We would like to thank the reviewer for the valuable comments and for considering the potential and contribution of this study. Our response to the comments are below:

1. The paper is based on perceptions but there is no attempt to link it with observed conduct. Doing so might have important implications as it might imply that policies should work on these perceptions or on the behavior of these actors (As a farmer only interacts with a number of small traders, it might be hard for him to assess the whole universe and his perceptions might be wrong (or right)). Some of these factors are hard to measure, such as reputation. However, location and prices might be more straightforward. If possibilities exist, it might be good to test out to what extent these perceptions link with reality. With the dataset available, one could possibly do interesting additional assessments linking these observables with perceptions.

*Thank you for this excellent idea, which we have incorporated in the paper. We added Table 4 (page 15) to show the correlation between perceptions and a more objective measure of the underlying attribute that was rated. While, as the reviewer said, it may not be possible to find a viable proxy for each rating, we were able to construct objective measures for relative location of the actors, quality of the actor, and price competitiveness. We find significant correlations between the perceptions of the farmers (ratings) and the observed location and quality. However, no significant association between the price-based rating and the actual prices could be observed. Table 4 is discussed in the last paragraph of the section 4.3 “Reliability of ratings” (page 14):*

*“*The fact that we find reasonable inter-rater agreement seems to suggest that ratings are valid proxies for the attributes of the value chain actors being rated. In Table [4](#_bookmark74), we test validity further by correlating average ratings received by actors to observable characteristics of the actor. Some of these dimensions, like reputation, are hard to measure, but for others like location, quality and prices charged, we are able to construct proxies. In the first column of Table [4](#_bookmark74), we correlate the location based rating to a measure that attempts to capture the location of the actor (dealer or miller) relative to where the customers are, and find that actors that are less centrally located are scored lower on the location attribute. To test if quality ratings are associated to observable quality characteristics of the actors, we first compute an index that is based on various observables. For instance, for agro-input dealers, the index measures if various seed quality related attributes such as shelf life and moisture content were checked over the course of the previous year by official inspectors. For traders, the index includes whether the trader uses improved storage bags, as well as a number of services they provide to farmers. For millers, quality is proxied by looking at the structure where the mill is located in (type of roof, wall and floor). Using this quality index, we also find that there is a positive correlation between observed quality and the quality ratings actors get. The last column shows that there is no significant correlation between the price that value chain actors charge for their services and products, and the price competitiveness ratings*.“*

1. I am a bit unclear about the role of traders and processors in the value chain and the extent to which farmers are able to rate them. It is said that traders buy produce and link with processors. If that is the overall structure of the value chain and farmers are asked to rate processors, what problems does this create as they might not know this person. It would be good to analyze how pervasive this problem is in this dataset. Additionally, it might possibly be interesting to analyze what farmers might say about somebody that might not really know well and compare to people that do. If data are there, it might also be good to bring intensity of interactions in there.

*We agree that it was perhaps not entirely clear that we consider (also) small processors and that farmers use these small millers to process the maize for their own consumption. As such, farmers are likely to have a good idea about the millers in their vicinity. We have added some more text in the paper (section 2 “THE STRUCTURE OF THE MAIZE SUPPLY CHAIN”) on what the value chain typically looks like (Page 4).*

*The second point related to interactions between the farmer and the actor is an important one and we ran some additional analysis.*

*First, we checked if farmers that have interacted with a particular actor score significantly different from farmers that did not have first hand experience with the actor. We find that farmers that report interactions also score higher – which was to be expected as farmers would self-select into relationships with actors.*

*However, the correlation between ratings and interaction may also confound the relationships we are interested in. For instance, with respect to hypothesis 4, if gender of the actor is also correlated to the likelihood of interaction, then the estimate of the relationship between gender of the actor and the rating may be biased. Similarly, for hypothesis 2, if the gender of the rater is also correlated with the likelihood of interaction (e.g., men are more likely to interact), this may again bias the relationship between gender of the rater and rating (men rate higher but not because they are men, but because they interact).*

*To take this into account, we control for interaction between farmer and actor in the relevant hypotheses. For hypothesis 2, we find that indeed men interact more, and as a result, the positive relationship between the farmer being a woman and the rating becomes more outspoken (see table 7 on page 21 and 9 on page 23). For hypothesis 4, we do not find that there is a correlation between the gender of the actor and the likelihood of interaction, and so including interaction between the farmer and the actor has no effect on actor’s gender. We also find that previous results related to gender homophily effects were robust to the inclusion of interaction as a control variable (table 7 and 9).*

1. There is some literature that indicates that ordering of questions and anchoring bias might be an issue, i.e. if you have said “excellent” already twice, you might want to change the next answer. Some discussion or thoughts on this potential issue might also be useful.

*Thank you for pointing this out. We have added “*If the same individual needs to rate various actors on various attributes, the resulting ratings may suffer from some kind of anchoring bias if, for instance, a farmer that gives two consecutive positive ratings is more (or less) likely to give a third positive rating ([Furnham & Boo](#_bookmark32), [2011](#_bookmark32); [Tversky & Kahneman](#_bookmark67), [1974](#_bookmark67)). As the direction of this bias is unclear a priori and likely depends on where in the rating distribution the farmer starts (that is, if a farmer (the rater) starts with a five (one), the farmer may be likely to adjust downward (upward), making the direction of adaptive adjustment in the ratings unpredictable). Hence, it is also not clear how this feature in the data will affect our findings. Although anchoring bias can thus result in within-farmer and within-actor correlations as consecutive ratings may be correlated, the fact that we cluster standard errors at the level of the farmer and the actor (see below) is expected to diminish concerns related to heteroscedasticity resulting from this correlation.*” as footnote 3 on page 11.*