0.003)

Note: Column (1) reports sample means at mid- or endline and standard deviations below; columns (2)-(3) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; **, *, and + denote significance at the 1, 5 and 10% levels.

persistently low in sub-Saharan Africa. We study one particular constraint to technology adoption: the perceived quality of agricultural inputs. We hypothesize that seed quality deteriorates because agro-input dealers lack knowledge and/or because asymmetric information results in excessive search costs for farmers and reduced incentives for dealers.

To assess the importance of these potential constraints to agricultural technology adoption, we tested two interventions in the market for improved maize seed varieties in eastern Uganda using an RCT. A training informed agro-input dealers about correct seed handling and storage practices. An information clearinghouse based on crowd-sourced ratings of the seed that agro-input dealers sell was expected to reduce the information asymmetry between seller and buyer by making the quality of maize seed observable.

The results of our analyses show that training agro-input dealers does not change their practices, and consequently, their operations remain unaffected. We also do not find any impact on farmers that live in the catchment areas of dealers that were trained: they do not perceive seed quality differently nor have higher adoption rates than farmers that were not exposed to trained dealers. These null results show that investing in trainings only may not be an effective strategy as long as dealers are not incentivized. In fact, our results suggest that if agro-input dealers have the right incentives, they actively seek out knowledge necessary to improve and keep ahead of competitors. If dealers are exposed to

both, incentives and information, they handle and store seed better and attract more business.

The information clearinghouse clearly affected the Ugandan market for maize seed as sellers and buyers started behaving in line with our theory of change. Agro-input dealers report more business and smallholders report increased use of improved maize varieties. This effect seems to partly originate from dealers who increase effort and expand service provision to outperform their competition. There is also evidence that the clearinghouse improves the opinions that farmers hold about agro-input shops and their products. In areas where the clearinghouse was implemented, farmers are also more likely to switch between agro-input dealers, but our evidence that they move from lower rated dealers to higher rated ones is at most suggestive. However, as clearinghouse treated farmers report higher yields than control farmers, we can safely conclude that this intervention affected smallholders positively, which is important in terms of welfare implications.

We conclude that quality considerations are important constraints to the adoption of agricultural inputs. This has notable implications for the public sector. Ideally, it needs to safeguard seed supply chains to prevent mishandling and secure the delivery of high quality inputs to farmers. One way to do so would be controls and checks along the supply chain, including inspections of agro-input shops. A complementary strategy would be to inform farmers how to examine bags of seed and identify characteristics that indicate mishandling, such as expired dates, or even tampering or fraud, such as torn seals and seams, and scratched e-verification labels.

We also see important implications for the private sector. If seed companies want to successfully market their products, they need to acknowledge that many of their potential customers have concerns about quality. Their current efforts to make products identifiable and traceable have not been very effective: we bought seed bags from dealers in our sample, and only 8% of them have a certification sticker, and only 3% show an electronic verification label. To assure farmers that they are buying authentic, high-quality seed, companies need stronger signals of product quality. Some of them started labeling their bags with information about germination testing, and in Kenya, some companies started marketing their seed using novel packaging features to signal product authenticity (Gharib et al., 2021). Any strategy to reduce information asymmetries between sellers and buyers by making input quality more observable is

likely to benefit the market for quality inputs. These are clear opportunities for innovations in the private sector that could build consumer confidence in the quality of the improved maize seed available at agro-input dealers.

A crowd-sourced information clearinghouse can be an important institutional innovation to solve the problem of asymmetric information in the market for agricultural inputs. It may be preferable to alternative strategies due to its likely lower cost and self-sustaining nature, and helps to overcome problems such as insufficient public investment in regulatory systems, regulatory enforcement, and market surveillance. Furthermore, peers, who are familiar with the heterogeneous conditions farmers face, provide the clearinghouse ratings, and their opinion may be more useful and trustworthy for smallholders than the judgment of an inspector or a seed certification or verification system.

The objective of this paper is to prove a concept and to test whether a prediction of the theory occurs in practice, namely whether making quality observable improves the market for maize seed. If the clearinghouse would be scaled up, several challenges would need to be addressed. For example, one should decide how long SeedAdvisor certificates remain valid and whether they have to be removed at some point, as old certificates could provide incorrect signals, and dealers who received a good certificate at some point in time would not have any incentive to keep up the good work if the outdated certificate remains displayed in their store. Furthermore, farmers may favor dealers they personally know well or discriminate against minorities, and these preferences, instead of quality differences, may be reflected in the ratings. In this case, the ratings would need to be monitored and potentially adjusted, to avoid that the clearinghouse amplifies bias or discrimination. Another danger to the functioning of the clearinghouse could be agro-input dealers who start to influence ratings in a dishonest manner, for example, by faking ratings or by bribing farmers. Note that during this trial, rating and rated participants were connected by asking dealers where their customers come from, and collecting ratings from ten randomly selected farmers in that village. Dealers would need to understand our experimental design in order to know which farmer they would need to influence, so that it is almost impossible that they cheat during this trial, but details like this need to be considered before scaling up the experiment.

4.10 Acknowledgments

We would like to thank Charles Marc Wanume, Leocardia Nabwire, Richard Ariong, and Wilberforce Walukano for their support in the field, and Martin Paul Jr. Tabe-Ojong for his comments and suggestions as our discussant at the Centre for the Study of African Economies Conference 2023. This project received funding from The Netherlands - Consultative Group for International Agricultural Research research programme on Seed Systems Development, grant number W08.240.105, funded by the Netherlands Organisation for Scientific Research. The research was also supported by the Fonds Wetenschappelijk Onderzoek - Vlaanderen and the Fonds de la Recherche Scientifique under EOS project G0G4318N. This project received clearance from Makerere University's School of Social Sciences Research Ethics Committee (08.20.436/PR1) as well as from the International Food Policy Research Institute's Institutional Review Board (DSGD-20-0829). The research was also registered at the Ugandan National Commission for Science and Technology (SS603ES).

4.11 Appendix

4.11.1 Can farmers assess maize seed quality?

The objective of the clearinghouse is to make maize seed quality observable, so that its functioning would be endangered if farmers' ratings do not actually measure quality but, for example, the personal relationship between a farmer and a dealer. Some could even argue that farmers cannot assess the quality of maize seed at all, even not after using it, because there are so many factors at play in agricultural production: if farmers experience a disappointing harvest, they cannot safely conclude that the seed material was poor because it could have also been poor soil, insufficient, late or too much rain, or own mismanagement like late planting or insufficient weeding. Misattribution occurs when farmers mistakenly ascribe bad outcomes to bad inputs, rather than to other possible causes. Learning about the true quality becomes more difficult when this misattribution is present (Hoel et al., 2021). Tjernström et al. (2021) agree that sub-Saharan Africa's soil heterogeneity hampers farmer learning about the returns to inputs. Maize seed would then be a credence good instead of an experience good and clearinghouse ratings of farmers would be useless. Even though farmers and dealers might still change their behavior in the short run because they expect the clearinghouse to work, this effect would fade out as soon as both actors learn that the ratings are unreliable.

We argue that farmers' ability to infer maize seed quality is indeed not perfect but they can assess it to some extent. Shortly after planting, farmers can observe germination rates, i.e., the proportion of seeds that germinate, and later how fast the seed matures. Some seed may also be more susceptible to pests and diseases, while other seed may be particularly tolerant in terms of drought. After harvest, the farmer can observe the yield. All these attributes of seed quality can be judged to some extent after one agricultural season.

To support our claim, we test if the ratings are correlated with objective indications of seed quality. We find that specialized agro-input shops which only sell farm inputs have higher ratings, so do dealers with better seed handling practices. Furthermore, farmers did not only rate the seed of agro-input shops but also the seed they used on a randomly selected maize field, using the same questions. We find that these seed ratings are positively correlated with farmers' yield. All this shows that the clearinghouse ratings do measure maize seed quality to some extent.

Table 4.17: Correlating ratings & quality indicators - dealer level

<i>independent variables:</i>		<i>dependent variable:</i> rating index _{endline}
Shop only sells farm inputs _{endline}		0.205**
Index of labor-intensive seed handling practices _{endline}		0.164 ⁺
Index of capital-intensive seed handling practices _{endline}		0.204 [*]
Index of all seed handling practices _{endline}		0.225 ⁺
Shop received seed related complaint from customer _{endline}		-0.108
Shop received a warning after inspection _{endline}		-0.034
Randomly selected seed bag shows packaging date _{endline}		0.052
Days since packaging date/expiry date minus 6 months _{endline}		0.000
Seed is in the original bag without any signs of damage _{endline}		0.101
Randomly selected seed bag shows lot number _{endline}		0.111
Moisture in randomly selected seed bag in % _{endline}		0.027

Table 4.18: Correlating ratings & quality indicators - farmer level

<i>independent variables:</i>		<i>dependent variable:</i>
		Yield in kg/acre on randomly selected field _{endline}
Index of farmer's ratings of seed used on randomly selected field _{midline}		144.971**
Index of farmer's ratings of seed used on randomly selected field _{endline}		35.225**

4.11.2 Rating computation: details

What to do if a treated dealer does not receive a single rating? If a shop in a treated catchment area is not rated by a single farmer, for example, because no farmer in our sample knows him or her, we could fill in the catchment area mean as his or her rating. However, this is not as innocent as it seems because it is likely that the lack of ratings is not random. Poor quality dealers have less customers, so their likelihood to get rated is lower. Giving them average catchment area ratings inflates the ratings of these low quality dealers. Instead, we simply told farmers that we do not have information about this shop (implicitly informing the farmer that it exists). 16 of 193 treated dealers were not rated by a single farmer in the first round.

Should more ratings lead to better ratings? Some shops were not rated by any smallholder in the first round, while others were rated by up to 22 smallholders. If dealer A is rated by ten farmers and gets rating 3,5 and dealer B is rated by one farmer and gets rating 3,6, we treat dealer B as the better dealer. Even though receiving many (few) ratings can be related to good (poor) quality (the lack of ratings could be nonrandom, see previous paragraph), there could be other reasons why dealers are rated by many (few) farmers. Furthermore, giving higher ratings to better-known dealers could harm new dealers entering the market and dealers who are discriminated, for example, due to their gender. Also on TripAdvisor, having more reviews than a rival hotel does not lead to a better rating.

Should ratings depend on catchment area dealer performance? The following examples show that ratings should not depend on catchment area averages. In an area with poor quality dealers in which one dealer is a bit better than the rest but still poor, we do not want this dealer to be rated well (i.e., expose farmers to poor quality dealers). Similarly, in an area with good dealers in which one dealer is a bit worse than the rest but still good, we do not want this dealer to be rated poorly (which would be unfair towards him or her). On the other hand, less than 9% of shops received a rating below 3 out of 5, so we would throw away valuable data if we would only disseminate good scores without any variation. Therefore, we take the distribution of ratings into account by using quintiles. Consequently, less dealers receive rating 4 or 5, more dealers receive rating 1 or 2. This could strengthen the effect of the

treatment on dealer effort. If dealers get ratings 1 or 2 instead of 4 or 5, they could feel more inclined to improve their scores. Consequently, also the effect on seed quality itself could be larger. However, the clearinghouse should also have a signaling effect, which might be weaker if more dealers are rated 1 or 2 instead of 4 or 5 (dealers would seem to be of worse quality to farmers). Therefore, we chose words with a positive connotation as the quintile names for rating dissemination. As most dealers received a good or very good rating before taking the distribution into account, we ensure that even a 2 is still communicated as “good” to farmers to not weaken the signaling effect. That is why the first quintile is translated to “okay” and gets one star, the second one is named “good” and receives two stars, the third quintile is “very good” and gets three stars, the fourth and fifth one are “excellent” and awarded with four and five stars. This way of considering the distribution of the original ratings when choosing the names also helps us to disseminate ratings as truthfully, purely and as closely to reality as possible.

Are female dealers rated worse than male dealers? Because we found significant differences between the ratings of female (41% of dealers) and male agro-input dealers (59% of dealers) after controlling for some potentially confounding variables like education and for several indications of quality, we have no reason to believe that these differences in perception can be explained by differences in real quality. Instead, it is likely that women are perceived to be worse due to discrimination (De, Mieke, and Van Campenhout, 2022), so that we adjusted the ratings of female dealers accordingly to prevent that they are harmed by our intervention. We regressed all seed quality attributes on the gender dummy and added the resulting coefficients to the initial ratings of female dealers.

4.11.3 Rating dissemination: details

Table 4.19: Text messages to disseminate ratings to farmers

treatment SMS	Hello from AgroAdvisor!
	Did you know that customers from <i>shop name</i> rate the quality of maize seed sold there as okay/good/very good/excellent?
control SMS	Hello from AgroAdvisor!
	Did you know that you can get quality maize seed in your area from <i>shop name</i> ?

4.11.4 Outcome variables and results as they were pre-registered

Table 4.20: Effects on primary dealer outcomes

	baseline mean	training	midline CH	obs.	training	endline CH	obs.
Quantity of maize seed sold in kg ^{§†}	695.503 (1497.183)	-0.092 (0.220)	0.284 (0.227)	292	-0.499 ⁺ (0.250)	0.239 (0.253)	286
Sales price of maize seed in UGX/kg	4273.897 (955.073)	-192.784 ⁺ (114.934)	99.272 (113.292)	275	-33.867 (143.152)	145.861 (138.816)	264
Revenue from maize seed in mln UGX ^{§†}	2.890 (6.286)	-0.069 (0.104)	0.185 ⁺ (0.108)	292	-0.227 ⁺ (0.118)	0.143 (0.118)	286
Number of maize seed customers per day ^{§†}	19.764 (20.689)	-0.056 (0.098)	0.127 (0.101)	294	-0.190 (0.116)	0.310 ^{**} (0.112)	288
Moisture in randomly selected seed bag in %	13.563 (1.442)	0.017 (0.142)	-0.122 (0.144)	175	-0.041 (0.198)	-0.220 (0.197)	261
Index of capital-intensive seed handling practices ^{1†}	0.000 (0.508)	-0.019 (0.063)	0.000 (0.072)	270	-0.087 (0.092)	0.070 (0.081)	265
Index of labor-intensive seed handling practices ^{2†}	0.010 (0.484)	0.058 (0.070)	0.099 (0.065)	285	0.083 (0.067)	0.074 (0.068)	274
Index of all seed handling practices ³	0.009 (0.382)	0.042 (0.051)	0.052 (0.053)	251	0.021 (0.063)	0.083 (0.059)	248
Index of dealer's efforts and services ^{4†}	0.000 (0.454)	-0.063 (0.062)	0.066 (0.060)	243	-0.031 (0.051)	0.086 ⁺ (0.048)	297
Index of shop's maize seed ratings by farmers ⁵	-0.018 (0.595)				0.020 (0.102)	0.122 (0.101)	327
Overall index	0.007 (0.591)	-0.004 (0.130)	0.214 ⁺ (0.121)	215	-0.058 (0.128)	0.239 [*] (0.117)	258
Max. number of obs. for dealer survey outcomes				306			297

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

[§]Due to the skewness of this variable, the regression was run after an Inverse Hyperbolic Sine transformation. Coefficient estimates can therefore be interpreted as percentage changes. The baseline mean column shows the untransformed variable.

¹The index of capital-intensive seed handling and storage practices contains 6 variables: whether roof is leak-proof, whether roof is insulated, whether walls are insulated, whether shop is ventilated, whether any official certificate is displayed, whether expired seed is handled correctly.

²The index of labor-intensive seed handling and storage practices contains 6 variables: whether seed is stored in dedicated area, whether shop has no pest problem, whether seed is stored in correct lighting, whether seed is stored on correct surface, whether seed is not stored in open containers, cleanliness and professionalism rating by enumerator.

³The index of all seed handling and storage practices contains 12 variables: the ones included in the index of capital-intensive practices and the ones included in the index of labor-intensive practices.

⁴The index of dealer's efforts and services contains 7 variables: whether shop offers explanations, complementary input recommendations, extension/training, discounts for larger quantities, credit, did not receive seed related customer complaint, accepts mobile money.

⁵The index of shop's maize seed ratings by farmers contains 6 ratings: general quality, yield, drought tolerance, pest/disease tolerance, time of maturity, germination. Ratings are aggregated at shop level (one shop is rated by multiple farmers), then the index is computed.

Table 4.21: Effects on secondary dealer outcomes: Indices

	<i>baseline</i>		<i>midline</i>		<i>endline</i>	
	mean	training	CH	obs.	training	CH
Index of dealer's motivation and satisfaction ¹	0.000 (0.674)	0.033 (0.082)	0.000 (0.085)	306	-0.109 (0.082)	-0.076 (0.086)
Index of dealer's self-ratings ²	0.000 (0.651)	-0.068 (0.084)	-0.002 (0.079)	306	-0.132 (0.086)	0.080 (0.079)
Index of dealer's efforts and services according to farmers ³	-0.027 (0.583)	-0.151* (0.074)	0.301** (0.069)	259	0.006 (0.092)	0.086 (0.084)
Index of dealer's knowledge about seed storage ⁴	0.000 (0.482)	0.091 (0.076)	0.115 (0.075)	306	0.030 (0.053)	0.124* (0.055)
Index of dealer's knowledge about seed ⁵	0.000 (0.533)	0.102 (0.072)	0.065 (0.070)	306	-0.009 (0.080)	-0.007 (0.078)
Max. number of obs.				306		297

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
¹The index of dealer's motivation and satisfaction contains 3 variables: whether dealers see themselves working as agro-input dealers in future, would recommend working as dealers, how happy dealers feel when they come to work. We report the mean and standard deviation at midline because these variables were not collected at baseline.
²The index of dealer's self-ratings contains 5 ratings: location, price, product quality, stock, reputation.
³The index of dealer's efforts and services according to farmers contains 7 variables: whether shop offers refund/insurance, credit, training/advice, delivery, after-sales service, accepts different payment methods, sells small quantities. The answers are aggregated at dealer level, then the index is computed.
⁴The index of dealer's knowledge about seed storage contains 5 variables: whether dealer knows how long seed can be carried over, how seed should be stored after repackaging, what the min. distance between floor and seed is, how seed should be stored in storeroom, whether seed should be repackaged.
⁵The index of dealer's knowledge about seed contains 4 variables: whether dealer knows which seed variety to recommend if farmer complains about poor soil, if farmer complains about little rain, if farmer is late for planting, what to tell clients about yield benefits of hybrid seed.

Table 4.22: Effects on primary farmer outcomes

	<i>baseline</i>		<i>midline</i>		<i>endline</i>	
	mean	training	CH	obs.	training	CH
Farmer planted improved maize seed on any plot [†]	0.492 (0.500)	-0.021 (0.020)	0.035 ⁺ (0.020)	3206	-0.009 (0.020)	0.042* (0.020)
Farmer bought maize seed at agro-input shop for any plot [†]	0.325 (0.468)	-0.014 (0.021)	0.059** (0.021)	3145	0.004 (0.019)	0.031 (0.020)
Amount of this seed farmer bought at agro-input shop in kg	9.519 (6.920)	0.512 (0.348)	-0.105 (0.358)	599	0.457 (0.419)	0.378 (0.431)
Index of farmer's maize seed ratings of shops within catchment area ¹	0.000 (0.637)				0.021 (0.054)	0.092 ⁺ (0.054)
Index of farmer's general ratings of shops within catchment area ²	0.000 (0.657)				-0.026 (0.043)	-0.005 (0.042)
Index of services of shops within catchment area according to farmers ³	-0.037 (0.609)	-0.138 ⁺ (0.073)	0.161* (0.067)	312	0.034 (0.081)	0.131 ⁺ (0.077)
Farmer switched to different agro-input shop ^{4†}	0.168 (0.374)	-0.013 (0.014)	0.042** (0.014)	3407	-0.024 (0.015)	0.026 ⁺ (0.015)
Index of farmer's practices on randomly selected plot ^{5†}	0.008 (0.400)	0.011 (0.019)	-0.026 (0.019)	2929	0.001 (0.021)	0.016 (0.021)
Farmer thinks maize seed at agro-input shops is adulterated	0.685 (0.465)	-0.033 (0.027)	-0.041 (0.027)	2113	-0.041 (0.028)	0.020 (0.028)
Farmer planted land race maize seed on randomly selected plot [†]	0.448 (0.497)	0.015 (0.021)	-0.013 (0.020)	2954	0.009 (0.022)	-0.024 (0.022)
Overall index⁶	0.009 (0.698)	0.008 (0.033)	0.017 (0.034)	2933	-0.023 (0.034)	0.063 ⁺ (0.034)
Max. number of obs.				3407		3441

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; [†] indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

¹The index of farmer's maize seed ratings contains 6 ratings: general quality, yield, drought tolerance, pest/disease tolerance, time of maturity, germination. The ratings are aggregated at farmer level (one farmer rates multiple shops), then this index is computed.

²The index of farmer's general ratings contains 6 ratings: general quality, location, price, product quality, stock, reputation. The ratings are aggregated at farmer level (one farmer rates multiple shops), then this index is computed.

³The index of services of shops within catchment area contains 7 variables: whether shop offers refund/insurance, credit, training/advice, delivery, after-sales service, accepts different payment methods, sells small quantities. The answers are aggregated at shop level, then the index is computed at farmer level. Only 320 farmers answered all 7 questions for at least one shop within the catchment area at baseline and at endline.

⁴We report the mean and standard deviation at midline because this variable was not collected at baseline.

⁵The index of farmer's practices contains 10 variables: whether farmer spaced seed correctly, sowed correct number of seeds/hill, applied organic manure, DAP/NPK, Urea, pesticides/herbicides/fungicides, weeded sufficiently, weeded at correct time, planted at correct time, re-sowed.

⁶We report the mean and standard deviation at midline because not all variables in this index were collected at baseline.

Table 4.23: Effects on secondary farmer outcomes: Adoption on randomly selected maize plot

	baseline		midline		endline		
	mean	training	CH	obs.	training	CH	obs.
Farmer planted hybrid seed†	0.264	0.002	0.009	2654	-0.023	0.032	2700
	(0.441)	(0.022)	(0.022)		(0.023)	(0.023)	
Farmer planted open-pollinated seed†	0.260	-0.017	0.002	2654	0.010	-0.007	2700
	(0.439)	(0.022)	(0.022)		(0.020)	(0.021)	
Farmer planted farmer-saved seed†	0.579	0.020	-0.042+	3153	-0.009	-0.016	3240
	(0.494)	(0.022)	(0.022)		(0.020)	(0.020)	
Farmer planted seed bought at agro-input shop†	0.330	-0.010	0.047*	3153	0.012	0.036+	3240
	(0.470)	(0.022)	(0.022)		(0.019)	(0.019)	
Farmer planted hybrid or open-pollinated seed¹	0.432	-0.019	0.035	2954	0.009	0.030	3047
	(0.495)	(0.023)	(0.023)		(0.023)	(0.023)	
Overall index	-0.003	0.000	0.002	2867	-0.010	0.026	2963
	(0.553)	(0.024)	(0.024)		(0.025)	(0.025)	
Max. number of obs.				3407	3441		

Note: Column (1) reports baseline means and standard deviations below; columns (2), (3), (5), and (6) report differences between treatment and control groups and standard errors below; they are clustered at the level of randomization; columns (4) and (7) report number of observations; **, *, and + denote significance at the 1, 5 and 10% levels; [†] indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
¹For this variable, only seed which was not farmer-saved counted as hybrid seed and only seed which was not recycled too often counted as open-pollinated seed.

4.11.5 Multiple choice questions to measure dealer knowledge

Dealer knowledge about seed storage

1. How long can seed be carried over before losing viability?
 - (a) Seed can be carried over into the next seasons as you can store seed for twelve months.
 - (b) Seed cannot be carried over into the next seasons as six months is the longest seed can be stored.
 - (c) This depends on the seed: hybrids cannot be carried over, OPVs can be carried over for five seasons.
 - (d) I don't know.
2. How should seed best be stored after repackaging?
 - (a) Airtight in polyethylene bags.
 - (b) In paper bags or perforated polyethylene bags.
 - (c) In a sealed tin/plastic container.
 - (d) I don't know.
3. What is the minimum recommended distance between the floor and where seed is stored?
 - (a) 0 inches, seed should be stored directly on the floor for maximum stability.
 - (b) Minimum two inches from the floor.
 - (c) Minimum six inches from the floor.
 - (d) I don't know.
4. How should seed ideally be stored in your store room?
 - (a) In sealed cardboard boxes.
 - (b) Stacked on pallets.
 - (c) Arranged on shelves with sufficient space between packets.
 - (d) I don't know.

5. Which statement do you agree most with?
 - (a) You should repackage all your seed to visually verify that you are selling good quality seed.
 - (b) You should repackages all your seed so you can sell more to small farmers.
 - (c) You should avoid repackaging your seed as much as possible.
 - (d) I don't know.

Dealer knowledge about seed

1. If a farmer complains about poor soil, which maize variety do you recommend?
 - (a) Longe 5.
 - (b) Bazooka.
 - (c) Longe 10H.
 - (d) I don't know.
2. What do you tell clients who inquire about the yield benefits of hybrid seeds?
 - (a) Hybrid seeds double maize yields (increasing yield from about 4 to 8 bags/acre).
 - (b) Hybrid seeds triple maize yields (increasing yield from about 4 to 12 bags/acre).
 - (c) Hybrid seeds increase yields tenfold (increasing yield from about 4 to 40 bags/acre).
 - (d) I don't know.
3. If a farmer misses the rains or lives in an area that receives little rain, which maize variety do you recommend?
 - (a) Longe 10H.
 - (b) Longe 7H.
 - (c) Wema.
 - (d) I don't know.

4. If a farmer is late for planting in the short season and needs a fast maturing variety, which maize variety do you recommend?
- (a) Bazooka.
 - (b) Longe 10H.
 - (c) Myezi mitatu (mm3).
 - (d) I don't know.

Bibliography

- Abate, G. T., S. Rashid, C. Borzaga, and K. Getnet. 2016. "Rural finance and agricultural technology adoption in Ethiopia: Does the institutional design of lending organizations matter?" *World Development* 84: 235–253.
- Abate, G. T., T. Bernard, S. Makhija, and D. J. Spielman. 2023. "Accelerating technical change through ICT: Evidence from a video-mediated extension experiment in Ethiopia." *World Development* 161: 106089.
- Ainembabazi, J. H. and J. Mugisha. 2014. "The role of farming experience on the adoption of agricultural technologies: Evidence from smallholder farmers in Uganda." *Journal of Development Studies* 50 (5): 666–679.
- Akerlof, G. A. 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics* 84 (3): 488–500.
- Alibhai, S., A. A. Donald, M. Goldstein, A. A. Oguz, A. Pankov, and F. Strobbe. 2019. "Gender bias in SME lending: Experimental evidence from Turkey." *World Bank Policy Research Working Paper* (9100).
- Anderson, J. R. and G. Feder. 2004. "Agricultural Extension: Good Intentions and Hard Realities." *The World Bank Research Observer* 19 (1): 41–60.
- Anderson, M. L. 2008. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–1495.
- Angrist, J., E. Bettinger, and M. Kremer. 2006. "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia." *American Economic Review* 96 (3): 847–862.

- Arkedis, J., J. Creighton, A. Dixit, A. Fung, S. Kosack, D. Levy, and C. Tolmie. 2021. "Can transparency and accountability programs improve health? Experimental evidence from Indonesia and Tanzania." *World Development* 142: 1–14.
- Ashour, M., D. O. Gilligan, J. B. Hoel, and N. I. Karachiwalla. 2019. "Do beliefs about herbicide quality correspond with actual quality in local markets? Evidence from Uganda." *Journal of Development Studies* 55 (6): 1285–1306.
- Ashraf, N., X. Giné, and D. 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.
- Banerjee, A., A. Deaton, and E. Duflo. 2004. "Wealth, health, and health services in rural Rajasthan." *American Economic Review* 94 (2): 326–330.
- Banerjee, A., R. Hanna, J. Kyle, B. A. Olken, and S. Sumarto. 2018. "Tangible Information and Citizen Empowerment: Identification Cards and Food Subsidy Programs in Indonesia." *Journal of Political Economy* 126 (2): 451–491.
- Barrett, C. B. and M. R. Carter. 2010. "The Power and Pitfalls of Experiments in Development Economics: Some Non-random Reflections." *Applied Economic Perspectives and Policy* 32 (4): 515–548.
- Barriga, A. and N. Fiala. 2020. "The supply chain for seed in Uganda: Where does it go wrong?" *World Development* 130: 104928.
- Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova. 2009. "Powerful women: Does exposure reduce bias?" *The Quarterly Journal of Economics* 124 (4): 1497–1540.
- Beegle, K., N. Karachiwalla, T. J. Lybbert, H. Michelson, J. Sanabria, J. Stevenson, and E. Tjernstrom. 2021. "Devil in the details: measuring seeds." World Bank Blogs.
- Bell, R. M. and D. F. McCaffrey. 2002. "Bias reduction in standard errors for linear regression with multi-stage samples." *Survey Methodology* 28 (2): 169–181.

- Bellemare, M. F. and C. J. Wichman. 2020. “Elasticities and the Inverse Hyperbolic Sine Transformation.” *Oxford Bulletin of Economics and Statistics* 82 (1): 50–61.
- Bernheim, B. D. and M. Whinston. 1986. “Common Agency.” *Econometrica* 54 (4): 923–942.
- Bertsimas, D., M. Johnson, and N. Kallus. 2015. “The power of optimization over randomization in designing experiments involving small samples.” *Operations Research* 63 (4): 868–876.
- Björkman, M. and J. Svensson. 2009. “Power to the people: Evidence from a randomized field experiment on community-based monitoring in Uganda.” *The Quarterly Journal of Economics* 124 (2): 735–769.
- Björkman Nyqvist, M., D. de Walque, and J. Svensson. 2017. “Experimental evidence on the long-run impact of community-based monitoring.” *American Economic Journal: Applied Economics* 9 (1): 33–69.
- Bold, T., K. C. Kaizzi, J. Svensson, and D. Yanagizawa-Drott. 2017. “Lemon technologies and adoption: Measurement, theory and evidence from agricultural markets in Uganda.” *The Quarterly Journal of Economics* 132 (3): 1055–1100.
- Bulte, E., G. Beekman, S. Di Falco, J. Hella, and P. 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.
- Bulte, E., S. Di Falco, M. Kassie, and X. Vollenweider. 2023. “Low-quality seeds, labor supply and economic returns: Experimental evidence from Tanzania.” *Review of Economics and Statistics* 1–33.
- Byerlee, D. and E. H. De Polanco. 1986. “Farmers’ stepwise adoption of technological packages: Evidence from the Mexican Altiplano.” *American Journal of Agricultural Economics* 68 (3): 519–527.
- Cameron, A. C. and D. L. Miller. 2015. “A practitioner’s guide to cluster-robust inference.” *Journal of Human Resources* 50 (2): 317–372.

- Campos, F., M. Frese, M. Goldstein, L. Iacovone, H. C. Johnson, D. McKenzie, and M. Mensmann. 2017. “Teaching personal initiative beats traditional training in boosting small business in West Africa.” *Science* 357 (6357): 1287–1290.
- Card, D., S. DellaVigna, P. Funk, and N. Iriberry. 2019. “Are Referees and Editors in Economics Gender Neutral?” *The Quarterly Journal of Economics* 135 (1): 269–327.
- Chari, A. and P. Goldsmith-Pinkham. 2017. *Gender representation in economics across topics and time: Evidence from the NBER summer institute*. Tech. rep., National Bureau of Economic Research.
- Chen, M., C. Hu, and R. J. Myers. 2022. “Understanding transient technology use among smallholder farmers in Africa: A dynamic programming approach.” *Agricultural Economics* .
- Cochrane, W. W. 1958. *Farm prices: Myth and reality*. U of Minnesota Press.
- Conley, T. G. and C. R. Udry. 2010. “Learning about a new technology: Pineapple in Ghana.” *American Economic Review* 100 (1): 35–69.
- Correll, S. J., K. R. Weisshaar, A. T. Wynn, and J. D. Wehner. 2020. “Inside the Black Box of Organizational Life: The Gendered Language of Performance Assessment.” *American Sociological Review* 85 (6): 1022–1050.
- Curzi, D., P. Nota, and S. Di Falco. 2022. “Post-Harvest Losses and Climate Conditions in Sub-Saharan Africa.” *Contributed Paper prepared for presentation at the 96th Annual Conference of the Agricultural Economics Society* .
- Custodio, M. C., M. Demont, A. Laborte, and J. Ynion. 2016. “Improving food security in Asia through consumer-focused rice breeding.” *Global Food Security* 9: 19–28.
- De, A., C. Mieke, and B. Van Campenhout. 2022. “Gender bias in consumer perceptions: The case of agro-input dealers in Uganda.” *IFPRI Discussion Paper* .
- de Souza Leão, L. and G. Eyal. 2019. “The rise of randomized controlled trials (RCTs) in international development in historical perspective.” *Theory and Society* 48 (3): 383–418.

- Dillon, B., J. C. Aker, and J. E. Blumenstock. 2020. "How Important is the Yellow Pages? Experimental Evidence from Tanzania." *CEPR Discussion Paper* .
- Dolan, C. 2001. "The 'Good Wife': Struggles over Resources in the Kenyan Horticultural Sector." *The Journal of Development Studies* 37 (3): 39–70.
- Donato, K. and A. G. Mosqueira. 2019. "Information Improves Provider Behaviour: A Replication Study of a Community-Based Monitoring Programme in Uganda." *The Journal of Development Studies* 55 (5): 967–988.
- Drexler, A., G. Fischer, and A. Schoar. 2014. "Keeping It Simple: Financial Literacy and Rules of Thumb." *American Economic Journal: Applied Economics* 6 (2): 1–31.
- Duflo, E. and A. Banerjee. 2011. *Poor economics*. PublicAffairs.
- Duflo, E., R. Glennerster, and M. Kremer. 2007. "Using randomization in development economics research: A toolkit." *Handbook of Development Economics* 4: 3895–3962.
- Duflo, E., M. Kremer, and J. Robinson. 2011. "Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya." *American Economic Review* 101 (6): 2350–90.
- Emerick, K., A. de Janvry, E. Sadoulet, and M. H. Dar. 2016. "Technological innovations, downside risk, and the modernization of agriculture." *American Economic Review* 106 (6): 1537–61.
- Evenson, R. E. and D. Gollin. 2003. "Assessing the impact of the Green Revolution, 1960 to 2000." *Science* 300 (5620): 758–762.
- Fafchamps, M. and B. Minten. 2012. "Impact of SMS-based agricultural information on Indian farmers." *The World Bank Economic Review* 26 (3): 383–414.
- Feder, G., R. E. Just, and D. Zilberman. 1985. "Adoption of agricultural innovations in developing countries: A survey." *Economic Development and Cultural Change* 33 (2): 255–298.

- Ferrara, E. L., A. Chong, and S. Duryea. 2012. "Soap operas and fertility: Evidence from Brazil." *American Economic Journal: Applied Economics* 4 (4): 1–31.
- Foster, A. D. and M. R. Rosenzweig. 1995. "Learning by doing and learning from others: Human capital and technical change in agriculture." *Journal of Political Economy* 103 (6): 1176–1209.
- Francis, P. and R. James. 2003. "Balancing rural poverty reduction and citizen participation: The contradictions of Uganda's decentralization program." *World Development* 31 (2): 325–337.
- Fujiwara, T. and L. Wantchekon. 2013. "Can informed public deliberation overcome clientelism? Experimental evidence from Benin." *American Economic Journal: Applied Economics* 5 (4): 241–255.
- Garnett, T., M. C. Appleby, A. Balmford, I. J. Bateman, T. G. Benton, P. Bloomer, B. Burlingame, M. Dawkins, L. Dolan, D. Fraser et al. 2013. "Sustainable intensification in agriculture: Premises and policies." *Science* 341 (6141): 33–34.
- Gars, J. and P. S. Ward. 2019. "Can differences in individual learning explain patterns of technology adoption? Evidence on heterogeneous learning patterns and hybrid rice adoption in Bihar, India." *World Development* 115: 178–189.
- Gharib, M. H., L. H. Palm-Forster, T. J. Lybbert, and K. D. Messer. 2021. "Fear of fraud and willingness to pay for hybrid maize seed in Kenya." *Food Policy* 102: 102040.
- Giné, X. and G. Mansuri. 2021. "Money or Management? A Field Experiment on Constraints to Entrepreneurship in Rural Pakistan." *Economic Development and Cultural Change* 70 (1): 41–86.
- Giroux, S., J. Blekking, K. Waldman, D. Resnick, and D. Fobi. 2021. "Informal vendors and food systems planning in an emerging African city." *Food Policy* 103: 101997.
- Gorton, M., J. Sauer, and P. Supatpongkul. 2011. "Wet Markets, Supermarkets and the "Big Middle" for Food Retailing in Developing Countries: Evidence from Thailand." *World Development* 39 (9): 1624–1637.

- Govender, V., T. Aveling, and Q. Kritzing. 2008. "The effect of traditional storage methods on germination and vigour of maize (*Zea mays* L.) from northern KwaZulu-Natal and southern Mozambique." *South African Journal of Botany* 74 (2): 190–196.
- Goyal, A. 2010. "Information, Direct Access to Farmers, and Rural Market Performance in Central India." *American Economic Journal: Applied Economics* 2 (3): 22–45.
- Griliches, Z. 1957. "Hybrid corn: An exploration in the economics of technological change." *Econometrica, Journal of the Econometric Society* 501–522.
- Grossman, G. and K. Michelitch. 2018. "Information dissemination, competitive pressure, and politician performance between elections: A field experiment in Uganda." *American Political Science Review* 112 (2): 280–301.
- Hanna, R., S. Mullainathan, and J. Schwartzstein. 2014. "Learning through noticing: Theory and evidence from a field experiment." *The Quarterly Journal of Economics* 129 (3): 1311–1353.
- Hasanain, S. A., M. Y. Khan, and A. Rezaee. 2023. "No bulls: Experimental evidence on the impact of veterinarian ratings in Pakistan." *Journal of Development Economics* 161: 102999.
- Haugerud, A. 1995. *The Culture of Politics in Modern Kenya*. Cambridge University Press.
- Hickey, S. and S. King. 2016. "Understanding Social Accountability: Politics, Power and Building New Social Contracts." *The Journal of Development Studies* 52 (8): 1225–1240.
- Hoffmann, V., S. K. Mutiga, J. W. Harvey, R. J. Nelson, and M. G. Milgroom. 2021. "Observability of food safety losses in maize: Evidence from Kenya." *Food Policy* 98: 101895.
- Hornik, J., R. Shaanan Satchi, L. Cesareo, and A. Pastore. 2015. "Information dissemination via electronic word-of-mouth: Good news travels fast, bad news travels faster!" *Computers in Human Behavior* 45: 273–280.
- Hospido, L. and C. Sanz. 2021. "Gender Gaps in the Evaluation of Research: Evidence from Submissions to Economics Conferences." *Oxford Bulletin of Economics and Statistics* 83 (3): 590–618.

- Humphreys, M., R. S. De la Sierra, and P. Van der Windt. 2013. "Fishing, commitment, and communication: A proposal for comprehensive nonbinding research registration." *Political Analysis* 1–20.
- Hünermund, P. and B. Louw. 2020. "On the nuisance of control variables in regression analysis." *arXiv preprint, arXiv:2005.10314* .
- Jayachandran, S. 2021. "Social norms as a barrier to women's employment in developing countries." *IMF Economic Review* 69 (3): 576–595.
- Jensen, R. 2010. "The (perceived) returns to education and the demand for schooling." *The Quarterly Journal of Economics* 125 (2): 515–548.
- Kabunga, N. S., T. Dubois, and M. Qaim. 2012. "Yield effects of tissue culture bananas in Kenya: Accounting for selection bias and the role of complementary inputs." *Journal of Agricultural Economics* 63 (2): 444–464.
- Kahneman, D. 2017. "Thinking, fast and slow." Macmillan.
- Karlan, D. and M. Valdivia. 2011. "Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions." *The Review of Economics and Statistics* 93 (2): 510–527.
- Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry. 2014. "Agricultural decisions after relaxing credit and risk constraints." *The Quarterly Journal of Economics* 129 (2): 597–652.
- Kasy, M. 2016. "Why experimenters might not always want to randomize, and what they could do instead." *Political Analysis* 24 (3): 324–338.
- Klein, G., Z. Shtudiner, and M. Zwilling. 2021. "Uncovering gender bias in attitudes towards financial advisors." *Journal of Economic Behavior & Organization* 189: 257–273.
- Kosec, K. and L. Wantchekon. 2020. "Can information improve rural governance and service delivery?" *World Development* 125: 1–13.
- Lane, G., D. Schonholzer, and E. Kelley. 2022. "Information and Strategy in Lemon Markets: Improving Safety in Informal Transit." *PEDL Research Paper* .

- Le Barbanchon, T. and J. Sauvagnat. 2021. “Electoral Competition, Voter Bias, and Women in Politics.” *Journal of the European Economic Association* 20 (1): 352–394.
- Leathers, H. D. and M. Smale. 1991. “A Bayesian approach to explaining sequential adoption of components of a technological package.” *American Journal of Agricultural Economics* 73 (3): 734–742.
- Lecoutere, E., D. J. Spielman, and B. Van Campenhout. 2023. “Empowering women through targeting information or role models: Evidence from an experiment in agricultural extension in Uganda.” *World Development* 167: 106240.
- Ledgerwood, A. and A. E. Boydston. 2014. “Sticky prospects: Loss frames are cognitively stickier than gain frames.” *Journal of Experimental Psychology: General* 143 (1): 376.
- Leonard, D. K., G. Bloom, K. Hanson, J. O’Farrell, and N. Spicer. 2013. “Institutional Solutions to the Asymmetric Information Problem in Health and Development Services for the Poor.” *World Development* 48: 71–87.
- Liang, K.-Y. and S. L. Zeger. 1986. “Longitudinal data analysis using generalized linear models.” *Biometrika* 73 (1): 13–22.
- Lin, W. 2013. “Agnostic notes on regression adjustments to experimental data: Reexamining Freedman’s critique.” *Annals of Applied Statistics* 7 (1): 295–318.
- Lipton, M. and R. Longhurst. 1989. *New seeds and poor people*. Routledge.
- Loimeier, R. 2005. “The Baraza - A Grassroots Institution.” *ISIM Review* (16): 26–27.
- López-Moctezuma, G., L. Wantchekon, D. Rubenson, T. Fujiwara, and C. Pe Lero. 2020. “Policy Deliberation and Voter Persuasion: Experimental Evidence from an Election in the Philippines.” *American Journal of Political Science* 66 (1).
- Lybbert, T. J. and A. Bell. 2010. “Stochastic benefit streams, learning, and technology diffusion: Why drought tolerance is not the new Bt.” *AgBioForum* 13 (1): 13–24.

- Lybbert, T. J. and D. A. Sumner. 2012. "Agricultural technologies for climate change in developing countries: Policy options for innovation and technology diffusion." *Food Policy* 37 (1): 114–123.
- Ma, X. and G. Shi. 2015. "A dynamic adoption model with Bayesian learning: An application to U.S. Soybean farmers." *Agricultural Economics* 46 (1): 25–38.
- Magruder, J. R. 2018. "An Assessment of Experimental Evidence on Agricultural Technology Adoption in Developing Countries." *Annual Review of Resource Economics* 10 (1): 299–316.
- McKenzie, D. and C. Woodruff. 2013. "What Are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World?" *The World Bank Research Observer* 29 (1): 48–82.
- Michelson, H., A. Fairbairn, B. Ellison, A. Maertens, and V. Manyong. 2021. "Misperceived quality: Fertilizer in Tanzania." *Journal of Development Economics* 148: 102579.
- Miehe, C., B. Van Campenhout, R. Sparrow, and D. Spielman. 2023. "Lifting quality constraints to agricultural technology adoption in the Ugandan market for maize seed." *Unpublished manuscript* .
- Mitchell, K. M. and J. Martin. 2018. "Gender bias in student evaluations." *PS: Political Science & Politics* 51 (3): 648–652.
- Moser, C. M. and C. B. Barrett. 2006. "The complex dynamics of smallholder technology adoption: The case of SRI in Madagascar." *Agricultural Economics* 35 (3): 373–388.
- Muralidharan, K., M. Romero, and K. Wüthrich. 2023. "Factorial designs, model selection, and (incorrect) inference in randomized experiments." *The Review of Economics and Statistics* .
- Mustalahti, I. and O. S. Rakotonarivo. 2014. "REDD+ and empowered deliberative democracy: Learning from Tanzania." *World Development* 59: 199–211.
- Naanyu, V., J. E. Sidle, R. M. Frankel, D. Ayuku, W. M. Nyandiko, and T. S. Inui. 2011. "Rooting Inquiry in Tradition: The Health Baraza as a Tool for Social Research in Kenya." *Qualitative Health Research* 1 (21): 14–26.

- Nguyen, T. 2008. "Information, role models and perceived returns to education: Experimental evidence from Madagascar." *Unpublished manuscript* 6.
- Ntakyo, P. R. and M. Van Den Berg. 2022. "The Unintended Side-Effects of a Major Development Strategy: Commercialization of Smallholder Production and Women Empowerment in Uganda." *The Journal of Development Studies* 58 (8): 1605–1626.
- Olken, B. A. 2007. "Monitoring corruption: Evidence from a field experiment in Indonesia." *Journal of Political Economy* 115 (2): 200–249.
- Omotilewa, O. J., J. Ricker-Gilbert, and J. H. Ainembabazi. 2019. "Subsidies for agricultural technology adoption: Evidence from a randomized experiment with improved grain storage bags in Uganda." *American Journal of Agricultural Economics* 101 (3): 753–772.
- OPM. 2013. *Implementation of the Baraza Initiative: Progress Report*. Tech. rep.
- Orr, A., T. Tsusaka, S. H. Kee-Tui, and H. Msere. 2016. "What do we mean by 'women's crops'? Commercialisation, gender and the power to name." *Journal of International Development* 28 (6): 919–937.
- Pair, E., N. Vicas, A. M. Weber, V. Meausoone, J. Zou, A. Njuguna, and G. L. Darmstadt. 2021. "Quantification of Gender Bias and Sentiment Toward Political Leaders Over 20 Years of Kenyan News Using Natural Language Processing." *Frontiers in Psychology* 12.
- Pandey, P., S. Goyal, and V. Sundararaman. 2009. "Community participation in public schools: Impact of information campaigns in three Indian states." *Education Economics* 17 (3): 355–375.
- Porter, C. and D. Serra. 2020. "Gender Differences in the Choice of Major: The Importance of Female Role Models." *American Economic Journal: Applied Economics* 12 (3): 226–54.
- Quiñones, M. A., N. E. Borlaug, and C. R. Dowswell. 1997. *A fertilizer-based green revolution for Africa*, chap. 4, 81–95. John Wiley & Sons, Ltd.
- Raffler, P., D. N. Posner, and D. Parkerson. 2020. "Can citizen pressure be induced to improve public service provision." *Unpublished manuscript* .

- Reimers, I. and J. Waldfogel. 2021. "Digitization and pre-purchase information: The causal and welfare impacts of reviews and crowd ratings." *American Economic Review* 111 (6): 1944–71.
- Reinikka, R. and J. Svensson. 2004. "Local capture: Evidence from a central government transfer program in Uganda." *The Quarterly Journal of Economics* 119 (2): 679–705.
- Riley, E. 2022. "Role models in movies: The impact of queen of Katwe on students' educational attainment." *Review of Economics and Statistics* 1–48.
- Rosenberg, S. W. 2007. *Deliberation, participation and democracy*. Springer.
- Sheahan, M. and C. B. Barrett. 2017. "Ten striking facts about agricultural input use in sub-Saharan Africa." *Food Policy* 67: 12–25.
- Shiferaw, B., T. Kebede, M. Kassie, and M. Fisher. 2015. "Market imperfections, access to information and technology adoption in Uganda: Challenges of overcoming multiple constraints." *Agricultural Economics* 46 (4): 475–488.
- Stifel, D. and B. Minten. 2008. "Isolation and agricultural productivity." *Agricultural Economics* 39 (1): 1–15.
- Stiglitz, J. E. 2000. "The Contributions of the Economics of Information to Twentieth Century Economics." *The Quarterly Journal of Economics* 115 (4): 1441–1478.
- . 2002. "Information and the Change in the Paradigm in Economics." *The American Economic Review* 92 (3): 460–501.
- Sumberg, J., D. Keeney, and B. Dempsey. 2012. "Public agronomy: Norman Borlaug as 'brand hero' for the Green Revolution." *Journal of Development Studies* 48 (11): 1587–1600.
- Sunding, D. and D. Zilberman. 2001. "The agricultural innovation process: Research and technology adoption in a changing agricultural sector." *Handbook of Agricultural Economics* 1: 207–261.
- Suri, T. 2011. "Selection and comparative advantage in technology adoption." *Econometrica* 79 (1): 159–209.

- Suri, T. and C. Udry. 2022. “Agricultural technology in Africa.” *Journal of Economic Perspectives* 36 (1): 33–56.
- Thomas, D. 1990. “Intra-Household Resource Allocation: An Inferential Approach.” *The Journal of Human Resources* 25 (4): 635–664.
- Tilman, D., C. Balzer, J. Hill, and B. L. Befort. 2011. “Global food demand and the sustainable intensification of agriculture.” *Proceedings of the National Academy of Sciences* 108 (50): 20260–20264.
- Tjernström, E., T. J. Lybbert, R. F. Hernández, and J. S. Correa. 2021. “Learning by (virtually) doing: Experimentation and belief updating in smallholder agriculture.” *Journal of Economic Behavior & Organization* 189: 28–50.
- Tripp, R. 2002. “Can biotechnology reach the poor? The adequacy of information and seed delivery.” In “Economic and Social Issues in Agricultural Biotechnology,” 237–250. Wallingford UK: CABI Publishing.
- Tripp, R. and D. Rohrbach. 2001. “Policies for African seed enterprise development.” *Food Policy* 26 (2): 147–161.
- Van Campenhout, B. 2021. “The role of information in agricultural technology adoption: Experimental evidence from rice farmers in Uganda.” *Economic Development and Cultural Change* 69 (3): 1239–1272.
- Van Campenhout, B. and A. De. 2023. “Gendered Perceptions in Maize Supply Chains: Evidence from Uganda.” *Development Policy Review* 41: e12662.
- Van Campenhout, B., E. Bizimungu, J. Smart, and N. Kabunga. 2018. “Impact pathways of a participatory local governance initiative in Uganda: A qualitative exploration.” *Development in Practice* 28 (8): 1046–1056.
- Van Campenhout, B., D. J. Spielman, and E. Lecoutere. 2021. “Information and communication technologies to provide agricultural advice to smallholder farmers: Experimental evidence from Uganda.” *American Journal of Agricultural Economics* 103 (1): 317–337.
- Waddington, H., A. Sonnenfeld, J. Finetti, M. Gaarder, D. John, and J. Stevenson. 2019. “Citizen engagement in public services in low-and middle-income countries: A mixed-methods systematic review of participation, inclusion, transparency and accountability (PITA) initiatives.” *Campbell Systematic Reviews* 15 (1-2): 1–90.

- Wossen, T., K. A. Abay, and T. Abdoulaye. 2022. “Misperceiving and misreporting input quality: Implications for input use and productivity.” *Journal of Development Economics* 157: 102869.
- Wu, A. H. 2020. “Gender bias among professionals: An identity-based interpretation.” *Review of Economics and Statistics* 102 (5): 867–880.
- Xie, Y. 2017. *Dynamic Documents with R and knitr*. CRC Press.

Doctoral dissertations from the Faculty of Economics and Business, see:
<https://www.kuleuven.be/english/research/doctoraldefences/archive>.