Study Protocol: Why do farmers sell immediately after harvest when prices are lowest?

Joachim De Weerdt*, Brian Dillon†, Emmanuel Hami*, Bjorn Van Campenhout‡

February 23, 2022

Motivation

It is often observed that smallholder farmers sell most—if not all—of their marketable surplus immediately after the harvest to itinerant traders at the farm gate. Selling immediately after the harvest is not optimal. Thin and poorly integrated markets mean that immediately post harvest, prices in excess supply areas drop. Later, during the lean season when some of the farmers run out of stock, prices have recovered, or even increase further since farmers start to buy back. This leads to the "sell low buy high" puzzle (Stephens and Barrett, 2011). In addition to high supply immediately post harvest, agricultural commodities are often not yet in optimal condition. For instance, in the case of maize, fresh grains are generally not dry enough, requiring further processing and leading to increased risk of rot by the trader. Often, this is used by buyers as a reason to further drive down the price paid to the farmer.

There are many possible reasons why farmers choose the sell early at low prices instead of waiting a few months until prices recover. Farmers may simply not have the space and infrastructure available to safely store large quantities of maize for extended periods of time (Omotilewa et al., 2018). Farmers may be in urgent need of cash after the lean season (Burke, Bergquist, and Miguel, 2018; Dillon, 2021). Price movements may be unpredictable and farmers may be too risk averse to engage into intertemporal arbitrage (Cardell and Michelson, 2020). It may be that traders only visit villages immediately after harvest, and farmers do not have the means to transport maize to markets themselves. Furthermore, issues related to social taxation may mean farmers convert maize

*Development Strategy and Governance Division, International Food Policy Research Institute, Lilongwe, Malawi

†Dyson School, Cornell University, Ithaca, United States

‡Development Strategy and Governance Division, International Food Policy Research Institute, Leuven, Belgium

to cash, which is easier to hide from friends and family. There may also be behavioural factors, such as present bias, anchoring of prices to past experience where negative experiences are more salient than positive experiences, etc.

In this study, we zoom in on two potential explanations why farmers seemingly sell at sub-optimal time. A first is related to budget neglect, whereby farmers underestimate expenses later in the season and as a result sell too much of their harvest too soon. To test this hypothesis, we implement a field experiment that aims to reduce budget neglect by drawing up expenditure plans up to the next harvest. A second potential explanation is that farmers consistently underestimate the expected return to waiting, reducing the incentive to store. We conjecture that farmers focus too much on the uncertainty related to the price in isolation, a cognitive bias known as "narrow bracketing", and that this makes farmers particularly loss averse. To overcome this bias, we develop a treatment where we point out that in addition to uncertainty related to the price, there are many other sources of uncertainty that play simultaneously.

Literature

Why do farmers sell low and buy high? One of the most obvious, is related to credit constraints. Using observational data, Stephens and Barrett (2011) find that to meet consumption needs later in the year, many farmers end up buying back grain from the market a few months after selling it, in effect using the maize market as a high-interest lender of last resort. Burke, Bergquist, and Miguel (2018) show that in a field experiment in Kenya, credit market imperfections limit farmers' abilities to move grain intertemporally. Providing timely access to credit allows farmers to buy at lower prices and sell at higher prices, increasing farm revenues and generating a return on investment of almost 30%. Dillon (2021) uses the fact that in Malawi, primary school began 3 months earlier in 2010 than in 2009, and notes that this prompted households with children to sell maize when prices are particularly low. To identify the impacts of liquidity during the lean season, Fink, Jack, and Masiye (2020) offered subsidized loans in randomly selected villages in rural Zambia and conclude that liquidity constraints contribute to inequality in rural economies. While credit constraints thus seems to be an important reason why farmers sell immediately post harvest at low prices, we feel this is not the entire story. If farmers need urgent cash, it would make more sense to only sell part of the harvest to cover most urgent expenses. However, farmers generally sell all maize immediately post harvest at low prices.

Risk averse farmers may also fail to delay sales if there is considerable uncertainty about the price in the future. A recent article argues that the "sell low buy high" puzzle is not a puzzle at all, as price movements are insufficient for farmers to engage in inter-temporal arbitrage (Cardell and Michelson, 2020). However, their analysis use prices obtained from market centers, which may be a poor proxy for the farm gate prices that farmers face: prices in main markets are generally much better integrated in the wider national, regional and even global economy, and so will be less prone to extreme spikes and slumps. While we agree that uncertainty about prices is indeed an important reason to sell immediately post harvest for loss averse farmers (and indeed loss aversion lies at the core of one of our research hypotheses), we do feel that this is not a sufficient explanation in the face of large recurrent seasonal price movements.

A third reason that is often heard in the field is that farmers have nowhere to store, so they just sell. This could be a lack of space, as the average smallholder often harvest 10-20 bags of 100kg of maize. But there are also risk related to pests and diseases affecting the stored maize. If storage is the main reason why farmers do not engage in intertemporal arbitrage more, then providing storage technology should delay sales. Omotilewa et al. (2018) indeed find that households that received PICS bags stored maize for a longer period, reported a substantial drop in storage losses. Again, we feel storage is indeed part of the reason, but it does not explain everything. For instance ACE in Malawi provides storage technology but still fails to fill its warehouses.

Another reason may be related to social taxation. If a farmer has a lot of maize stored in his house, this is visible for family and neighbours, and it will be very hard to deny if they come and ask for help. Therefore, farmers may choose to convert their harvest to money, which is easier to hide, even though this comes at a cost. Social taxation has been found important in a similar marketing decisions where household seem to forgo the benefits of buying in bulk (Dillon, De Weerdt, and O'Donoghue, 2020).

Model

Consider a very simple 2 period model where a farmer maximizes utility by choosing how much to consume in the first and second period (c_1, c_2). We further add an expenditure that is only relevant in period 2 (e_2), and that is affected by a parameter $0 \leqslant \gamma \leqslant 1$ that represents the probability that the agent remembers the expenditure. δ is the discount rate.

$$\underset{c_1,c_2,e_2}{max}\, u_1\left(c_1\right) + \delta\left(u_2\left(c_2\right) + \gamma v_2\left(e_2\right)\right) \tag{1}$$

$$st\; c_1 + c_2 + \gamma e_2 = \left(\lambda p_L + (1-\lambda) p_H\right) Q$$

On the right hand side of the budget constraint, we model beliefs about the future price, which can be low (P_L) or high (P_H). The parameter λ represents the subjective probability that the price will be low. The parameter λ is assume to increase due to narrow bracketing. The model will be used to compare a benchmark case, where there is no budget neglect $(\gamma = 1)$ and the future price is high $(\lambda = 0)$, against alternatives and formulate a series of predictions that can be taken to the data.

Hypotheses and Interventions

In a first hypothesis, we assume that farmers suffer from budget neglect ($\gamma < 1$ on the left side of the budget constraint), which may lead to an overoptimistic view of the future. In particular, farmers may neglect some future expenditures when deciding on how much to sell immediately after the harvest. For example, immediately after harvest, they may budget for fresh seed from the agro-input dealer and for fertilizer, but they may forget that they also need pesticides and insecticides.

This hypothesis is related to the planning fallacy, where individuals typically underestimate the time it takes to complete a task, despite extensive experience with failure to complete the same task in a similar time frame in the past (Buehler, Griffin, and Peetz, 2010). Part of it may also be related to optimism bias if farmers neglect or underestimate the risk that adverse effects will happen to them. For instance, farmers may not budget for pesticides or insecticides because they believe they will not be affected by pests or insects. Budget neglect is also found to be a main contributing factor to recurrent hungry seasons in Zambia.

In a second hypothesis, we borrow from the literature on why people fail to invest in high return ventures. There is an emerging behavioural economics literature that blames "narrow bracketing" as an important reason why loss averse individuals fail to invest in potentially high return activities. Narrow bracketing refers to an individual considering each choice or source of uncertainty they face in isolation, failing to integrate it with other choices and risk from other sources (Tversky and Kahneman, 1981). Narrow bracketing is implicitly assumed in a range of economic models and analyses. It has bite in the case of loss aversion, due to the importance of the sharp kink in the utility function at the reference point, which would effectively be smoothed out if individuals were considering many sources of uncertainty simultaneously (Kremer, Rao, and Schilbach, 2019).

The decision to sell in the future at an uncertain price can be seen as risky prospect. At the time of the harvest, the farmer evaluates this risky project to decide if they should sell or wait. However, the fact that the farmer considers this risky prospect in isolation (due to for instance cognitive capacity limitations) may lead them to refrain from taking the bet (that is wait), while they may take the bet if they concurrently consider the variety of other risky prospects they are faced with. For instance, the utility of the farmer will also depend on the realized price of all the goods in the consumption bundle. Especially when the different risky prospects are negatively correlated, which is likely to be the case since when the price of maize happens to be low during the lean season, the price of foodstuff in the consumption bundle will also be low, the consequences of narrow bracketing may be particularity severe (Read et al., 1999).

To test the first hypothesis related to budget neglect, we will design an intervention that takes the farmer through a detailed budgeting exercise. The budget exercise will involve three components. A first component uses recall to provide a first approximation of what will be necessary in the future. A second component consists of segmentation, which involves defining categories of expenditures for cognitive ease. Finally, we will look at a range of risks, which involve expenses that are not certain but may materialize. We try as much as possible to attach objective probabilities to these risks and also incorporate this in the budget.

To test the second hypothesis, a second intervention will be designed. In one arm, the control, we will present expected price distributions at different points in time that the farmer can expect when supplying maize. In a second intervention, we will not only provide information on the sales price distribution of maize, but also of retail prices commonly bought foodstuff. We will also make the negative correlation between maize sold and food bought salient by expressing revenue of maize sales in terms of purchasing power.

Experimental design and power calculations

We propose parallel design with one control group and two treatment arms. Kaur et al (personal communication) find that, in a similar budget neglect experiment, treated farmers enter the hungry season with 20 percent more maize (valued by current prices at 405 zambian kwacha instead of 335 zambian kwacha in the control group). If we assume that standard deviation is about 592 (1.6 times the mean of treatment and control means – the 1.6 is derived from maize production data in Uganda), we get a sample size of 1123 in each sample. If we run a factorial design and want to power all treatment cell for similar effect sizes, we will need about 3369 observations. To account for attrition, we will increase this to 3500. Note for designs with a common control group, maximize power, is attained when allocating approximately 42% of the sample to the control units, and then equally (29% each) to the two treatments. In our case, this means about 1470 farmers in the control group and 1015 in each treatment arm. We will implement this in 100 villages, such that in each village we will have 15 farmers in the control group, and 10 farmers in each treatment group.

The intervention will be implemented at the individual level, which raises concerns about potential spillover effects. We expect that farmers that are allocated to the control group may learn from treated farmers. This will thus affect outcomes of control farmers in the same direction as outcomes for treated farmers, making it harder to detect an effect if there is one. To get a sense of the importance of these spillover effects, we stratify on village size. Using census data, we subdivide villages in the study area in small and large villages based on the median number of inhabitants and then make sure we implement the study in 50 small and 50 large villages.

Burke, Bergquist, and Miguel (2018) document significant effects of a credit intervention on seasonal price fluctuations in local grain market. To study these kind of general equilibrium effects of our intervention, we will also record prices in at regular intervals in the villages we work in. These prices will then be compared with prices collected in similar villages nearby where the intervention is not implemented. We will collect prices in 50 additional villages.

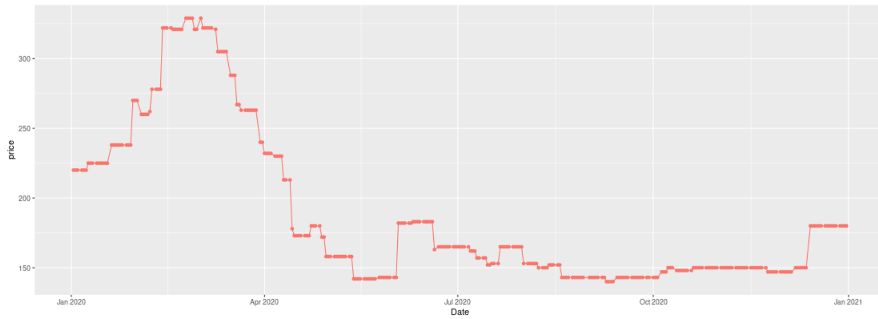

Figure 1: Price of maize in Rumphi

Context and study area

The study focuses on the Central and Northern Region of Malawi (Kasungu, Mzimba, Ntchisi, Rumphi, Dowa and Mchinji). These areas are characterized by rained agriculture with a single season. The resulting seasonal price movements is illustrated in Figure 1 that shows maize price in kwatcha per Kg in Rumpi over 2020. Planting of maize starts in December, and maize becomes increasingly scarce during the growing season. Harvesting starts around April 2020, which takes the pressure off the prices when farmers start consuming from their own maize. However, farm gate sales are still low as traders wait for maize to dry. This results in a relatively long period of low prices all the way to the start of the planting season towards the end to the year. The aim of the study is to encourage farmers to wait just a few months longer before they sell.

Figure 2 shows that most sales happen only around August. So farmers do seem to hold on to their maize for reasonably long periods (suggesting some of the other explanations like lack of storage space or social taxation are less likely).

Taken together, the figures suggest that the best time for the interventions would be around April or May, immediately before farmer start to sell.

The study population consists of semi-subsistence small scale maize farmers in the study districts defined above. As mentioned in Section , we plan to recruit a fixed number of households in each of a total of 100 villages (stratified by size—50 large and 50 small villages). Using census data, within each village, we will randomly sample 35 maize farming households to participate in the study using a random number generator of a computer. We work with the assumption that all households in the area produce maize and produce on a small scale, so all households meet the inclusion condition. In a second stage, these 35 households will be allocated to the three treatment groups (control, treatment 1 and treatment 2). This will also be done by a random number of a computer. These households will then be visited and asked to participate. We will provide study participants with a small token of gratitude (a bar of soap) regardless of whether the household accepts to participate or not.

Figure 2: Quantities of maize bought and sold

Data collection and endpoints

We will not organize a dedicated baseline survey, but rather ask a limited number of questions immediately prior to the interventions. This information can then be used to demonstrated balance, to control for baseline outcomes for the primary outcome variables in and ANCOVA regression, and to explore heterogeneous treatment effects.

Instead of relying on a baseline and endline, we will evaluate the interventions through multiple rounds of data collection, often using phone interviews. There are different reasons for this. First, when measuring noisy and relatively less autocorrelated outcomes such as amounts of maize to sell or household expenditure, one can increase power by taking multiple measurements at relatively short intervals to average out noise (McKenzie, 2012). Furthermore, it will allow us to assess the effect of the interventions at multiple points in time instead of just at endline.

Limited baseline information will be collected in April/May 2022 during the intervention phase. Intermediate data will be collected in July 2022, September 2022, November 2022, and January 2023. A slightly more elaborate in-person endline survey will be organized in March 2023.

Primary outcomes in this study include amounts of maize stocks and and maize sold at different points in time. As the main aim of waiting is intertemporal arbitrage, we will also compare prices obtained between treatment and control farmers. We will equally look at purchases. Further down the impact pathway, we compare welfare, both subjective and through consumption expenditure, between treatment and control households.

To investigate impact pathways, we will also include a range of questions related to expenditure, and how easy it was for farmers. For instance, did treated households have less issues in meeting expenditures for eg. fertilizer or improved seed for the next season? We will also elicit a full set of future prices to look at impact pathways for the second hypothesis.

References

Stephens, Emma C. and Christopher B. Barrett (2011). "Incomplete Credit Markets and Commodity Marketing Behaviour". In: *Journal of Agricultural Economics* 62.1, pp. 1–24. DOI: https://doi.org/10.1111/j.1477-9552.2010.00274.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1477-9552.2010.00274.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1477-9552.2010.00274.x.

Dillon, Brian, Joachim De Weerdt, and Ted O'Donoghue (Jan. 2020). "Paying More for Less: Why Don't Households in Tanzania Take Advantage of Bulk Discounts?" In: *The World Bank Economic Review* 35.1, pp. 148–179. ISSN: 0258-6770. DOI: 10.1093/wber/lhz020. eprint: https://academic.oup.com/wber/article-pdf/35/1/148/36158732/lhz020.pdf. URL: https://doi.org/10.1093/wber/lhz020.

Omotilewa, Oluwatoba J. et al. (2018). "Does improved storage technology promote modern input use and food security? Evidence from a randomized trial in Uganda". In: *Journal of Development Economics* 135, pp. 176–198. ISSN: 0304-3878. DOI: https://doi.org/10.1016/j.jdeveco.2018.07.006. URL: https://www.sciencedirect.com/science/article/pii/S030438781830926X.

Fink, Gunther, B. Kelsey Jack, and Felix Masiye (2020). "Seasonal Liquidity, Rural Labor Markets, and Agricultural Production". In: *American Economic Review* 110.11, pp. 3351–92. DOI: 10.1257/aer.20180607. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20180607.

Dillon, Brian (2021). "Selling Crops Early to Pay for School: A Large-Scale Natural Experiment in Malawi". In: *Journal of Human Resources* 56.4, pp. 1296–1325. DOI: 10.3368/jhr.56.4.0617-8899R1. eprint: http://jhr.uwpress.org/content/56/4/1296.full.pdf+html. URL: http://jhr.uwpress.org/content/56/4/1296.abstract.

McKenzie, David (2012). "Beyond baseline and follow-up: The case for more T in experiments". In: *Journal of Development Economics* 99.2, pp. 210–221. ISSN: 0304-3878. DOI: https://doi.org/10.1016/j.jdeveco.2012.01.002. URL: https://www.sciencedirect.com/science/article/pii/S030438781200003X.

Read, Daniel et al. (1999). "Choice bracketing". In: *Elicitation of preferences*. Springer, pp. 171–202.

Burke, Marshall, Lauren Falcao Bergquist, and Edward Miguel (Dec. 2018). "Sell Low and Buy High: Arbitrage and Local Price Effects in Kenyan Markets*". In: *The Quarterly Journal of Economics* 134.2, pp. 785–842. ISSN: 0033-

5533. DOI: 10.1093/qje/qjy034. eprint: https://academic.oup.com/qje/article-pdf/134/2/785/28289352/qjy034.pdf. URL: https://doi.org/10.1093/qje/qjy034.

Cardell, Lila and Hope Michelson (July 2020). *"Sell Low, Buy High?" - A New Explanation for a Persistent Puzzle*. 2020 Annual Meeting, July 26-28, Kansas City, Missouri 304448. Agricultural and Applied Economics Association. DOI: 10.22004/ag.econ.304448. URL: https://ideas.repec.org/p/ags/aaea20/304448.html.

Tversky, Amos and Daniel Kahneman (1981). "The Framing of Decisions and the Psychology of Choice". In: *Science* 211.4481, pp. 453–458. DOI: 10.1126/science.7455683. eprint: https://www.science.org/doi/pdf/10.1126/science.7455683. URL: https://www.science.org/doi/abs/10.1126/science.7455683.

Buehler, Roger, Dale Griffin, and Johanna Peetz (2010). "The planning fallacy: Cognitive, motivational, and social origins". In: *Advances in experimental social psychology*. Vol. 43. Elsevier, pp. 1–62.

Kremer, Michael, Gautam Rao, and Frank Schilbach (2019). "Behavioral development economics". In: *Handbook of Behavioral Economics: Applications and Foundations 1*. Vol. 2. Elsevier, pp. 345–458.