Gendered Perceptions in Maize Supply Chains: Evidence from Uganda

Bjorn Van Campenhout*†, Anusha De†

Abstract

In situations with imperfect information, the way that value chain actors perceive each other is an important determinant of the value chain's structure and performance. Inaccurate perceptions may result in inefficient value chains, and systematic bias in perceptions may affect inclusiveness. In a case study on perceptions in Ugandan maize supply chains, a random sample of farmers were asked to rate upstream and downstream value chain actors—agro-input dealers, traders, and processors—on a set of important attributes that included ease of access, quality of services rendered, price competitiveness, and overall reputation. These value chain actors were then tracked and asked to assess themselves on the same set of attributes. We find that input dealers, traders, and processors assess themselves more favourably than farmers do. We also focus on heterogeneity in perceptions related to gender and find that for self-assessments, the gender of the value chain actor does not matter. However, the difference between how actors assess themselves and how farmers perceive them is larger for male than for female farmers, as female farmers appear to rate dealers, traders, and processors significantly higher in several dimensions. The gender of the actor being rated does not affect the rating they receive, and gender-based homophily among women is not present in rating behaviour.

1 Introduction

Neoclassical economic theory is built on the premise that rational agents interact in a context of full information. However, in the real world, both consumers and producers face substantial information frictions. Sometimes, agents lack the skills to correctly assess information about their counterparts. In other cases, agents may strategically decide to hide valuable information. As a result, when decisions need to be made, economic agents usually rely on incomplete information that is combined with heuristic techniques prone to bias and updated as new information becomes available. Information frictions may likewise exist in commodity supply chains, especially in informal value chains where quality is hard to track and agreements difficult to enforce. As a result, value chain actors base a substantial part of their decisions on perceptions and beliefs about actors up- and downstream.

Perceptions that do not align with reality can have significant consequences for the entire supply chain. Erroneous perceptions may lead to inefficient supply chains and can hamper value chain innovation. More

*Development Strategy and Governance Division, International Food Policy Research Institute, Kampala, Uganda
†LICOS Center for Institutions and Economic Performance, KU Leuven, Belgium

importantly, systematic bias in perceptions may hamper the inclusiveness of value chains. For instance, if traders of a certain clan or tribe are traditionally regarded as being good, other actors may experience barriers to entry. Despite the importance of beliefs and perceptions for transactions within food supply chains, there are few studies that track perceptions throughout the chain, partly because perceptions are not easy to measure.

In this paper, we study how perceptions align throughout maize value chains in Uganda, with a particular focus on heterogeneity related to gender. To do so, a representative sample of 1,526 maize farmers were asked to rate—on a scale of 1 (very poor) to 5 (excellent)—agro-input dealers, maize traders, and maize processors, on dimensions such as ease of access, quality of services rendered, price competitiveness, and overall reputation. These agro-input dealers, traders, and processors were then traced and asked to assess themselves on the same dimension, resulting in self-assessments of 78 agro-input dealers, 341 assembly traders, and 174 processors. This information was then used to document (in)consistencies between how farmers perceive input dealers, traders, and processors, and how these actors perceive themselves. To investigate systematic bias along gender lines, we further tested if the gender of the farmer and/or actor that was being rated had an impact on the ratings.

Ratings are often used to reveal perceptions. Advances in information and communication technology have facilitated the use of simple rating applications at a large scale to reveal consumer perceptions of a variety of products and services (Reimers and Waldfogel, 2021). Self-ratings have also been used to assess actors' own performance in various settings (e.g. Horswill et al., 2013). Perceptions have been found to correlate with innovations in food supply chains and help identify performance issues of actors involved (Ola and Menapace, 2020; Odongo et al., 2016).

Women are important actors in food supply chains. In many places, women's participation in agricultural labor is increasing, creating more space for women to engage in a sector that is important for poverty reduction (Kawarazuka et al., 2021). Furthermore, although women often only participate in agricultural production as unpaid family laborers, they often have considerably more agency in other nodes of the value chain. For instance, Maertens and Swinnen (2012) find that in Senegal's emerging high-value horticulture supply chains, women participate as paid wage laborers on large-scale estates and in agro-industrial processing. A range of studies examines the opportunities of and constraints faced by women in agricultural value chains (Barrientos, 2019; Kruijssen et al., 2018; Mnimbo et al., 2017). While the inclusion of women in food supply chains is important for food and nutrition security, more inclusive agricultural value chains also hold intrinsic value.

In light of the importance of perceptions in informal value chains, gender-related biases may be an important barrier to the increased participation of women. Systematic differences in ratings related to gender, where women are rated significantly lower than their male counterparts, have been found in a variety of contexts (Mitchell and Martin, 2018; Furnham, 2005; Patiar and Mia, 2008). Gender bias in perceptions can constitute important barriers to market access for women, leading to mistrust, lower access to credit, and other challenges. Bias in self-rating along gender lines may also constrain women's entry and performance by affecting aspirations, which have been found to be an important determinant for success (e.g. Riley, 2017).

The first objective of this study is to establish how the ratings given by farmers align with self-ratings of dealers, traders, and processors. The objective considers a range of questions, such as how agro-input dealers and farmers each assess the quality of the seeds sold by agro-input dealers. Do farmers agree that traders offer a good price when they buy at the farm gate? If farmers think it is difficult to physically access processors, are processors aware of this problem? Second, we focus on the effect of gender on raters and test whether male farmers' ratings are systematically different than those of female farmers. The third objective is to test for gender related differences in self-rating of the agro-input dealers, traders, and processors. A fourth research objective tests whether dealers, traders, and processors are rated differently because of their gender. The final objective tests for the presence of gender-related homophily among women, wherein female farmers give higher ratings to female dealers, traders, and processors (McPherson et al., 2001).

We find that agro-input dealers, traders, and processors consistently rate themselves higher than farmers do except for one attribute that is easily observable by both those who rate and those who are rated. We find that gender does not play a significant role in self-assessments. The gender of the actor being rated does not affect the rating that they receive either, and we find no signs of gender-based homophily among women in the ratings. There is some evidence that female farmers rate more favorably than male farmers but only in some dimensions. In the sections that follow, we expound on the study's context and hypotheses, explain the data and econometric models, present the results, and finally provide a conclusion and reflect on the consequences.

2 The structure of the maize supply chain

We focus on the maize value chain in Uganda. Maize is both a staple and cash crop in Uganda, prioritized by the government for food security and household income. Judged in terms of land area, maize is the most important agricultural commodity in Uganda, covering 30 percent of total cropped land, followed by beans (15 percent of cropped land).

The government's interventions in the maize sub-sector over the past decade have focused on increasing on-farm productivity and production, yet productivity remains low. On-farm maize production data from the Uganda Annual Agricultural Survey (2018) reports average yields of about 600 kg per acre, a figure that falls almost midway between the yield range of 270 and 995 kg per acre found in a recent study focused on estimating maize yields in Uganda (Gourlay et al., 2019). However, these average yields are considerably lower than the figures that research stations report, which range between 730 kg per acre and 1,820 kg per acre (Fermont and Benson, 2011).

Various factors constrain the development of efficient and inclusive value chains in Uganda. Limited use of improved inputs by producers, especially improved seed, constitutes a key supply-side constraint (Gollin et al., 2021). Another important demand-side constraints is access to markets, as farmers are generally located in remote areas with poor roads that become impassable during the rainy season (Stifel and Minten, 2008). Limited processing capacity also leads to issues related to quality and shelf life, further depressing demand for the commodity. As input dealers, farmers, maize traders, and maize processors are intricately linked within Ugandan maize supply chains, performance issues in one node can have consequences for the entire value chain.[1]

2.1 Agro-input dealers

There is ample evidence of the key role of modern agricultural inputs, especially improved seed varieties and inorganic fertilizers, in improved agricultural productivity, poverty reduction, and structural transformation more generally (Gollin et al., 2021; McArthur and McCord, 2017; Evenson and Gollin, 2003). Despite

[1] Throughout this study we will differentiate between farmers and actors, where the latter is used to refer to the agro-input dealers, traders and processors as a group.

decades of policy and institutional reforms to promote their use in low- and middle-income countries, adoption of these inputs remains low, especially in south of the Sahara (Sheahan and Barrett, 2017; Benin, 2016). Some studies point out the limited availability of improved input technologies in low- and middle-income countries (Asfaw et al., 2012; Maredia et al., 2000). However, due to increased government action in the area of research and breeding, privatization/liberalization of the inputs market, and, in some cases, input subsidy programs, improved inputs have become more available in low- and middle-income countries overtime (Minten et al., 2013; Jayne and Rashid, 2013). Lately, the poor quality of purchased inputs is emerging as an additional explanation for limited adoption (Barriga and Fiala, 2020; Ashour et al., 2019; Bold et al., 2017). As such, perceptions related to the conduct and performance of agro-input dealers and the quality of the products they sell will have important consequences for the value chain structure and performance.

In our sample, there is significant heterogeneity in agro-input dealers. Some are large shops located in town centers with several branches that specialize in farm inputs and implements. Others are small shops in villages that only stock seed during the planting season, but generally sell food and other consumables. In our sample, 41 percent of shops are formal businesses operating with required licenses. Agro-input shops are often clustered in towns or trading centers.

Shop owners in our sample are generally well-educated. The average age of agro-input dealers is 36 years, which is also younger than other actors in the value chain. About 29 percent of dealers are women. On average, an agro-input dealer sells three different types of improved maize seed varieties. The average shop sold about 438 kg of hybrid seed and 522 kg of Open Pollinated Varieties during the first agricultural season of 2018.

2.2 Traders

Maize traders link producers to processors and consumers. Using bicycles or light motorbikes, local assembly traders visit several farmers in a day to buy maize at the farm gate. These traders then aggregate and sell further to larger traders or large-scale processors. Larger traders also often use (shared) storage facilities to engage in arbitrage over time, as maize prices are known to display significant seasonality (Van Campenhout et al., 2015).

Assembly trader performance has important consequences for quality downstream. The bulking and mixing of smallholder supply dilutes incentives to supply high-quality maize (Anissa et al., 2021). The procurement of sufficiently dried maize and proper storage and handling are important to reduce aflatoxin contamination (Bauchet et al., 2021). Furthermore, while small traders are often cast in a negative light and many development interventions attempt to bypass middlemen, most studies find that small traders, in sufficiently large numbers, are essential for smallholder market participation (Barrett, 2008). For example, studying market access in southern and eastern Africa, Mather et al. (2013) note that access to assembly traders has increased over time and created important opportunities for remote areas to access maize markets. Sitko and Jayne (2014) find that marketing margins and number of traders create highly competitive conditions for trading in eastern and southern Africa.

Generally, the traders in our sample are not just service providers offering transport services. They also become owners of the commodity during trading, hence internalizing the risks inherent to trading. On a typical day immediately after harvest when prices are typically lowest and most traders are active, traders visit 12 smallholder farmers to collect 1,308 kgs of maize. Downstream, the average trader delivers approximately four different buyers during the peak season. The average storage capacity of a typical trader is about 13,000 kgs. Ninety-three percent of traders indicate that they also trade in other agricultural commodities aside from maize. Only 7 traders out of 341 were women.

2.3 Processors

A third important actor in the maize supply chain is the processor. In general, processors operate maize mills that remove the bran from the maize kernels and mill the maize into maize flour for direct consumption. Some processors also have packing facilities to produce maize flour for supermarkets or export. The smaller mills often provide milling as a service, wherein farmers pay a fee to have bags of maize milled.

There are again significant differences between these processors. Some maize mills in remote rural areas are diesel engine-powered mills that are only able to produce low-grade maize meal for home consumption. However, larger processors use machines powered by three-phase electric power and mill multiple times to obtain fine export-grade maize flour. In our sample, we find that about 57 percent of the millers use diesel-powered mills, while 37 percent rely on three-phase electric power. The quality of the end product is indicated by grades, from highest (grade 1) to lowest (grade 3). The grade that can be obtained depends on a various factors, including the number of times the product is passed through the mill, the quality of the grain used, and the type of mill. About 44 percent of processors indicate that they can also produce grade 1 flour. In our sample, only 6.9 percent of processors are women.

2.4 Farmers

Central to our study are smallholder maize farmers, who buy inputs such as seed and fertilizer from agro-input dealers, sell to traders, and/or use millers to process their maize for own consumption. These are generally small farmers, cultivating maize on one or two plots that correspond to 1.81 acres on an average. The average age of the farmers is 44.5 years, and 43 percent only finished primary education. The average distance from the farm homestead to a tarmac road is 6.54 km and to an all-weather feeder road is 0.51 km.

About 53 percent of farmers used improved seed on their plots, while 21 percent used inorganic fertilizer. Yields in our sample amount to 468 kg/acre. More than 95 percent of farmers in our sample report that they took part of the maize they harvested to a miller. In terms of market participation, we find that 64 percent of farmers sold at least part of their crop. Those who reported selling sold on average 706 kg, which corresponds to about 55 percent of the total maize harvest. In our sample, 49 percent of farmers are women.

3 Study hypotheses

This section describes the five hypotheses that we tested and the theory in which these hypotheses are grounded.

- Hypothesis 1: Self-ratings of dealers, traders, and processors are higher than ratings given to them by farmers.

The first hypothesis revolves around how the agro-input dealers, processors, and traders rate themselves as compared to the scores that are given to them by farmers. A significant (positive) difference could mean that agro-input dealers systematically overestimate their own performance, perhaps as a result of an overconfidence effect. However, research has shown that agents are generally good at assessing their own performance (Clark and Friesen, 2008). At the same time, the difference can also increase if farmers systematically underestimate the performance of other value chain actors.

Perceptions in this context may lead to inefficient value chains and slower upgrades to value chains. If service and input providers perceive themselves to be performing better than they actually are, there may be no incentive to improve. If farmers rate input dealers, traders, and processors lower than their self-ratings, it would indicate that actors are not meeting the farmers' expectations. Cheng et al. (2017) discusses how more favourable self-assessments can be a result of leniency in assessing self-performance. Such a leniency creates a gap between the farmers' perceptions of the input and service providers' performance and the input and service providers' self-perceptions of their performance. If farmers underestimate service quality of other value chain actors, this may lead to depressed demand for the services (Michelson et al., 2021).

- Hypothesis 2: Female farmers rate more favourably than male farmers.

In a second hypothesis, we test whether female farmers rate input dealers, traders, and processors systematically higher than male farmers do. There is some evidence that women generally rate more positively than men (Furnham, 2005; Winquist et al., 1998). More favourable ratings from female farmers may indicate that they received better services and inputs from the input dealers, maize processors, and traders. Alternatively, women may show greater leniency in rating the service and input providers. However, the literature does not mention the statistically significant presence of leniency for ratings provided by female raters (Thornton III et al., 2019).

Women who are more forgiving toward other value chain actors may suffer if these other actors do not reciprocate and provide high quality goods and services. At the same time, relatively more positive feedback from women may mean that service providers also strive to offer high quality goods and services when dealing with women. The fact that female farmers rate relatively higher also provides more scope for disappointment. This may lead to a higher likelihood of women exiting the value chain if reality does not match up with higher (perceived) quality of services and inputs.

- Hypothesis 3: Self-ratings from women are less favourable than self-ratings from men.

The third hypothesis compares the self-ratings given by female agro-input dealers, processors, and traders with the self-rating of their male counterparts. While we saw in hypothesis 1 that actors tend to overestimate their own performance and in hypothesis 2 that women tend to rate others higher than men do, studies suggest that women generally tend to underrate themselves (relative to men). For instance, Patiar and Mia (2008) found that in the hotel industry, male department managers tend to hold inflated self-assessments, while the women assessed themselves lower. Similar patterns have consistently been found in a variety of contexts (e.g., Bengtsson et al., 2005; Beyer, 1990; Rosenkrantz et al., 1968). Furthermore, Braddy et al. (2020) found that women tend to experience harsher consequences than men when they overrate themselves.

Lower self-ratings of women as compared to men may signal a lack of confidence, which may hamper aspirations and restrict women from thriving in business. Cultural norms, societal expectations, and gender stereotyping will also be reflected in self-ratings. The gendered ideas reflected in self-assessments can thus be an important barrier to entry for women, leading to exclusively male-dominated value chains.

- Hypothesis 4: Male agro-input dealers, traders, and processors receive more favourable ratings than their female counterparts.

In hypothesis 4, the ratings given by farmers to female agro-input dealers, traders, and processors are compared to ratings given by farmers to their male counterparts. The fact that women are held to more stringent standards than men has been repeatedly established over time. Lyness and Heilman (2006) found that female managers received lower performance ratings compared to male managers. Basow and Silberg (1987) found that students rated female professors lower than male professors. Bias against female professors has been replicated numerous times (e.g., Feldman, 1993; Mengel et al., 2018; Miller and Chamberlin, 2000). A recent study by Wu (2020) finds that there is a gender bias in how women are perceived in professional circumstances: perceptions about women are generally lower in the professional sphere and higher in the domestic sphere or when physical appearance is judged. Cultural factors play an important role in the creation of bias, given that it is more present in male-dominated sectors.

In the context of this study, bias against women may again create a barrier for entry for women in agro-input dealing, trading, and processing. When women recognize that they are perceived to be less capable of these business activities, they might not enter the sector to avoid criticism and performance obstacles. However, a bias in rating may also lead to actual differences in the quality of service, as women who are perceived to be inferior struggle to attract credit to invest in their activities or secure desirable locations to set up their business. More generally cultural context might limit women from performing well and being equally competitive with men. A study of female entrepreneurs in Kigali, Rwanda by Nsengimana et al. (2017) reports many impediments to the success of women's businesses. Similarly, Guma (2015) discusses barriers faced by female entrepreneurs in Uganda. Some of the prominent issues faced by women are gender-related stereotypes (risk-taking behavior and lower aggressiveness), under-capitalization (limited credit access and availability of collateral), balance across multiple responsibilities (childcare and family responsibilities), inadequate skills and business knowledge, lack of respect from the male-dominated business community, time investment constraints, and reputation and work credibility challenges. These can significantly impact the ratings received by the female input and service providers in the supply chain and can undermine their perceived performance in the sector.

- Hypothesis 5: Female farmers give higher ratings to female agro-input dealers, traders, and processors.

Finally, in hypothesis 5 we test the interaction between the gender of the farmer and the other value chain actors in order to investigate whether same-gender status has a significant impact on the ratings. This hypothesis is motivated by the literature on homophily in social networks. The homophily principle essentially focuses on network ties based on the relationships and characteristics of the actors involved. In the context of this study, gender homophily effects for women in rating would exist if ratings in female rater-ratee pairs (e.g., female farmers rating female agro-input dealers, female traders, or female processors) are consistently higher than ratings in male-male or mixed gender rater-ratee linkages. McPherson et al. (2001) discusses the causes and consequences of such preferences, such as limitations in the social world, biased information, attitudes influenced by the characteristics of the ties, and interactions limited to these homogeneous networks, arguing that gender-based homophily can strongly divide personal environments.

Gender-based homophily in food value chains may lead to the development of several co-existing value chains aligned by gender. If a female farmer gives higher ratings to a female trader, she may always interact with traders of the same sex. The higher ratings and lower levels of competition enjoyed by dealers, traders, and processors may reduce effort and delay innovations. However, recent research suggests that increased competition in value chains characterized by relational contracts (for instance, contracts based on gender-based preferences) is not always good (Macchiavello and Morjaria, 2020). Indeed, the increased trust in relationships mediated by gender homophily may make it easier for women to enter into business.

4 Data

This section explains sampling and data collection, describes how perceptions were measured and the variables constructed, and discusses reliability of the ratings.

4.1 Sample

We use survey data from 1,526 farming households, 78 agro-input dealer shops, 341 assembly traders, and 174 processors operating in the maize growing districts of Iganga, Bugiri, and Namutumba in eastern Uganda. Data were collected in July 2019. The farmer household sample was drawn from 63 villages in the three districts. The villages were selected through a process of stratified random sampling at the sub-county level. In each of the three districts, the sub-counties from which the villages were sampled were purposely selected based on their distance (km) from the main district town, in a range of 10 kms, 20 kms, and 30 kms, from the main town. A map of the study area is provided in figure 1. In each selected village, several households were then randomly selected. The number of households was determined proportionate to the village population using the 2012 sampling frame of Uganda National Bureau of Statistics (UBoS). The input dealer shops, assembly traders, and processors interviewed were those referred to by farmers, either because they had purchased agro-inputs from these agro-input dealers, sold harvested products to these traders, or used these processors to mill maize.

4.2 Ratings used to measure perceptions

This section focuses on the ratings, which were the central indicators used in this study. To obtain the ratings, each farmer was asked to rate between one and three of each of the three value chain actors (input dealers, traders, and processors). Farmer's perceptions about the other value chain actors and the other value chain actors' perceptions of themselves are derived from scores based on four dimensions: (1) location, in which we asked participants to reflect on ease of reaching the actor; (2) quality, in which we asked them to rate the quality of the service rendered and/or product sold; (3) price, in which we asked whether the price charged for the service rendered or product sold was reasonable in relation to what others charged;[2] and (4)

[2]Note that for input dealers and processors, price competitiveness would be rated higher if they charged lower prices to farmers. For traders, price competitiveness would be rated higher if they paid higher prices to farmers at the farm gate.

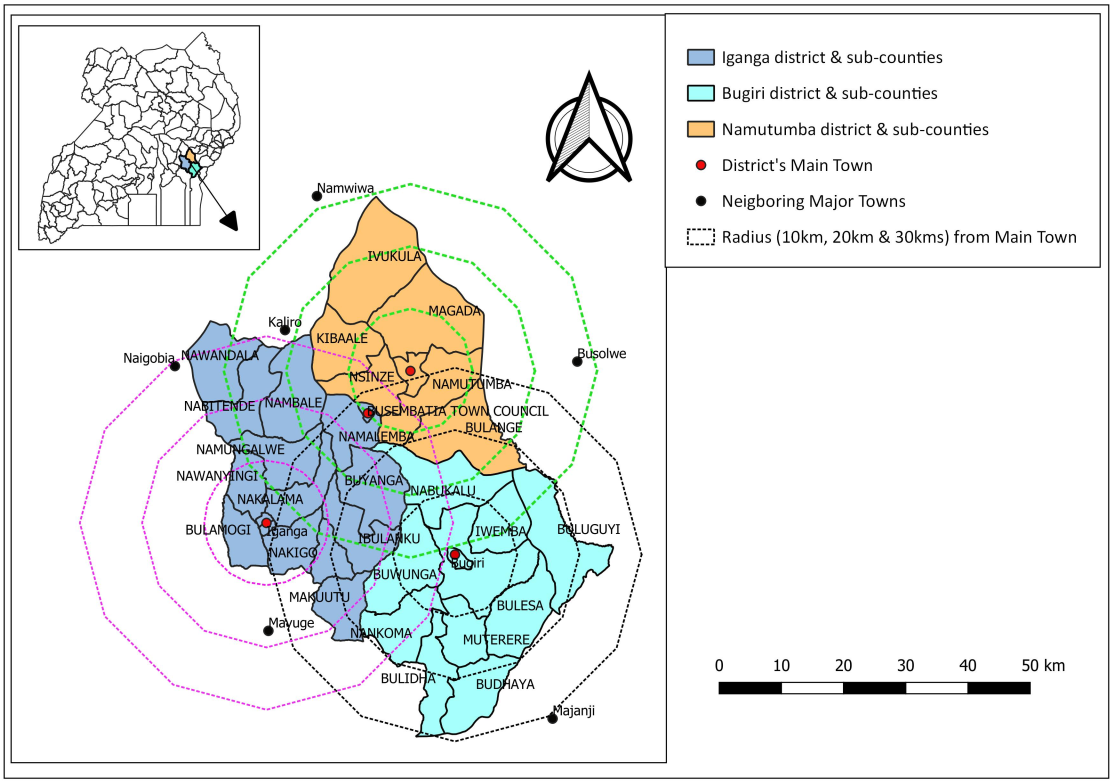

Figure 1: Map of the study area.

reputation, in which we probed about how other farmers think about the actor being rated. The averages of these dimension-based ratings were calculated to form an overall rating.

Several methods have been proposed to measure perceptions, attitudes, or beliefs in social science research. Delavande et al. (2011) survey the literature on the measurement of subjective beliefs in low- and middle-income countries and categorize possible methods into three groups: Likert-style questions, elicitation of the "most likely" outcome, and a full elicitation of the distribution of beliefs, most often conducted with visual aids. The ratings in this study reflect the Likert scale assessment where scores range from 1 to 5, with 1 being the worst score and 5 being the best. This is the case for both farmer ratings and self-ratings from agro-input dealers, traders, and processors.

Table 1 presents ratings obtained from farmers. It shows average scores for all value chain actors (top panel), and average scores for each actor separately in each of the four dimensions. Farmers are generally positive about dealers, traders and processors. For example, only about 6 percent of all ratings given by farmers were the lowest score of 1 while about 38 percent of ratings were a 5. Interestingly, traders seem to receive slightly better ratings than the other actors. Price competitiveness is the dimension that always scores the lowest. Location is scored highest when actors are pooled, which seems to be driven by ease of access to traders. Judging by this table, the biggest constraint to the adoption of commercial seed and other purchased inputs is not quality, but rather price. It is notable that all actors seem to score very well in terms of reputation.

Table 2 shows summary statistics for self-ratings from agro-input dealers, traders, and maize processors. These actors seem very confident about their reputation because among all the dimensions, the majority

(59 percent) give a self-score of 5 for reputation. They seem to be the least confident about their price competitiveness as among all the dimensions, the majority adhere to a score of 3 at most for this dimension.

4.3 Reliability of ratings

In this section we test whether the ratings are actually meaningful (as opposed to just noise). To do so, we consider intra-class correlation (ICC) coefficients determining the level of agreement between the ratings. We include both inter-rater agreement and intra-rater agreement (Gwet, 2014). Inter-rater agreement considers the correlation between ratings given by different farmers to a single actor, while intra-rater agreement is judged by the correlation between ratings received by different actors from a single farmer. Intra-class correlation coefficients range between 0 and 1, with 0 being low agreement and 1 being total agreement. Only farmers who rated more than six times are considered in this analysis.

Table 3 presents the results for the ICC analysis. In the left panel, results for inter-rater agreement are shown. Judging by average ratings, farmers rate different actors fairly consistently. However, farmers disagree more with each other when quality is assessed, or when reputation needs to be rated. This is expected, as location and prices are observable factors and hence, ratings for these factors should be more similar compared to non-observable attributes like quality and reputation.

The right panel of Table 3 shows the results for intra-rater agreement. The results also show that ratings for different actors are consistent among farmers. The fact that intra-rater agreement is higher than inter-rater agreement may indicate some degree of assortative matching within the chain, whereby farmers who select better-performing agro-input dealers also tend to go to better processors and sell to better traders. However, it could also be that ratings are determined more by farmer-level characteristics (such as the education level of the farmer) than by actor-level characteristics, leading farmers to rate different actors in a more consistent way. At the same time, because farmers rate different actors, it also seems reasonable that there is no complete intra-rater agreement.

5 Econometric analysis

We start the analysis with a descriptive section based on simple averages in subgroups of the data. This is followed by a section that presents more formal tests of the hypotheses. To test the first hypothesis, we use simple t-tests. The other hypotheses will be tested in a regression framework. In this section, we elaborate on the specifications we will estimate.

We start from the following multivariate regression model:

$$y_{f,a} = \alpha + \beta_1 g_f^F + \beta_2 g_a^A + \beta_3 g_f^F * g_a^A + \gamma X_{f,a} + \varepsilon_{f,a} \tag{1}$$

Here, $y_{f,a}$ is the primary outcome variable, which will be the rating that farmer i gave to actor a (hypotheses 2, 4 and 5). We will run separate regressions for the ratings for the different dimensions, as well as a regression where $y_{f,a}$ is the average of the ratings of the four dimensions that farmer i gave to actor a.[3] The main variables of interest are the sex of the farmer (g_f^F, a dummy variable that takes the value of

[3] When we look at ratings on a particular dimension, this will be an integer number ranging between 1 and 5. When we look at average ratings, this can also be a rational number. While we agree that the outcome variable is likely to be non-normal, we nevertheless proceed with Ordinary Least Squares, as this gives the conditional mean under minimal assumptions.

1 if the farmer is a woman and 0 otherwise), which varies at the farmer level *f*, and sex of the actor (g_a^A, a dummy variable taking the value of 1 for female actors and 0 otherwise), which varies at the actor level *a*. $X_{f,a}$ is a vector of control variables, some of which vary at the farmer level, like farmer's age (in years), dummy variable indicating if the farmer has finished primary education, distance of farmer's homestead to tarmac and feeder roads (in km), and marital status. Other control variables included in $X_{f,a}$ vary at the level of the actor, like age, education (if primary education is finished), and a dummy variable for marital status of the dealer, trader, or processor. The error term in the model is $\varepsilon_{f,a}$. As the number of women in some actor categories are very low, we decided to run the analysis on the pooled dataset and include three fixed effects for the actor type (dealer, trader, processor). We use the approach proposed by Cameron et al. (2011) to allow for two-way non-nested clustering at the farmer and actor levels.

The coefficients of interest in these models are β_1, β_2, and β_3. In particular, finding that $\beta_1 > 0$ would confirm hypothesis 2, while finding that $\beta_2 < 0$ would confirm 4. Gender homophily among women would mean that $\beta_3 > 0$ (hypothesis 5).

To test hypothesis 3, a regression based on the actors' self-rating data only is used:

$$y_a = \alpha + \beta_1 g_a^A + \gamma X_a + \varepsilon_a \tag{2}$$

Here, the primary outcome variable is the self-rating y_a of actor a which is regressed on the sex of the actor $\left(g_a^A\right)$. Finding that $\beta_1 < 0$ would confirm hypothesis 3. Also here, we include a range of control variables (X_a), including fixed effects for the type of actor (dealer, trader, or processor). The error term in this model is ε_a.

As we rely on observational data, we control for confounding bias through the inclusion of exogenous control variables. Men are likely to be better educated than women. Better levels of education and knowledge will probably mean that farmers have a better understanding of what to expect from service and input providers and thus may rate more or less favourably, as scores given will be better informed. The impact pathway must be controlled for or else the gender and education effects will be conflated. The age of the farmer may also affect ratings in some way. In our sample of farm households, women are likely to be younger than men (Jensen and Thornton, 2003), so age effects need to be purged from the model.

An interviewee's marital status may also be correlated with ratings. It may be that single household heads are more likely to interact with lower-rated agents (for instance, predatory traders who target households with only one head). At the same time, the women we interviewed in our sample are also more likely to be unmarried, so we need to control for the effect of marital status on ratings that are affected by gender. Distance to murram and tarmac roads are proxies of remoteness. In remote areas, input and service providers face many challenges, such as larger transaction costs and poor access to services such as electricity. For instance, in semi-urban areas, mills often run on three-phase electricity, while in remote areas, combustion engines are used to power the mills. The latter produce inferior maize products. If women are also more likely to reside in remote areas, this may lead to biased coefficient estimates.

Similar arguments can be made for the gender of the dealers, traders, and processors. Since men are likely to be better educated than women, their education and knowledge might define what kind of services and inputs they provide to their customers. This would lead to more or less favourable ratings from farmers, so we need to control for education to disentangle the effects of actor gender and education levels. Men are likely to be older, with more experience in service providing and input dealing. Older individuals might have more experience in business, which can impact the ratings. Controlling for marital status of input dealers, traders, and processors is necessary, as female value chain actors may be more likely to be single.

6 Results

6.1 Average ratings

In this section, we provide a descriptive account of the hypotheses outlined in Section 3 and based on the subgroup averages reported in Table 4. The table shows scores aggregated over all input dealers (first three columns on the left), but also scores for each actor type separately (columns 4–6 for agro-input dealers, columns 7–9 for traders, and columns 10–12 for processors). We further differentiate between scores received by male and female actors.

The rows represent the different dimensions for which actors were scored (or were asked to rate themselves). We again start with an overall rating (rows 1–4) and then provide separate ratings for location (5–8), quality (9–12), and reputation (rows 13–16). We also differentiate between the gender of the farmer and also add a line for self-ratings.

In line with hypothesis 1, Table 4 shows that the mean overall self-rating given by the actors (4.22) is substantially higher than the mean overall rating that farmers give to actors (3.6). This pattern is consistent across all the different rating dimensions. In examining individual groups of input dealers, traders, and processors in Table 4, self-ratings are also always higher.

In line with hypothesis 2, we find that the mean overall rating provided by female farmers (3.62) is slightly higher than the mean overall rating given by male farmers (3.58). We similarly see that location-, price-, and reputation-based ratings are higher among female farmers than among male farmers. However, for quality-based rating, male farmers give a higher rating.

Looking at average ratings across actor types, we again find that women consistently rate higher than men, but the margin is small. Women rate traders rate more favourably in all dimensions. Women also generally rate processors more favorably, except for reputation where ratings between men and women are effectively the same. Female farmers rate dealers lower on both reputation and price competitiveness than male farmers do. In all, out of 20 comparisons, 16 show agreement with hypothesis 2.

Next, we focus on self-ratings from female and male value chain actors (hypothesis 3). For overall average self-rating, female actors rate themselves lower than how male actors rate themselves (4.16 vs. 4.23). We also find women rate themselves lower on location and reputation dimensions. However, for quality and price, men assess themselves worse than women do.

There is no clear pattern when we consider different actors' self-ratings. While female traders assess themselves higher than their male counterparts, male input dealers rate themselves higher than female dealers' self-assessments. When evaluating across the four dimensions, we find that for agro-input dealers, men rate themselves higher on three of the four dimensions. For processors, the opposite trend emerges: women rate themselves higher on two of the four dimensions and also, for the overall average. For traders, women consistently rate themselves higher. However, due to the small number of female traders, this result needs to be interpreted with caution. Overall, we see that from the 20 comparisons, 9 are in line with hypothesis 3, indicating that the hypothesis is likely to be false.

In hypothesis 4, we test whether the gender of the actor leads to systematically different ratings from farmers. Judging from the overall score, male value chain actors receive lower scores than female actors, but the difference is negligible (3.59 vs. 3.61). When all actors are pooled, we see that male actors are scored higher on location and price competitiveness, but they are scored lower on the dimensions of quality and reputation.

When we examine the different actors in more detail, we find no systematic difference in the overall ratings between men and women. As is the case when all actors are pooled, female dealers are viewed more favorably for quality and reputation, while male dealers receive higher scores for location and price. For traders, women consistently receive higher scores than men, but again, these results need to be interpreted with care given the very small number of female traders in the sample. For processors, men seem to get somewhat higher ratings, except perhaps on the quality dimension. Overall, the descriptive analysis provides little support for hypothesis 4.

Finally, we seek indications of female gender homophily (hypothesis 5). Aggregating across actors and dimensions, we see that female farmers score female actors higher (3.63) than any other gender combination. But if we examine the different dimensions, there is no indication of female gender homophily. For location and price competitiveness, female farmers score male actors highest; for quality, male farmers score female actors highest. For reputation, male and female farmers give the same score to female actors. When looking at dealers, traders, and processors separately, we only find signs of female gender homophily for traders. But again, these results likely suffer from the small sample size. For other actors, there also seems to be no indication of female gender homophily, leading us to reject hypothesis 5.

6.2 Regressions

To test hypothesis 1 formally, we test whether the difference between an actor's self-rating and a rating of the actor is significantly larger than zero. Table 5 shows that we reject the null hypothesis that the difference is equal to zero in favour of the alternative hypothesis that actors rate themselves systematically higher than farmers do.

Formal testing of hypotheses 2, 4, and 5 is done by estimating Regression Model 1 outlined in Section 5, the results of which are reported in Table 6. Taking overall ratings as the dependent variable in column 1, we cannot reject the null hypothesis that the sex of the farmers does not affect the rating given (hypothesis 2). However, if we review the different components of the rating index, we see that female farmers rate actors significantly higher when asked to assess location and price competitiveness (columns 2 and 4).

The gender of the actor being rated does not seem to be significantly correlated with ratings given by the farmer. As such, we do not find evidence for hypothesis 4 in our data.[4] Looking at the interaction between the gender of the farmer and the gender of the actor, we also do not find any significant correlation, suggesting there is no female gender homophily effect (hypothesis 5).

Formal testing of hypothesis 3 is done by estimating Regression Model 2 outlined in Section 5, the results of which are reported in Table 7. While female actors seem to rate themselves somewhat higher on the quality dimension, the difference with men is not significant. We certainly do not find that men rate themselves higher than women, leading us to reject hypothesis 3.

[4]In fact, there are some indications that women score higher on quality and reputation, which runs against hypothesis 4, but differences are not significant.

We also ran an additional regression similar to Model 1, but used the difference between actor self-ratings and farmer ratings as the dependent variable. Results are in Table 8. This provides an alternative way to test hypothesis 1 by looking at the significance of the constant in Model 1. Interestingly, in a regression framework that controls for a range of farmer- and actor-level characteristics, there is no significant difference between actor ratings and farmer ratings for the location dimension. This seems plausible, as location can be easily observed by both farmer and actor. The table also shows that the gender of the farmer now also becomes significantly negative, which makes sense as women rate more positively (hypothesis 2), making the gap between actor and farmer ratings smaller.

The impact of the sex of the actor, although often not statistically significant, is also interesting. The difference between self-ratings and ratings from farmers increases substantially for quality and price if the actor is a woman, while the difference declines in the case of reputation-based ratings. Women rate themselves higher on quality compared to men, but farmer ratings on quality are not dependent on the sex of the actor being rated. On price, the coefficient of the actor's sex is also positive for self-ratings, but here the lower rating from the farmers of female actors seems to make the gap larger between self-ratings and ratings given by farmers. For reputation, female actors rate themselves lower compared to the self-ratings of male actors: the combination of a negative gender effect on self-ratings and a positive gender effect on farmer ratings significantly reduces the gap.

7 Conclusion

In informal food supply chains, perceptions about the quality of services of value chain actors such as agro-input dealers, traders, and processors are an important input in the decision-making processes underlying the structure, conduct, and performance of the value chain. We thus investigated perceptions of maize farmers about input and service providers in informal maize value chains, as well as these input and service providers' self-perceptions. We were particularly interested in gender-based heterogeneity in these perceptions. Perceptions were captured through ratings given on dimensions that included ease of access, quality of service, price competitiveness, and reputation.

We find that agro-input dealers, traders, and processors consistently rate themselves higher than farmers rate them, except for one attribute—gender—that is easily observable by both those who rate and those who get rated. We do not find that gender plays a significant role in self-assessments, except perhaps for the fact that women seem to rate themselves somewhat higher on the quality dimension. The sex of the actor being rated does not affect the rating that they receive and we find no signs of gender-based homophily for women in the ratings. There is some evidence that female farmers rate more favorably than male farmers, but only in some dimensions. Taken together, women actors rate themselves relatively higher and farmers rate female actors relatively lower when price and quality considerations are concerned. Female actors rate themselves lower in regard to reputation.

The data did not support one of study's key hypotheses that female agro-input dealers, traders, and processors are systematically rated lower than male actors. Still, given the extensive literature that finds discrimination in a variety of contexts (eg. Mengel et al., 2018; Mitchell and Martin, 2018; Lyness and Heilman, 2006), we caution against sweeping conclusions. Heterogeneous effects between actors may potentially result from low sample size and limited variation in the gender of the actor. For instance, we do find that male agro-input dealers get higher scores for location than women, an effect which may become significant if the sample size grows.

The fact that self-assessments are always higher than farmer ratings may either mean that actors are overconfident or farmers are overcritical. Overconfidence of value chain actors may mean that actors do not see the nee to improve, which may delay innovations within the chain. Farmers that expect more from value chain actors are likely to demand more limited services from these actors. As such, policy interventions aimed at reducing the gap between actor self-assessments and farmer ratings are likely to increase efficiency in value chains. Examples of such policy interventions include certification by independent agencies or non-centralized clearinghouse mechanisms based on crowdsourcing (Reimers and Waldfogel, 2021; Hasanain et al., 2019).

Even though we did not find that farmers rate female actors differently, gender may still affect the inclusiveness of value chains. For instance, the tendency of female farmers to rate more favourably may result in input and service providers treating female farmers differently.[5] The fact that female farmers provide higher ratings also means that the gap between the self-assessments of actors and ratings by female farmers is smaller than the gap between the self-assessments of actors and ratings by male farmers. This gap may lead to differences in the efficiency of the chain conditional on the gender of the farmers.

Finally, judging by farmer perceptions, there seem to be issues related to price competitiveness within the maize value chain, as agro-input dealers, traders, and processors are consistently scored lowest. Quality, on the other hand, seems to be rated fairly well. Input dealers, for example, are perceived to be performing poorly in terms of price competitiveness but receive the highest scores for quality. This seems to contradict recent studies that blame the poor quality of inputs as the root cause for low adoption by farmers (Bold et al., 2017). Rather, policies that encourage market entry and competition between agro-input dealers, traders, and processors are likely to increase the price competitiveness of value chain actors.

Acknowledgments

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 (https://www.cgiar.org/funders/). 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). We thank Richard Ariong, Leocardia Nabwire, Wilberforce Walukano, and Marc Charles Wanume for excellent field support. The analysis contained here is the sole responsibility of the authors, and does not reflect the views of any funding agency or organization mentioned here.

References

Anissa, B. P., Abate, G., Bernard, T., and Bulte, E. (2021). Is the local wheat market a 'market for lemons'? Certifying the supply of individual wheat farmers in Ethiopia. *European Review of Agricultural Economics*.

Asfaw, S., Shiferaw, B., Simtowe, F., and Lipper, L. (2012). Impact of modern agricultural technologies on smallholder welfare: Evidence from Tanzania and Ethiopia. *Food Policy*, 37(3):283–295.

Ashour, M., Gilligan, D. O., Hoel, J. B., and Karachiwalla, N. I. (2019). Do beliefs about herbicide quality

[5] The leniency of female farmers may mean actors are not compelled to provide high-quality inputs or services. Alternatively, positive feedback may encourage actors to provide higher-quality inputs and services.

correspond with actual quality in local markets? Evidence from Uganda. *The Journal of Development Studies*, 55(6):1285–1306.

Barrett, C. B. (2008). Smallholder market participation: Concepts and evidence from eastern and southern Africa. *Food Policy*, 33(4):299–317.

Barrientos, S. (2019). *Gender and work in global value chains: Capturing the gains?* Cambridge, UK: Cambridge University Press.

Barriga, A. and Fiala, N. (2020). The supply chain for seed in Uganda: Where does it go wrong? *World Development*, 130:104928.

Basow, S. A. and Silberg, N. T. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology*, 79(3):308.

Bauchet, J., Prieto, S., and Ricker-Gilbert, J. (2021). Improved drying and storage practices that reduce aflatoxins in stored maize: experimental evidence from smallholders in Senegal. *American Journal of Agricultural Economics*, 103(1):296–316.

Bengtsson, C., Persson, M., and Willenhag, P. (2005). Gender and overconfidence. *Economics Letters*, 86(2):199–203.

Benin, S. (2016). *Agricultural productivity in Africa: Trends, patterns, and determinants.* Washington, DC: International Food Policy Research Institute.

Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology*, 59(5):960–970.

Bold, T., Kaizzi, K. C., Svensson, J., and Yanagizawa-Drott, D. (2017). Lemon technologies and adoption: Measurement, theory and evidence from agricultural markets in Uganda. *The Quarterly Journal of Economics*, 132(3):1055–1100.

Braddy, P. W., Sturm, R. E., Atwater, L., Taylor, S. N., and McKee, R. A. (2020). Gender bias still plagues the workplace: Looking at derailment risk and performance with self-other ratings. *Group & Organization Management*, 45(3):315–350.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2):238–249.

Cheng, K. H. C., Hui, C. H., and Cascio, W. F. (2017). Leniency bias in performance ratings: The big-five correlates. *Frontiers in Psychology*, 8:521.

Clark, J. and Friesen, L. (2008). Overconfidence in forecasts of own performance: An experimental study. *The Economic Journal*, 119(534):229–251.

Delavande, A., Giné, X., and McKenzie, D. (2011). Measuring subjective expectations in developing countries: A critical review and new evidence. *Journal of Development Economics*, 94(2):151–163.

Evenson, R. E. and Gollin, D. (2003). Assessing the impact of the green revolution, 1960 to 2000. *Science*, 300(5620):758–762.

Feldman, K. A. (1993). Gender and university teaching: A negotiated difference. *The Journal of Higher Education*, 64(1):117–120.

Fermont, A. and Benson, T. (2011). *Estimating yield of food crops grown by smallholder farmers*. Washington, DC: International Food Policy Research Institute.

Furnham, A. (2005). Gender and personality differences in self- and other ratings of business intelligence. *British Journal of Management*, 16(2):91–103.

Gollin, D., Hansen, C. W., and Wingender, A. M. (2021). Two blades of grass: The impact of the green revolution. *Journal of Political Economy*, 129(8):2344–2384.

Gourlay, S., Kilic, T., and Lobell, D. B. (2019). A new spin on an old debate: Errors in farmer-reported production and their implications for inverse scale-productivity relationship in Uganda. *Journal of Development Economics*, 141:102376.

Guma, P. K. (2015). Business in the urban informal economy: Barriers to women's entrepreneurship in Uganda. *Journal of African Business*, 16(3):305–321.

Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters*. Maryland, USA: Advanced Analytics, LLC.

Hasanain, A., Khan, M. Y., and Rezaee, A. (2019). No bulls: Experimental evidence on the impact of veterinarian ratings in Pakistan. Under Review.

Horswill, M. S., Sullivan, K., Lurie-Beck, J. K., and Smith, S. (2013). How realistic are older drivers' ratings of their driving ability? *Accident Analysis & Prevention*, 50:130–137.

Jayne, T. S. and Rashid, S. (2013). Input subsidy programs in sub-Saharan Africa: A synthesis of recent evidence. *Agricultural Economics*, 44(6):547–562.

Jensen, R. and Thornton, R. (2003). Early female marriage in the developing world. *Gender & Development*, 11(2):9–19.

Kawarazuka, N., Doss, N., Farnworth, C., and Pyburn, R. (2021). Myths about the feminization of agriculture: Implications for global food security. Technical report, Washington, DC: International Food Policy Research Institute.

Kruijssen, F., McDougall, C. L., and van Asseldonk, I. J. (2018). Gender and aquaculture value chains: A review of key issues and implications for research. *Aquaculture*, 493:328–337.

Lyness, K. S. and Heilman, M. E. (2006). When fit is fundamental: Performance evaluations and promotions of upper-level female and male managers. *Journal of Applied Psychology*, 91(4):777.

Macchiavello, R. and Morjaria, A. (2020). Competition and relational contracts in the Rwanda coffee chain. *The Quarterly Journal of Economics*, 136(2):1089–1143.

Maertens, M. and Swinnen, J. F. (2012). Gender and modern supply chains in developing countries. *The Journal of Development Studies*, 48(10):1412–1430.

Maredia, M. K., Byerlee, D., and Pee, P. (2000). Impacts of food crop improvement research: Evidence from sub-Saharan Africa. *Food Policy*, 25(5):531–559.

Mather, D., Boughton, D., and Jayne, T. (2013). Explaining smallholder maize marketing in southern and eastern Africa: The roles of market access, technology and household resource endowments. *Food Policy*, 43:248–266.

McArthur, J. W. and McCord, G. C. (2017). Fertilizing growth: Agricultural inputs and their effects in economic development. *Journal of Development Economics*, 127:133–152.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444.

Mengel, F., Sauermann, J., and Zölitz, U. (2018). Gender bias in teaching evaluations. *Journal of the European Economic Association*, 17(2):535–566.

Michelson, H., Fairbairn, A., Ellison, B., Maertens, A., and Manyong, V. (2021). Misperceived quality: Fertilizer in Tanzania. *Journal of Development Economics*, 148:102579.

Miller, J. and Chamberlin, M. (2000). Women are teachers, men are professors: A study of student perceptions. *Teaching Sociology*, 28(4):283–298.

Minten, B., Koru, B., and Stifel, D. (2013). The last mile(s) in modern input distribution: Pricing, profitability, and adoption. *Agricultural Economics*, 44(6):629–646.

Mitchell, K. M. W. and Martin, J. (2018). Gender bias in student evaluations. *PS: Political Science & Politics*, 51(3):648–652.

Mnimbo, T. S., Lyimo-Macha, J., Urassa, J. K., Mahoo, H. F., Tumbo, S. D., and Graef, F. (2017). Influence of gender on roles, choices of crop types and value chain upgrading strategies in semi-arid and sub-humid Tanzania. *Food Security*, 9(6):1173–1187.

Nsengimana, S., Tengeh, R. K., and Iwu, C. G. (2017). The sustainability of businesses in Kigali, Rwanda: An analysis of the barriers faced by women entrepreneurs. *Sustainability*, 9(8).

Odongo, W., Dora, M., Molnár, A., Ongeng, D., and Gellynck, X. (2016). Performance perceptions among food supply chain members. *British Food Journal*, 118(7):1783–1799.

Ola, O. and Menapace, L. (2020). Smallholders' perceptions and preferences for market attributes promoting sustained participation in modern agricultural value chains. *Food Policy*, 97:101962.

Patiar, A. and Mia, L. (2008). The effect of subordinates' gender on the difference between self-ratings, and superiors' ratings, of subordinates' performance in hotels. *International Journal of Hospitality Management*, 27(1):53–64.

Reimers, I. and Waldfogel, J. (2021). Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings. *American Economic Review*, 111(6):1944–71.

Riley, E. (2017). Increasing students' aspirations: the impact of Queen of Katwe on students' educational attainment. *CSAE Working Paper WPS/2017-13, Centre for the Study of African Economies, University of Oxford*.

Rosenkrantz, P., Vogel, S., Bee, H., Broverman, I., and Broverman, D. M. (1968). Sex-role stereotypes and self-concepts in college students. *Journal of Consulting and Clinical Psychology*, 32(3):287.

Sheahan, M. and Barrett, C. B. (2017). Ten striking facts about agricultural input use in sub-Saharan Africa. *Food Policy*, 67:12–25.

Sitko, N. J. and Jayne, T. (2014). Exploitative briefcase businessmen, parasites, and other myths and legends: Assembly traders and the performance of maize markets in eastern and southern Africa. *World Development*, 54:56–67.

Stifel, D. and Minten, B. (2008). Isolation and agricultural productivity. *Agricultural Economics*, 39(1):1–15.

Thornton III, G. C., Rupp, D. E., Gibbons, A. M., and Vanhove, A. J. (2019). Same-gender and same-race bias in assessment center ratings: A rating error approach to understanding subgroup differences. *International Journal of Selection and Assessment*, 27(1):54–71.

Van Campenhout, B., Lecoutere, E., and D'Exelle, B. (2015). Inter-temporal and spatial price dispersion patterns and the well-being of maize producers in southern Tanzania. *Journal of African Economies*, 24(2):230–253.

Winquist, L. A., Mohr, C. D., and Kenny, D. A. (1998). The female positivity effect in the perception of others. *Journal of Research in Personality*, 32(3):370–388.

Wu, A. (2020). Gender bias in rumors among professionals: An identity-based interpretation. *Review of Economics and Statistics*, 102(5):867–880.

Tables

Table 1: Summary statistics of variables related to farmers

	Summary statistics (Farmers)					
	Mean	Standard Deviation	Minimum	Maximum	First Quartile	Third Quartile
Overall rating (all actors)	3.6	0.77	1	5	3.2	4.2
Location rating (all actors)	3.88	1.17	1	5	3	5
Quality rating (all actors)	3.5	1.1	1	5	3	4
Price rating (all actors)	3.04	1.08	1	5	2	4
Reputation rating (all actors)	3.83	1.02	1	5	3	5
Overall rating (dealers)	3.59	0.74	1	5	3.2	4
Location rating (dealers)	3.65	1.27	1	5	3	5
Quality rating (dealers)	3.64	1.02	1	5	3	4
Price rating (dealers)	2.99	1.08	1	5	2	4
Reputation rating (dealers)	3.84	0.96	1	5	3	5
Overall rating (traders)	3.67	0.8	1	5	3.2	4.2
Location rating (traders)	4.09	1.02	1	5	4	5
Quality rating (traders)	3.54	1.01	1	5	3	4
Price rating (traders)	3.07	1.05	1	5	2	4
Reputation rating (traders)	3.84	1.04	1	5	3	5
Overall rating (processors)	3.54	0.75	1	5	3	4
Location rating (processors)	3.8	1.21	1	5	3	5
Quality rating (processors)	3.41	1.19	1	5	3	4
Price rating (processors)	3.02	1.11	1	5	2	4
Reputation rating (processors)	3.82	1.03	1	5	3	5

Table 2: Summary statistics of variables related to dealers, traders, and processors

	Self-ratings of value chain actors					
	Agro-input dealers					
	Mean	Standard Deviation	Minimum	Maximum	First Quartile	Third Quartile
Overall self-ratings	4.13	0.43	2.8	5	3.85	4.4
Location self-ratings	4.22	0.88	2	5	4	5
Quality self-ratings	4.58	0.61	3	5	4	5
Price self-ratings	4.05	0.82	2	5	3	5
Reputation self-ratings	4.4	0.86	1	5	4	5
	Assembly Traders					
Overall self-ratings	4.29	0.5	2.2	5	4	4.6
Location self-ratings	4.11	0.85	1	5	4	5
Quality self-ratings	4.33	0.77	1	5	4	5
Price self-ratings	3.91	0.83	1	5	3	5
Reputation self-ratings	4.45	0.77	2	5	4	5
	Processors					
Overall self-ratings	4.18	0.52	3	5	3.8	4.6
Location self-ratings	3.99	0.97	1	5	3	5
Quality self-ratings	4.16	0.84	2	5	4	5
Price self-ratings	3.84	0.95	1	5	3	5
Reputation self-ratings	4.5	0.69	2	5	4	5

Table 3: ICC coefficients for inter-rater reliability and intra-rater reliability

	Intraclass correlation coefficients	
	Inter-rater reliability (Agreement)	*Intra-rater reliability (Agreement)*
Overall	0.54	0.64
Location	0.47	0.62
Quality	0.15	0.31
Price	0.43	0.43
Reputation	0.24	0.68

Table 4: Average ratings (all dimensions) from farmers and average self-ratings (all dimensions) from dealers, traders, and processors, grouped by gender

	Average ratings (Mean)											
	Overall average (All dimensions)											
	All actors			Agro-input dealers			Assembly traders			Millers		
	Men	*Women*	*All*	*Men*	*Women*	*All*	*Men*	*Women*	*All*	*Men*	*Women*	*All*
Farmer is male	3.58	3.61	3.58	3.59	3.57	3.58	3.64	3.92	3.65	3.51	3.59	3.52
Farmer is female	3.62	3.63	3.62	3.6	3.64	3.61	3.68	4.09	3.69	3.58	3.44	3.57
Farmer is either male or female	3.59	3.61	3.6	3.59	3.59	3.59	3.66	4	3.66	3.54	3.53	3.54
Self-ratings	4.23	4.16	4.22	4.06	4.02	4.05	4.28	4.53	4.28	4.23	4.29	4.24
	Location											
Farmer is male	3.85	3.54	3.83	3.61	3.33	3.53	4.05	4.41	4.06	3.76	3.74	3.76
Farmer is female	3.98	3.82	3.97	3.93	3.77	3.89	4.13	4.4	4.13	3.87	3.67	3.86
Farmer is either male or female	3.91	3.64	3.88	3.72	3.46	3.65	4.08	4.41	4.09	3.81	3.71	3.8
Self-ratings	4.11	4.07	4.11	4.08	4.01	4.06	4.1	4.97	4.12	4.12	3.91	4.11
	Quality											
Farmer is male	3.49	3.7	3.51	3.71	3.65	3.69	3.53	3.82	3.54	3.37	3.77	3.39
Farmer is female	3.47	3.65	3.49	3.48	3.64	3.52	3.54	3.93	3.55	3.41	3.56	3.42
Farmer is either male or female	3.48	3.68	3.5	3.63	3.65	3.64	3.53	3.88	3.54	3.39	3.69	3.41
Self-ratings	4.24	4.68	4.28	4.48	4.62	4.52	4.3	4.88	4.31	4.12	4.71	4.16
	Price											
Farmer is male	3.01	2.95	3	2.96	2.92	2.95	3.05	3.24	3.05	2.99	2.93	2.98
Farmer is female	3.1	3	3.09	3.08	3.09	3.08	3.09	3.47	3.1	3.1	2.69	3.08
Farmer is either male or female	3.04	2.97	3.04	3	2.97	2.99	3.07	3.34	3.07	3.04	2.83	3.02
Self-ratings	3.9	4.06	3.92	3.82	4.05	3.88	3.93	3.94	3.93	3.91	4.14	3.92
	Reputation											
Farmer is male	3.82	3.93	3.83	3.82	3.96	3.86	3.81	4.06	3.82	3.83	3.84	3.83
Farmer is female	3.83	3.93	3.84	3.78	3.89	3.81	3.85	4.4	3.87	3.82	3.79	3.82
Farmer is either male or female	3.82	3.93	3.83	3.81	3.94	3.84	3.83	4.22	3.84	3.83	3.82	3.82
Self-ratings	4.48	4.34	4.47	4.53	4.4	4.49	4.38	4	4.37	4.55	4.33	4.54

Table 5: T-test results for differences between self-ratings and farmer ratings

	T-tests: Differences between self-ratings and farmer ratings > 0		
	Self-ratings	Farmer ratings	P-value
Overall	4.22	3.596	<0.001
Location	4.106	3.884	<0.001
Price	3.917	3.036	<0.001
Quality	4.282	3.499	<0.001
Reputation	4.467	3.833	<0.001

Note: The average self-ratings for each dimension are mentioned in the first column and the average farmer ratings for each dimension are mentioned in the second column. The p-value indicating the significance of each t-test is also presented. The alternative hypothesis is that the differences between the self-ratings and the farmer ratings are greater than 0.

Table 6: Regression results for the impact of farmer's and actor's gender on the ratings given to actors by farmers

	Dependent variable: Ratings from farmers				
	Overall	Location	Quality	Price	Reputation
	(1)	(2)	(3)	(4)	(5)
Constant	3.535	3.878	3.238	3.064	3.83
	(0.106)	(0.195)	(0.174)	(0.143)	(0.122)
Farmer is female	0.034	0.113^{**}	−0.034	0.07^{*}	−0.011
	(0.031)	(0.053)	(0.044)	(0.039)	(0.039)
Actor is female	0.039	−0.174	0.138	−0.034	0.122
	(0.072)	(0.133)	(0.104)	(0.088)	(0.098)
Farmer has finished	−0.003	0.01	−0.039	0.011	−0.003
primary education	(0.03)	(0.044)	(0.042)	(0.042)	(0.038)
Farmer's age	0	0.001	−0.001	0	−0.001
(in years)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
Farmer's distance to	−0.002	−0.002	-0.006^{*}	−0.004	0.001
tarmac road (in km)	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)
Farmer's distance to	-0.016^{*}	-0.028^{*}	−0.002	−0.013	0.005
murram road (in km)	(0.01)	(0.017)	(0.014)	(0.014)	(0.013)
Farmer is married	−0.064	−0.068	−0.035	−0.089	−0.085
	(0.043)	(0.073)	(0.06)	(0.067)	(0.059)
Actor's age	0.003	0.001	0.006^{**}	0.002	0.003
(in years)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)
Actor is married	-0.117^{**}	−0.134	−0.146	-0.136^{*}	−0.083
	(0.057)	(0.113)	(0.1)	(0.073)	(0.069)
Actor has finished	0.089^{**}	−0.022	0.295^{***}	0.106^{**}	0.057
primary education	(0.041)	(0.072)	(0.075)	(0.049)	(0.046)
Actor is a dealer	0.019	−0.124	0.123	−0.05	−0.03
	(0.048)	(0.105)	(0.078)	(0.062)	(0.054)
Actor is a trader	0.156^{***}	0.304^{***}	0.192^{***}	0.081	0.032
	(0.041)	(0.073)	(0.071)	(0.056)	(0.055)
Interaction: Farmer is female*	−0.014	0.156	0.001	−0.049	−0.021
Actor is female	(0.093)	(0.173)	(0.137)	(0.128)	(0.125)
Number of obs.	3588	3588	3588	3588	3588

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Note: Standard errors are clustered at the actor level (agro-input dealers, traders, and processors) and the farmer level.

Table 7: Regression results looking at the impact of actor gender on self-ratings

	Dependent variable: Self-ratings by dealers, traders, and millers				
	Overall	Location	Quality	Price	Reputation
	(1)	(2)	(3)	(4)	(5)
Constant	4.063	3.708	3.883	3.843	4.441
	(0.092)	(0.163)	(0.142)	(0.16)	(0.14)
Actor is female	0.003	−0.107	0.216	0.153	−0.105
	(0.086)	(0.153)	(0.133)	(0.15)	(0.131)
Actor's age	0	0.008^{**}	−0.001	−0.001	−0.002
(in years)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Actor is married	0.036	−0.154	0.21^{*}	−0.021	0.158
	(0.073)	(0.129)	(0.113)	(0.127)	(0.111)
Actor has finished	0.107^{**}	0.182^{**}	0.172^{**}	0.043	0.016
primary education	(0.044)	(0.078)	(0.068)	(0.076)	(0.067)
Actor is a dealer	−0.067	0.203	0.354^{***}	0.164	−0.07
	(0.071)	(0.126)	(0.11)	(0.124)	(0.108)
Actor is a trader	0.118^{**}	0.144^{*}	0.168^{**}	0.076	−0.079
	(0.048)	(0.085)	(0.074)	(0.083)	(0.073)
Number of obs.	592	592	592	592	592

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Note: Dependent variable is the self-rating given by the actors.

Table 8: Regression results for the impact of farmer and actor gender on the differences between actor self-ratings and farmer ratings

	Dependent variable: Differences between actor self-ratings and farmer ratings				
	Overall	Location	Quality	Price	Reputation
	(1)	(2)	(3)	(4)	(5)
Constant	0.546^{***}	−0.115	0.909^{***}	0.671^{**}	0.656^{**}
	(0.16)	(0.339)	(0.301)	(0.287)	(0.287)
Farmer is female	-0.094^{**}	-0.173^{**}	-0.135^{**}	−0.005	−0.053
	(0.04)	(0.076)	(0.058)	(0.052)	(0.052)
Actor is female	−0.069	0.085	0.284	0.194	-0.347^{**}
	(0.121)	(0.345)	(0.176)	(0.142)	(0.142)
Farmer has finished	0.005	−0.019	−0.059	0.094^{*}	−0.018
primary education	(0.037)	(0.06)	(0.049)	(0.052)	(0.052)
Farmer's age	0	0	−0.001	0.002	0
(in years)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Farmer's distance to	0.003	0.005	−0.003	0.006	−0.003
tarmac road (in km)	(0.003)	(0.007)	(0.005)	(0.005)	(0.005)
Farmer's distance to	0.011	0.043^{*}	0.01	−0.01	−0.014
murram road (in km)	(0.013)	(0.025)	(0.021)	(0.018)	(0.018)
Farmer is married	0.093^{*}	0.142^{*}	0.074	0.002	0.114^{*}
	(0.048)	(0.082)	(0.075)	(0.058)	(0.058)
Actor's age	0	0.007	0.004	−0.006	−0.001
(in years)	(0.003)	(0.007)	(0.005)	(0.005)	(0.005)
Actor is married	0.067	−0.149	0.085	0.251	0.071
	(0.113)	(0.257)	(0.204)	(0.19)	(0.19)
Actor has finished	0.025	0.262^{*}	-0.268^{**}	−0.106	0.018
primary education	(0.077)	(0.147)	(0.11)	(0.125)	(0.125)
Actor is a dealer	-0.245^{**}	−0.014	−0.031	0.141	0.009
	(0.097)	(0.268)	(0.17)	(0.123)	(0.123)
Actor is a trader	−0.096	-0.236^{*}	−0.074	−0.021	−0.199
	(0.078)	(0.141)	(0.13)	(0.13)	(0.13)
Interaction: Farmer is female*	0.147	−0.086	0.022	0.107	0.13
Actor is female	(0.112)	(0.255)	(0.151)	(0.133)	(0.133)
Number of obs.	3588	3588	3588	3588	3588

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Note: Standard errors are clustered at the actor (agro-input dealers, traders, and processors) and farmer level.