Why do farmers sell immediately after harvest when prices are lowest? A mock report

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

August 19, 2022

Abstract

It is often observed that smallholder farmers sell most—if not all— of their marketable surplus or cash crops immediately after the harvest during which seasonal price movement reach their lowest point. Most explanations for this apparently sub-optimal behavior focus on economic or infrastructural issue, such as credit constraints or lack of storage facilities. In this study, we zoom in on two potential behavioural explanations. One potential explanation is situated at the household expenditure side, and assumes that households face challenges in accurately predicting future expenditures. Such budget neglect leads farmer to sell more early on and save too little for later in the year. A second potential explanation is situated at the household income side. Here the assumption is that farmers face cognitive challenges in making inter-temporal cost benefit calculations and fail to commit to certain thresholds. This document, which supplements a preregistered pre-analysis plan, has been prepared after baseline data was collected and runs the analysis that on simulated data.

1 Introduction

It is often observed that smallholder farmers sell most of their marketable surplus or cash crops immediately after the harvest to itinerant traders at the farm gate. Selling immediately after the harvest is not optimal: Thin and poorly integrated markets often mean that immediately post harvest prices in excess supply areas drop to their seasonal low. In addition to high supply immediately post harvest,

*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

§Development Strategy and Governance Division, International Food Policy Research Institute, Kampala, Uganda

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 an additional reason to further drive down the price paid to the farmer.

Over time, prices gradually recover, reaching their seasonal high just before the next harvest. At this time, many farmers have run out of stock, and need to turn to the market to buy back maize at prices that are often a multiple of what they received, a phenomenon known as the "sell low buy high" puzzle (Stephens and Barrett, 2011; Burke, Bergquist, and Miguel, 2018). Van Campenhout, Lecoutere, and D'Exelle (2015) further argue how farmers also loose out twice, as transport costs are passed on by traders to farmers when traders buy commodities from farmers in rural areas to aggregate in consumer centers, and farmers need to incur transaction costs when they buy back (often from the same traders in towns or large warehouses in trading centers).

The literature suggest 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). They 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 to cash, which is easier to hide from friends and family (Jakiela and Ozier, 2015).

Most of the explanations above focus on neoclassical constraints to farmers' exploiting inter-temporal arbitrate. In this study, we zoom in on two potential behavioral explanations why farmers seemingly sell at sub-optimal time. One potential explanation is situated at the household expenditure side, and assumes that households face challenges in accurately predicting future expenditures. Such budget neglect leads farmer to sell more early on and save too little for later in the year. A second potential explanation is situates at the household income side. Here the assumption is that farmers face cognitive challenges in making inter-temporal cost benefit calculations (Drexler, Fischer, and Schoar, 2014) and fail to commit to certain thresholds (Ashraf, Karlan, and Yin, 2006; Duflo, Kremer, and Robinson, 2011).

The above hypotheses are tested with a field experiment among smallholder farmers in Malawi. These farmers produce (a mix of) maize, groundnuts and soybean for which at least part is destined for the market.

This document serves as a mock report. Like pre-analysis plans, mock reports are important tools against publication bias. It goes one step further than pre-analysis plans by making specifications and variables selection explicit through preregistering the actual code streams that will be run on the midline and endline data. To do so, this document uses the knitr engine to create a dynamic document that aims to be as close as possible to the eventual document that will be used to report the results of the experiment.[1] For now, most table will be filled with results based on simulated endline data. The ideas is that when mid and endline data is collected, it is simply a matter of connecting that data and re As such, it will provide a useful reference in evaluating the final results of the study (Humphreys, Sanchez de la Sierra, and van der Windt, 2013).

2 Related Research

Why do farmers sell low and buy high? One of the most obvious neo-classical explanations 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 by Cardell and Michelson (2020) argues that the "sell low buy high" puzzle is not a puzzle at all. Using 20 years of data from 787 markets in 26 countries, they argue that in many cases the price increase seems insufficient and too uncertain 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

[1] Ideally, the code that performs the analysis would be embedded in the latex source code. However, some statistical procedures (eg those that rely on simulation) may take a significant amount of time to run, making it impractical to rerun it each time the latex document is compiled. As such, we run the analysis in a separate script that writes results to matrices on disk. The knitr engine of the latex document then loads these matrices at startup and extracts data from these matrices to fill the tables. As all coded (R-scripts and latex source code) is under revision control and publicly available through git-hub transparency is preserved.

better integrated in the wider national, regional and even global economy, and so will be less prone to extreme spikes and slumps.

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, 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, this explanation is at odds with the fact that Agricultural Commodities Exchange (ACE) in Malawi 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 neighbors, and it will be very hard to deny if neighbors come and ask for help during the hunger season. 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).

3 Behavioural constraints to intertemporal arbitrage: Hypotheses and Interventions

As can be seen in the previous section, most research trying to explain the sell low buy high puzzle are neoclassical in nature. In this study, we want to test two potential explanations that are rooted in behavioral sciences. The first potential behavioral explanation is situated at the household expenditure side, and assumes that households face challenges in accurately predicting future expenditures. In particular, we assume that farmers systematically underestimate how much money they need in the future and as a result sell too much 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. Furthermore, farmers may underestimate the likelihood of, or simply forget to account for, unexpected events such illness within the family.

This hypothesis touches on cognitive limits of the household at the expenditure side. It is also 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 (Sharot, 2011). 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.

To test the first hypothesis, we designed an intervention that takes the farmer through a detailed budgeting exercise. In particular, the main decision maker within the household was provided with a template that needed to be filled in as detailed as possible. Expenditure categories included education expenditures (school fees, uniforms,...), agricultural investment expenditures (seed, fertilizer,...), investment expenditure in non-agricultural businesses (retail shop inventories,...), health and medical expenses (medicines, preventive doctor visit,...), household recurrent expenditure (food, utilities,...), household equipment and maintenance (furniture, renovation,...), and other expenditures (loan repayment, ceremonies,...). For each of these expenditure item lines, farmers were then asked to provide an estimate of the total cost for each month between May 2022 and April 2023 and write it down in the appropriate cell of the expenditure matrix. Farmers were also encouraged to provide their top 3 unexpected expenditures likely to occur between May 2022 and April 2023.

This second hypothesis is also related to cognitive limitations when planning, but this time at the income side of the farm household. Farmers may have difficulties in making the intertemporal cost-benefit calculations necessary to determine the optimal reservation price and/or storage period. They often lack precise information about the fixed and variable costs involved, about the level and variability of the future stream of income from sales, or about the time frame of both cost and income (Van Campenhout, 2021). The fact that farmers are faced with uncertain prices and uncertain expenditures often means they abandon plans and engage in impulsive or distress sales.

To test the second hypothesis, we develop, together with the farmer, a detailed sales plan for the year. Again using a template, we start by asking the farmer about the expected marketable surplus for maize, groundnuts and soybean. We then ask for each month between May 2022 and April 2023 how much the farmer is planning to sell from each of the crop and what the minimum expected price for this crop is.

4 Fieldwork

Farmers were visited between 21nd of may 2022 and 9th of april 2022. We visited a total of 3534 farmers that were located in

5 Baseline balance

We follow our pre-analysis plan states and demonstrate baseline balance by constructing a standard balance table consisting of the following variables household/demographic characteristics (inspired by balance tables in Duflo, Kremer, and Robinson (2011); Karlan et al. (2014)): household head is female (1=yes), household size (number of people), age of household head (years), number of years of education of the household head (years), material of roof (corrugated iron = 1), number of rooms in the house, cultivated acreage (maize+groundnuts+soybean), hired in agricultural labour (1=yes), distance to nearest all weather road (km), distance to nearest market (km). We also report t-tests comparing treatment and control (unadjusted for multiple hypothesis testing) as well as a joint F-test from a regression of the treatment assignment on all variables in the balance table. Results are summarized in Table 1.

The pre-analysis plan also states that we assess balance on a range of characteristics that we will use to investigate heterogeneous treatment effects. The variables mentioned there are access to credit, access to storage facility, membership of (marketing related) cooperative, livestock asset ownership, whether the household already makes a budget. Results for these variables are in Table 2.

6 Primary outcomes

A first primary outcome we consider is stocks of the three commodities at the time of the midline/endline. We look at absolute values in kg, as well as stock as a share of quantity harvested. We can not create a variable for stock at baseline to include as a control in the ANCOVA specification, as we only asked for production and sales transaction during baseline data collection. As such, we provide simple treatment control differences. For farmers that did not produce the crop, stocks are coded as zero. A second primary outcome is closely related to the first one and looks at stocks as a share of total production. Also here, we code farmers that did not produce the crop as zero.

A third primary outcome tests if the household sold at least once before a particular date. The binary variable that indicates if the household sold at baseline is constructed by checking if the household sold before first of October 2021.

The third primary outcome is quantity sold. While we may not find that the treatment reduced the fact that farmers sell early, we still may find that farmers sell less of the crop immediately after harvest.

Finally, we calculate the average price received for all transactions that the household made for a particular crop before October 2022. We control for the average price received for transactions over the same period in the 2021 season.

Note: We have a selection problem for some of the variables included in Table 3 are only recorded for a subset of the sample, and farmers may have entered or left this subset as a consequence of our treatment.

We use simulation to account for multiple comparisons.Simulation methods provide a flexible and intuitive way to think about multiple hypothesis testing. It accommodates the extent to which the multiple comparisons are correlated with one another and allows us to integrate design specific elements such as the blocking and multiple arms. This leads to a study-specific correction that will generally be more powerful than other methods to control the FWER. In particular, to determine target p-value cutoffs, we use the family-wise sharp null

Table 1: Balance table

	mean ctrl	T1	T2	nobs
Household head is female	0.219	-0.022	-0.011	3534
	(0.413)	(0.016)	(0.016)	
Household size (number of people)	5.011	-0.017	0.19*	3534
	(2.04)	(0.083)	(0.083)	
Age of household head (year)	43.138	-0.018	0.977	3414
	(14.885)	(0.608)	(0.61)	
Years of education of household head	6.237	0.248+	0.107	3428
	(3.457)	(0.14)	(0.14)	
Roof of main building is corrugated iron	0.37	0.029	0.038*	3534
	(0.483)	(0.019)	(0.019)	
Number of rooms in house	3.174	0.042	0.058	3534
	(1.17)	(0.046)	(0.046)	
Area of cultivated land (acres)	2.452	0.06	0.204**	3489
	(1.736)	(0.071)	(0.071)	
Hired labour for maize, soybean or gnut productions?	0.408	0.015	0.059**	3528
	(0.492)	(0.02)	(0.02)	
Distance to nearest all weather road (km)	1.415	0.348	-0.092	3350
	(4.585)	(0.257)	(0.258)	
Distance to nearest market (km)	4.342	0.584	-0.059	3251
	(8.407)	(0.377)	(0.377)	
F-test C/T1 (p-value)	1.064	(0.386)		
F-test C/T2 (p-value)	1.414	(0.168)		
Likelihood Ratio Test (p-value)	29.229	(0.083)		

Note: First column reports control group means (and standard deviations below); **, * and + denote significance at the 1, 5 and 10 percent levels. F-test test for joint significance in a regression with treatment status on the left hand side (T1/C or T2/C). Likelihood ratio test is derived from a multinomial model where the left hand side has three levels (T1,T2,C). All models include village fixed effects.

Table 2: Balance table for conditioning variables

	mean ctrl	T1	T2	nobs
Household has access to credit	0.384	0.002	-0.008	3534
	(0.487)	(0.019)	(0.019)	
Household has access to storage	0.592	0.004	0.02	3534
	(0.492)	(0.019)	(0.019)	
Cooperative member	0.137	-0.012	-0.001	3534
	(0.344)	(0.014)	(0.014)	
Has livestock	0.484	0.012	0.041*	3532
	(0.5)	(0.02)	(0.02)	
Makes a budget	0.694	0.008	0.012	3534
	(0.461)	(0.018)	(0.018)	
F-test C/T1 (p-value)	0.284	(0.922)		
F-test C/T2 (p-value)	1.176	(0.318)		
Likelihood Ratio Test (p-value)	7.452	(0.682)		

Note: First column reports control group means (and standard deviations below); **, * and + denote significance at the 1, 5 and 10 percent levels. F-test test for joint significance in a regression with treatment status on the left hand side (T1/C or T2/C). Likelihood ratio test is derived from a multinomial model where the left hand side has three levels (T1,T2,C). All models include village fixed effects.

of no effect for any unit on any dependent variable and for any hypothesis. More information, including R code on which our tests are based can be found here.

Multiple comparison arise from the number of treatment arms, the different primary outcomes and the different crops included in the analysis. Equation XXX tests 2 hypotheses (we will not directly test T1 against T2), Table 3 considers six outcomes, and look at 3 crops. This means a total of 36 hypothesis tests. Using a simple Bonferroni correction, this would mean that the 5 percent cut-off for the p-value would reduce to 0.001 and the 10 percent significance cutoff would reduce to 0.003. These values are only slightly higher on simulated data, but it is expected that with real data, correlation between outcomes is higher and difference between Bonferroni and the family wise null becomes larger.

7 Heterogeneity

7.1 Access to credit

7.2 Access to storage

References

Ashraf, N., D. Karlan, and W. Yin. 2006. "Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines*." *The Quar-*

Table 3: Primary outcomes

	mean ctrl	T1	T2	nobs
		maize		
Stocks (abs)	346.138	-44.591**	-119.838**	3398
	(408.135)	(14.315)	(14.354)	
Stocks (%)	0.437	-0.1**	-0.173**	3483
	(0.312)	(0.011)	(0.011)	
Sold (yes/no)	0.565	-0.07**	-0.021	3483
	(0.496)	(0.02)	(0.02)	
Quantity sold (abs)	413.212	-31.332	-6.579	3483
	(437.218)	(17.585)	(17.659)	
Quantity sold (%)	0.141	0.002	0.006	3483
	(0.258)	(0.01)	(0.01)	
Price sold (%)	198.495	4.071	3.526	946
	(63.04)	(4.295)	(4.107)	
		groundnuts		
Stocks (abs)	21.091	-1.717	-4.419**	3396
	(41.653)	(1.428)	(1.432)	
Stocks (%)	0.216	-0.027**	-0.073**	3483
	(0.312)	(0.01)	(0.01)	
Sold (yes/no)	0.694	-0.379**	-0.182**	3483
	(0.461)	(0.019)	(0.019)	
Quantity sold (abs)	162.851	-10.286	-1.521	3483
	(191.537)	(7.749)	(7.783)	
Quantity sold (%)	0.034	0.005	0.014*	3483
	(0.142)	(0.006)	(0.006)	
Price sold (%)	6296.058	132.136	-136.711	713
	(2026.94)	(176.049)	(172.227)	
		soybean		
Stocks (abs)	64.632	-5.183	-16.107**	3399
	(99.628)	(3.412)	(3.413)	
Stocks (%)	0.299	-0.051**	-0.107**	3483
	(0.332)	(0.011)	(0.011)	
Sold (yes/no)	0.937	-0.196**	-0.087**	3483
	(0.243)	(0.014)	(0.014)	
Quantity sold (abs)	421.622	-26.035	0.425	3483
	(488.226)	(19.809)	(19.889)	
Quantity sold (%)	0.043	0	0.02**	3483
	(0.162)	(0.007)	(0.007)	
Price sold (%)	549.566	11.543	12.932	1704
	(154.44)	(8.198)	(8.253)	

Note: First column reports control group means (and standard deviations below); **, * and + denote significance at the 1, 5 and 10 percent levels. All models include village fixed effects and controls for baseline characteristics that show imbalance at the 5 percent level or less in Table XXX.

terly Journal of Economics 121 (2): 635–672.

Buehler, R., D. Griffin, and J. Peetz. 2010. "The planning fallacy: Cognitive, motivational, and social origins." In "Advances in experimental social psychology," vol. 43, 1–62. Elsevier.

Burke, M., L. F. Bergquist, and E. Miguel. 2018. "Sell Low and Buy High: Arbitrage and Local Price Effects in Kenyan Markets*." *The Quarterly Journal of Economics* 134 (2): 785–842.

Cardell, L. and H. Michelson. 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.

Dillon, B. 2021. "Selling Crops Early to Pay for School: A Large-Scale Natural Experiment in Malawi." *Journal of Human Resources* 56 (4): 1296–1325.

Dillon, B., J. De Weerdt, and T. O'Donoghue. 2020. "Paying More for Less: Why Don't Households in Tanzania Take Advantage of Bulk Discounts?" *The World Bank Economic Review* 35 (1): 148–179.

Drexler, A., G. Fischer, and A. Schoar. 2014. "Keeping It Simple: Financial Literacy and Rules of Thumb." *American Economic Journal: Applied Economics* 6 (2): 1–31.

Duflo, E., M. Kremer, and J. Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *American Economic Review* 101 (6): 2350–90.

Fink, G., B. K. Jack, and F. Masiye. 2020. "Seasonal Liquidity, Rural Labor Markets, and Agricultural Production." *American Economic Review* 110 (11): 3351–92.

Humphreys, M., R. Sanchez de la Sierra, and P. van der Windt. 2013. "Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration." *Political Analysis* 21 (1): 1–20.

Jakiela, P. and O. Ozier. 2015. "Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies." *The Review of Economic Studies* 83 (1): 231–268.

Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry. 2014. " Agricultural Decisions after Relaxing Credit and Risk Constraints *." *The Quarterly Journal of Economics* 129 (2): 597–652.

Omotilewa, O. J., J. Ricker-Gilbert, J. H. Ainembabazi, and G. E. Shively. 2018. "Does improved storage technology promote modern input use and food security? Evidence from a randomized trial in Uganda." *Journal of Development Economics* 135: 176–198.

Sharot, T. 2011. "The optimism bias." *Current biology* 21 (23): R941–R945.

Stephens, E. C. and C. B. Barrett. 2011. "Incomplete Credit Markets and Commodity Marketing Behaviour." *Journal of Agricultural Economics* 62 (1): 1–24.

Van Campenhout, B. 2021. "The role of information in agricultural technology adoption: Experimental evidence from rice farmers in Uganda." *Economic Development and Cultural Change* 69 (3): 1239–1272.

Van Campenhout, B., E. Lecoutere, and B. D'Exelle. 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.