NBER WORKING PAPER SERIES

SELL LOW AND BUY HIGH:
ARBITRAGE AND LOCAL PRICE EFFECTS IN KENYAN MARKETS

Marshall Burke
Lauren Falcao Bergquist
Edward Miguel

Sell Low and Buy High: Arbitrage and Local Price Effects in Kenyan Markets
Marshall Burke, Lauren Falcao Bergquist, and Edward Miguel
NBER Working Paper No. 24476
April 2018
JEL No. D21,D51,G21,O13,O16,Q12

ABSTRACT

Large and regular seasonal price fluctuations in local grain markets appear to offer African farmers substantial inter-temporal arbitrage opportunities, but these opportunities remain largely unexploited: small-scale farmers are commonly observed to "sell low and buy high" rather than the reverse. In a field experiment in Kenya, we show that credit market imperfections limit farmers' abilities to move grain inter-temporally. 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 28%. To understand general equilibrium effects of these changes in behavior, we vary the density of loan offers across locations. We document significant effects of the credit intervention on seasonal price fluctuations in local grain markets, and show that these GE effects shape individual level profitability estimates. In contrast to existing experimental work, the results indicate a setting in which microcredit can improve firm profitability, and suggest that GE effects can substantially shape microcredit's effectiveness. In particular, failure to consider these GE effects could lead to underestimates of the social welfare benefits of microcredit interventions.

Marshall Burke
Department of Earth System Science
Stanford University
Stanford, CA 94305
and NBER
mburke@stanford.edu

Lauren Falcao Bergquist
Becker Friedman Institute
1126 E. 59th Street
SHFE 214
Chicago, IL 60637
lfbergquist@uchicago.edu

Edward Miguel
Department of Economics
University of California, Berkeley
530 Evans Hall #3880
Berkeley, CA 94720
and NBER
emiguel@econ.berkeley.edu

1 Introduction

African agricultural markets are thin and imperfectly integrated, resulting in substantial variation in staple commodity availability and prices (Fafchamps, 1992; Barrett and Dorosh, 1996; Minten and Kyle, 1999). Particularly pronounced are price fluctuations seen over time, with grain prices in major markets regularly rising by 25-40% between the harvest and lean seasons, and often more than 50% in more isolated markets. Underpinning these fluctuations is seemingly puzzling behavior at the farmer level: despite having access to relatively cheap storage technologies, farmers tend to sell their crops immediately after harvest and, in many areas, then later return to the market as consumers several months later in the lean season once prices have risen.

In this paper, we explore the role of financial market imperfections in contributing to farmers' apparent inability to exploit this arbitrage opportunity. Lack of access to credit markets has long been considered to play a central role in underdevelopment, but much of the literature has focused on the implications for firm growth and occupational choice (Banerjee and Newman, 1993; Galor and Zeira, 1993; Banerjee and Duflo, 2010). Here, we explore a novel channel by which poor financial market access may limit development: by restricting individuals' ability to engage in arbitrage and, at a macro level, by subsequently contributing to the large seasonal price fluctuations that characterize these markets.

Rural households' difficulty in using storage to move grain from times of low prices to times of high prices appears at least in part due to limited borrowing opportunities: lacking access to credit or savings, farmers report selling their grain at low post-harvest prices to meet urgent cash needs (e.g., to pay school fees). To meet consumption needs later in the year, many then 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 (Stephens and Barrett, 2011).

Working with a local agricultural microfinance NGO, we offer randomly selected smallholder maize farmers a loan at harvest, and study whether access to this loan improves their ability to use storage to arbitrage local price fluctuations relative to a control group. We find that farmers offered this harvest-time loan sell significantly less and purchase significantly more maize in the period immediately following harvest, and this pattern reverses during the period of higher prices 6-

9 months later. This change in their marketing behavior results in a statistically significant increase in revenues (net of loan interest) of 545 Ksh (or roughly 6 USD), suggesting that the loan produces a return on investment of 28% over a roughly nine month period. We replicate the experiment in two back-to-back years to test the robustness of these results and find remarkably similar results on primary outcomes in both years.

We also run a long-run follow-up survey with respondents 1-2 years after harvest-time credit intervention had been discontinued by the NGO, to test whether farmers are able to use the additional revenues earned from this loan product to "save their way out" of credit constraints in future years. We find no evidence of sustained shifts in the timing of farm sales in subsequent seasons, nor do we see long-run effects on sales or revenues in subsequent years.[1]

To explore whether this shift in sales behavior by individuals farmers has an effect on market-level prices, we experimentally vary the density of treated farmers across locations and tracked market prices at 52 local market points. We find that increase grain storage at the market level (induced by the credit intervention) led to significantly higher prices immediately after harvest and to lower (albeit not significantly so) prices during the lean season. One immediate implication of these observed price effects is that grain markets in the study region are highly fragmented.

We find that these general equilibrium effects also greatly alter the profitability of the loan. By dampening the arbitrage opportunity posed by seasonal price fluctuations, treated individuals in areas highly saturated with loans show diminished revenue gains relative to farmers in lower saturation areas. We find that while treated farmers in high-saturation areas store significantly more than their control counterparts, doing so is not significantly more profitable; the reduction in seasonal price dispersion in these areas may reduce the benefits of loan adoption. In contrast, treated farmers in low-density areas have both significantly higher inventories and significantly higher profits relative to control.

These general equilibrium effects — and their impact on loan profitability at the individual level — have lessons for both policy and research. In terms of policy, the general equilibrium effects shape the distribution of the welfare gains of the harvest-time loan: while recipients gain

---

[1]Though we do find some evidence of heterogeneity by treatment saturation.

relatively less than they would in the absence of such effects, we find suggestive evidence that non-recipients benefit from smoother prices, even though their storage behavior remains unchanged. Though estimated effects on non-treated individuals are relatively imprecisely estimates, a welfare calculation taking the point estimates at face-value suggests that 69% of overall gains in high-treatment-intensity areas accrue to program non-recipients. These gains to non-recipients, which cannot be readily recouped by private sector lending institutions, may provide some incentive for public provision of such products.

The eroding profitability of arbitrage that we observe in highly loan saturated areas also has implications for impact evaluation in the context of highly fragmented markets, such as the rural markets in this study. When general equilibrium effects are pronounced and the SUTVA assumption violated (Rubin 1986), the evaluation of an individually-randomized loan product may conclude there is a null effect even when there are large positive social welfare impacts. While this issue may be particularly salient in our context of a loan explicitly designed to enable arbitrage, it is by no means unique to our setting. Any enterprise operating in a small, localized market or in a concentrated industry may face price responses to shifts in own supply, and credit-induced expansion may therefore be less profitable than it would be in a more integrated market or in a less concentrated industry. Proper measurement of these impacts requires a study design with exogenous variation in treatment density.

The results speak to a large literature on microfinance, which finds remarkably heterogenous impacts of expanded credit access. Experimental evaluations have generally found that small enterprises randomly provided access to traditional microfinance products are subsequently no more productive on average than the control group, but that subsets of recipients often appear to benefit.[2]

Here we study a unique microcredit product designed to improve the profitability of small farms, a setting that has been largely outside the focus of the experimental literature on credit

---

[2]Experimental evaluations of microcredit include Attanasio et al. (2011); Crepon et al. (2011); Karlan and Zinman (2011); Banerjee et al. (2013); Angelucci et al. (2013) among others. See Banerjee (2013) and Karlan and Morduch (2009) for nice recent reviews of these literatures. A related literature on providing cash grants to households and small firms suggest high rates of return to capital in some settings but not in others. Studies finding high returns to cash grants include De Mel et al. (2008); McKenzie and Woodruff (2008); Fafchamps et al. (2013); Blattman et al. (2014). Studies finding much more limited returns include Berge et al. (2011) and Karlan et al. (2012).

constraints. Why do we find positive effects on firm profitability when many other experimental studies on microcredit do not? First, unlike most of the settings examined in the literature, using credit to "free up" storage for price arbitrage is a nearly universally available investment opportunity that does not depend on entrepreneurial skill.[3] Farmers do not even have to sell grain to benefit from credit in this context: a net-purchasing farm household facing similar seasonal cash constraints could use credit and storage to move its purchases from times of high prices to times of lower prices. Second, the terms of repayment on the loan we study are flexible, which has been showing to be important for encouraging investment (Field et al., 2012). Finally, as described above, the general equilibrium effects of credit expansion could alter individual-level treatment effect estimates in a number of ways, potentially shaping outcomes for both treated and untreated individuals (Breza and Kinnan, 2018)). This is a recognized but unresolved problem in the experimental literature on credit, and few experimental studies have been explicitly designed to quantify the magnitude of these general equilibrium effects (Acemoglu, 2010; Karlan et al., 2012).[4] Our results suggest that, at least in our rural setting, treatment density matters and market-level spillovers can substantially shape individual-level treatment effect estimates.[5]

Beyond contributing to the experimental literature on microcredit, our paper is closest to a number of recent papers that examine the role of borrowing constraints in households' storage decisions and seasonal consumption patterns.[6] Using secondary data from Kenya, Stephens and

---

[3]Existing studies have concluded that many small businesses or potential micro-entrepreneurs simply might not possess profitable investment opportunities (Banerjee et al., 2013; Fafchamps et al., 2013; Karlan et al., 2012; Banerjee, 2013) or may lack the managerial skill or ability to channel capital towards these investments (Berge et al., 2011; Bruhn et al., 2012).

[4]For instance, Karlan et al. (2012) conclude by stating, "Few if any studies have satisfactorily tackled the impact of improving one set of firms' performance on general equilibrium outcomes.... This is a gaping hole in the entrepreneurship development literature." Indeed, positive spillovers could explain some of the difference between the experimental findings on credit, which suggest limited effects, and the estimates from larger-scale natural experiments, which tend to find positive effects of credit expansion on productivity – e.g. Kaboski and Townsend (2012). Acemoglu (2010) uses the literature on credit market imperfections to highlight the understudied potential role of GE effects in broad questions of interest to development economists.

[5]Whether these GE effects also influenced estimated treatment effects in the more urban settings examined in many previous studies is unknown, although there is some evidence that spillovers do matter for microenterprises who directly compete for a limited supply of inputs to production. For example, see De Mel et al. (2008) and their discussion of returns to capital for bamboo sector firms, which must compete over a limited supply of bamboo. In any case, our results suggest that explicit attention to GE effects in future evaluations of credit market interventions is likely warranted.

[6]In a forebearer to this literature, McCloskey and Nash (1984) attribute the dramatic reduction in seasonal grain price fluctuations observed in England between the 14th and 17th centuries to a reduction in interest rates.

Barrett (2011) argue that credit constraints substantially alter smallholder farmers' marketing and storage decisions, and Basu and Wong (2012) show that allowing farmers to borrow against future harvests can substantially increase lean-season consumption. Similarly, Dillion (2017) finds that an administrative change in the school calendar that moved the timing of school fee payments to earlier in the year in Malawi forced credit constrained households with school-aged children to sell their crops earlier and at a lower price. Fink et al. (2014) shows that agricultural loans aimed at alleviating seasonal labor shortages can improve household welfare in Zambia, while Beaman et al. (2014) find in Mali that well-timed credit access can increase investment in agricultural inputs.

As in these related papers, our results show that financial market imperfections lead households to turn to increasingly costly ways to move consumption around in time. In our particular setting, credit constraints combined with post-harvest cash needs cause farmers to store less than they would in an unconstrained world. In this setting, even a relatively modest expansion of credit affects local market prices, to the apparent benefit of both those with and without access to this credit.

The remainder of the paper proceeds as follows. Section 2 describes the setting and the experiment. Section 3 describes our data, estimation strategy, and pre-analysis plan. Section 4 presents baseline estimates ignoring the role of general equilibrium effects. Section 5 presents the market level effects of the intervention. Section 6 shows how these market-level effects shape the individual-level returns to the loan. Section 7 concludes.

2 Setting and experimental design

2.1 Arbitrage opportunities in rural grain markets

Seasonal fluctuations in prices for staple grains appear to offer substantial intertemporal arbitrage opportunities, both in our study region of East Africa as well as in other parts of the developing world. While long-term price data do not exist for the small, rural markets where our experiment takes place, price series data are available for major markets throughout the region. Average seasonal price fluctuations for maize in available markets are shown in Figure 1. Increases in maize

prices in the six to eight months following harvest average roughly 25-40% in these markets; price increase reported elsewhere in Africa are consistent with these figures, if not higher.[7]

These increases also appear to be a lower bound on typical increase observed in the smaller markets in our study area, which (relative to these much larger markets) are characterized with much smaller "catchments" and less outside trade. We asked farmers at baseline to estimate average monthly prices of maize at their local market point over the five years prior to our experiment. As shown in Panel A of Figure 4, they reported a typical doubling in price between September (the main harvest month) and the following June.[8] We also collected monthly price data from local market points in our sample area during the two years of this study's intervention, as well as for a year after the intervention ended (more on this data collection below).[9] Panel B of Figure 4 presents the price fluctuations observed during this period. Because data collection began in November 2012 (two months after the typical trough in September), we cannot calculate the full price fluctuation for the 2012-2013 season. However, in the 2013-2014 and 2014-2015 seasons we observe prices increasing by 42% and 45% respectively. These are smaller fluctuations than those seen in prior years (as reported by farmers in our sample) and smaller than those seen in subsequent years, which saw increases of 53% and 125% respectively.[10] There is therefore some variability in the precise size of the price fluctuation from season to season. Nevertheless, we see price consistently rise by more than 40% and, in some years, by substantially more.

These fluctuations have meaningful and negative consequences for the welfare of rural households. Food price seasonality drives large fluctuations in consumption, with both food and non-food consumption dropping noticeably during the lean season (Kaminski and Gilbert, 2014; Basu and

---

[7]For instance, Barrett (2007) reports seasonal rice price variation in Madagascar of 80%, World Bank (2006) reports seasonal maize price variation of about 70% in rural Malawi, and Aker (2012) reports seasonal variation in millet prices in Niger of 40%.

[8]In case farmers were somehow mistaken or overoptimistic, we asked the same question of the local maize traders that can typically be found in these market points. These traders report very similar average price increases: the average reported increase between October and June across traders was 87%. Results available on request.

[9]The study period covers the 2012-2013 and 2013-2014 season. We also collect data for one year after the study period, covering the 2014-2015 season, in order to align with the long-run follow-up data collection on the farmer side.

[10]For the 2015-2016 season, we combine our data with that collected by Bergquist (2017) in the same county in Kenya and estimate that maize prices increased by 53% from November to June. For the 2016-2017 season, we thank Pascaline Dupas for her generosity in sharing maize price data collected in the same county in November 2016 and June 2017, from which we estimate an increase of 125%.

Wong, 2012). Barrett and Dorosh (1996) find that the greatest burden of such price fluctuations falls on the poorest of farmers.

These price fluctuations are surprising in light of the storable nature of staple commodities. Home storage is a simple technology available to farmers in this region. To store, farmers dry maize kernels on a tarp immediately after harvest, treat the crop with insecticide dust, and store it in locally-made sacks, kept on wooden pallets to allow for air circulation and typically located in farmers' homes or in small outdoor sheds. Our survey data suggests the cost of these storage materials is low, at around 3.5% of the value of the crop at harvest time. Post-harvest losses also appear minimal in this setting, with an average of 2.5% of the crop lost over a 6-9 month storage period (see Appendix B for further discussion). The low cost of storage, in conjunction with consistently large price increases over the course of the season, therefore appears to offer large opportunities for arbitrage.

However, farmers do not appear to be taking advantage of these apparent arbitrage opportunities. Figure A.1 shows data from two earlier pilot studies conducted either by our NGO partner (in 2010/11, with 225 farmers) or in conjunction with our partner (in 2011/12, with a different sample of 700 farmers). These studies tracked maize inventories, purchases, and sales for farmers in our study region. In both years, the median farmer exhausted her inventories about 5 months after harvest, and at that point switched from being a net seller of maize to a net purchaser as shown in the right panels of the figure. This was despite the fact that farmer-reported sales prices rose by more than 80% in both of these years in the nine months following harvest.

Why are farmers not using storage to sell grain at higher prices and purchase at lower prices? Our experiment is designed to test one specific explanation: that credit constraints limit farmers ability to arbitrage these price fluctuations.[11] In extensive focus groups with farmers prior to our experiment, credit constraints were the unprompted explanation given by the vast majority of these farmers as to why they sold the majority of their maize in the immediate post-harvest period. In particular, because nearly all of these farm households have school-aged kids, and a large percentage of a child's school fees are typically due in the few months after harvest in

---

[11]Other factors, mostly outside the scope of this paper, may also be at play in limiting farmers' ability to store. Appendix B explores these other factors in greater detail.

January, given the calendar-year school year schedule, many farmers report selling much of their harvest to pay these fees. Indeed, many schools in the area will accept in-kind payment in maize during this period. Farmers also report having to pay other bills they have accumulated throughout the year during the post-harvest period. Finally, many farmers also spend more on discretionary expenditures during this harvest period as well, which may be reflective of high levels of impatience or present-biased preferences. Regardless of the source, harvest is a time of large expenditure.

Why do these high harvest-time expenditures necessitate high harvest-time sales of maize? In the absence of functioning financial markets, the timing of production and consumption – or, more specifically, sale and expenditure – must be intimately tied. As with poor households throughout much of the world, farmers in our study area appear to have very limited access to formal credit. Only eight percent of households in our sample reported having taking a loan from a bank in the year prior to the baseline survey.[12] Informal credit markets also appear relatively thin, with fewer than 25% of farmers reporting having given or received a loan from a moneylender, family member, or friend in the 3 months before the baseline.

Absent other means of borrowing, and given the high expenditure needs they report facing in the post-harvest period, farmers end up liquidating grain rather than storing. Furthermore, a significant percentage of these households end up buying back maize from the market later in the season to meet consumption needs, and this pattern of "sell low and buy high" directly suggests a liquidity story: farmers are in effect taking a high-interest quasi-loan from the maize market (Stephens and Barrett, 2011). Baseline data indicate that 35% of our sample both bought and sold maize during the previous crop year (September 2011 to August 2012), and that over half of these sales occurred before January (when prices were low). 40% of our sample reported only purchasing maize over this period, and the median farmer in this group made all of their purchases after January. Stephens and Barrett (2011) report similar patterns for other households in Western Kenya during an earlier period.

---

[12]Note that even at the high interest rates charged by formal banking institutions (typically around 20% annually), storage would remain profitable, given the 40% plus (often much larger) increases in prices that are regularly observed over the 9-month post-harvest period and relatively small storage costs and losses (e.g., due to spoilage), which we estimate to be less than 6%.

2.2 Experimental design

To test the hypothesis that the limited availability of credit constrains farmers from taking advantage of the arbitrage opportunities presented by seasonal price fluctuations, we partner with the organization One Acre Fund (OAF) to offer farmers harvest-time loan. OAF is a microfinance NGO that makes in-kind, joint-liability loans of fertilizer and seed to groups of farmers, as well as providing training on improved farming techniques. OAF group sizes typically range from 8-12 farmers, and farmer groups are organized into "sublocations" – effectively clusters of villages that can be served by one OAF field officer.[13] OAF typically serves about 30% of farmers in a given sublocation.

The study sample is drawn from existing groups of One Acre Fund (OAF) farmers in Webuye and Matete districts in Western Kenya. The Year 1 sample consists of 240 existing OAF farmer groups drawn from 17 different sublocations in Webuye district, and our total sample size at baseline was 1,589 farmers. The Year 2 sample attempts to follow the same OAF groups as Year 1; however, some groups dissolved such that in Year 2 we are left with 171 groups. In addition, some of the groups experienced substantial shifting of the individual members; therefore some Year 1 farmers drop out of our Year 2 sample, and other farmers are new to our Year 2 sample.[14] Ultimately, of the 1,019 individuals in our Year 2 sample, 602 are drawn from the Year 1 sample and 417 are new to the sample.

Figure 4 displays the experimental design. There are two main levels of randomization. First, we randomly divided the 17 sublocations in our sample into 9 "high" intensity" sites and 8 "low intensity" sites. In high intensity sites, we enrolled 80% of OAF groups in the sample (for a sample of 171 groups), while in low intensity sites, we only enrolled 40% of OAF groups in the sample (for a sample of 69 groups). Then, within each sublocation, groups were randomized into treatment or control. In Year 1, two-thirds of individuals in each sublocation were randomized into

---

[13]A sublocation is a group of 4-5 villages, with a typical population of 400-500 people.

[14]Shifting of group members is a function of several factors, including whether farmers wished to participate in the overall OAF program from year to year. There was some (small) selective attrition based on treatment status in Year 1; treated individuals were 10 percentage points more likely to return to the Year 2 sample than control individuals (significant at 1%). This does slightly alter the composition of the Year 2 sample (see Table L.3 and Section L), but because Year 2 treatment status is stratified by Year 1 treatment status (as will be described below), it does not alter the internal validity of the Year 2 results.

treatment (more on this below) and one-third into control. In Year 2, half of individuals in each sublocation were randomized into treatment and half into control. As a result of this randomization procedure, high intensity sublocations have double the number of treated individuals as in low intensity sublocations.

The group-level randomization was stratified at the sublocation level (and in Year 1, for which we had administrative data, further stratified based on whether group-average OAF loan size in the previous year was above or below the sample median). In Year 2, we maintained the same saturation treatment status at the sublocation level,[15] but re-randomized groups into treatment and control, stratifying on their treatment status from Year 1.[16] Given the roughly 35% reduction in overall sample size in Year 2, overall treatment saturation rates (the number of treated farmers per sublocation) were effectively 35% lower in Year 2 as compared to Year 1.

In Year 1, there was a third level of randomization pertaining to the timing of the loan offer. In focus groups run prior to the experiment, farmers were split on when credit access would be most useful, with some preferring cash immediately at harvest, and others preferring it a few months later timed to coincide with when school fees were due (the latter preferences suggesting that farmers may be sophisticated about potential difficulties in holding on to cash between the time it was disbursed and the time it needed to be spent). In order to test the importance of loan timing, in Year 1, a random half of the treated group (so a third of the total sample) received the loan in October (immediately following harvest), while the other half received the loan in January (immediately before school fees are due, although still several months before the local lean season). As will be described in Section 4, results from Year 1 suggested that the earlier loan was more effective, and therefore in Year 2 OAF only offered the earlier timed loan to the full sample (though due to administrative delays, the actual loan was disbursed in November in Year 2).

Although all farmers in each loan treatment group were offered the loan, we follow only a randomly selected 6 farmers in each loan group, and a randomly selected 8 farmers in each of the

---

[15]Such that, for example, if a sublocation was a high intensity sublocation in Year 1 it remained a high intensity sublocation in Year 2.

[16]This was intended to result in randomized duration of treatment – either zero years of the loan, one year of the loan, or two years – however, because the decision to return to the Year 2 sample was affected by Year 1 treatment status, we do not use this variation here and instead focus throughout on one year impacts.

control groups.

Loan offers were announced in September in both years. The size of the loan for which farmers were eligible was a linear function of the number of bags they had in storage at the time of loan disbursal.[17] In Year 1, there was a cap of 7 bags for which farmers could be eligible; in Year 2, this cap was 5 bags. In Year 1, to account for the expected price increase, October bags were valued at 1500 Ksh, and January bags at 2000 Ksh. In Year 2, bags were valued at 2500 Ksh. Each loan carried with it a "flat" interest rate of 10%, with full repayment due after nine months.[18][19] These loans were an add-on to the existing in-kind loans that OAF clients received, and OAF allows flexible repayment of both – farmers are not required to repay anything immediately.

OAF did not take physical or legal position of the bags, which remained in farmers' home stores. Bags were tagged with a simple laminated tag and zip tie. When we mentioned in focus groups the possibility of OAF running a harvest loan program, and described the details about the bag tagging, many farmers (unprompted) said that the tags alone would prove useful in shielding their maize from network pressure: "branding" the maize as committed to OAF, a well-known lender in the region, would allow them to credibly claim that it could not be given out.[20] Because tags could represent a meaningful treatment in their own right, we wished to separate the effect of the credit from any effect of the tag, and therefore in the Year 1 study offered a separate treatment arm in which groups received only the tags.[21]

Finally, because self- or other-control problems might make it particularly difficult to channel cash toward productive investments in settings where there is a substantial time lag between when the cash is delivered and when the desired investment is made, in Year 1, we also cross-randomized

---

[17]However, there was no further requirement that farmers store beyond the date of loan disbursal. This requirement was set by OAF to ensure that farmers took a "reasonable" loan size that they would be able to repay.

[18]Annualized, this interest rate is slightly lower than the 16-18% APR charged on loans at Equity Bank, the main rural lender in Kenya.

[19]For example, a farmer who committed 5 bags when offered the October loan in Year 1 would receive 5*1500 = 7500 Ksh in cash in October (∼$90 at current exchange rates), and would be required to repay 8250 Ksh by the end of July.

[20]Such behavior is consistent with evidence from elsewhere in Africa that individuals take out loans or use commitment savings accounts mainly as a way to demonstrate that they have little to share with others(Baland et al., 2011; Brune et al., 2011).

[21]This is not the full factorial research design – there could be an interaction between the tag and the loan – but we did not have access to a sufficiently large sample size to implement the full 2 x 2 design to isolate any interaction effect.

a simple savings technology that had shown promise in a nearby setting (Dupas and Robinson, 2013). In particular, a subset of farmers in each loan treatment group in Year 1 were offered a savings lockbox (a simple metal box with a sturdy lock) which they could use as they pleased. While such a savings device could have other effects on household decision making, our hypothesis was that it would be particularly helpful for loan clients who received cash before it was needed.

The tags and lockbox treatments were randomized at the individual level during Year 1. These treatments were not included in Year 2 due to minimal treatment effects in Year 1 data (discussed below), as well as the somewhat smaller sample size in Year 2. Using the sample of individuals randomly selected to be followed in each group, we stratified individual level treatments by group treatment assignment and by gender. So, for instance, of all of the women who were offered the October loan and who were randomly selected to be surveyed, one third of them were randomly offered the lockbox (and similarly for the men and for the January loan). In the control groups, in which we were following 8 farmers, 25% of the men and 25% of the women were randomly offered the lockbox, with another 25% each being randomly offered the tags. The study design allows identification of the individual and combined effects of the different treatments, and our approach for estimating these effects is described below.

3   Data and estimation

The timing of the study activities is shown in Figure 3. In August/September 2012 (prior to the Year 1 experiment), a baseline survey was conducted with the entire Year 1 sample. The baseline survey collected data on farming practices, storage costs, maize storage and marketing over the previous crop year, price expectations for the coming year, food and non-food consumption expenditure, household borrowing, lending, and saving behavior, household transfers with other family members and neighbors, sources of non-farm income, time and risk preferences, and digit span recall. Table N.1 presents summary statistics for a range of variables at baseline; we observe balance on most of these variables across treatment groups, as would be expected from randomization. Table K.1 shows the analogous table comparing individuals in the high- and low-treatment-density areas; the samples appear balanced on observables here as well.

We then undertook three follow-up rounds over the ensuing 12 months, spanning the spring 2013 "long rains" planting (the primary growing season) and concluding just prior to the 2013 long rains harvest (which occurs August-September). The multiple follow-up rounds were motivated by three factors. First, a simple inter-temporal model of storage and consumption decisions suggests that while the loan should increase total consumption across all periods, the per-period effects could be ambiguous – meaning that consumption throughout the follow-up period needs to be measured to get at overall effects. Second, because nearly all farmers deplete their inventories before the next harvest, inventories measured at a single follow-up one year after treatment would likely provide very little information on how the loan affected storage and marketing behavior. Finally, as shown in McKenzie (2012), multiple follow-up measurements on noisy outcomes variables (e.g consumption) has the added advantage of increasing power. The follow-up surveys tracked data on storage inventory, maize marketing behavior, consumption, and other credit and savings behavior. Follow-up surveys also collected information on time preferences and on self-reported happiness.

Because the Year 2 experiment was designed to follow the sample sample as Year 1, a second baseline was not run prior to Year 2.[22] A similar schedule of three follow-up rounds over 12 months was conducted in Year 2 following the loan disbursal.

Attrition was relatively low across survey rounds. In Year 1, overall attrition was 8%, and not significantly different across treatment groups (8% in the treatment group and 7% in the control). In Year 2, overall attrition was 2% (in both treatment and control, with no significant difference).

Year 1 treatment status is predictive of Year 2 re-enrollment in the study (treated individuals were more likely to re-register for OAF in the second year, perhaps reflecting a positive appraisal of the value of the loan). However, because Year 2 treatment status was re-randomized and stratified by Year 1 treatment status, this does not alter the internal validity of the Year 2 results.[23, 24]

---

[22]In practice, due to the administrative shifts in farmer group composition described in greater detail in Section 2, 417 of the 1,019 individuals in the Year 2 sample were new to the study. For these individuals, we do not have baseline data (there was insufficient time between receiving the updated administrative records for Year 2 groups and the disbursal of the loan to allow for a second baseline to be run). Therefore, balance tables can only be run with the sample that was present in Year 1. Because the loan offer was randomized, however, this should not affect inference regarding the impacts of the loan.

[23]This does, however, mean that we cannot exploit the re-randomization in Year 2 to identify the effect of receiving the loan for multiple years or of receiving the loan and then having it discontinued, as an endogenously selected group did not return to the Year 2 sample and therefore was never assigned a Year 2 treatment status.

[24]This also slightly alters the composition of the Year 2 sample, relevant to external validity. Appendix L explores

In order to explore the long-run effects of the loan, we also ran a Long-Run Follow-Up (LRFU) survey from November-December 2015. This was two (one) years following loan repayment for the Year 1 (Year 2) treatment group. This survey followed up on the entire Year 2 sample (1,091 individuals) and a representative subset of the Year 1 only sample (another 481 individuals), for a total sample of 1500 individuals. The survey collected information on maize harvests, sales, purchases, and revenues by season from 2014-2015. It also collected data on farm inputs (labor and capital), food consumption and expenditure, household consumption, educational expenditure and attendance among children, non-farm employment and revenues, and a self-reported happiness measure. Attrition in the LRFU was 9%, with no differential attrition based on Year 2 treatment status and slight differential attrition based on Year 1 treatment status.[25] Appendix L provides further discussion.

In addition to farmer-level surveys, we also collected monthly price surveys at 52 market points in the study area. The markets were identified prior to treatment based on information from local OAF staff about the market points in which client farmers typically buy and sell maize. Data collection for these surveys began in November 2012 and continued through December 2015. Finally, we utilize administrative data on loan repayment that was generously shared by OAF.

3.1 Pre-analysis plan

To limit both risks and perceptions of data mining and specification search (Casey et al., 2012), we registered a pre-analysis plan (PAP) for Year 1 prior to the analysis of any follow-up data.[26] The Year 2 analysis follows a near identical analysis plan. The PAP can be found in Appendix N.

We deviate significantly from the PAP in one instance: the PAP specifies that we will analyze the effect of treatment saturation on the percent price spread from November to June. However, because in practice the loan was offered at slight different points in time (October and January in

---

this further.

[25]Being treated in Year 1 is associated with a 3 percentage point increase in the likelihood of being found in the long-run follow up survey, significant at 10%. This appears to be at least partially driven by the fact that Year 1 treated individuals were more likely to be in the Year 2 sample (and therefore had been more recently in touch with our survey team). After controlling for whether an individual was present in the Year 2 sample, Year 1 treatment status is no longer significantly correlated with attrition.

[26]The pre-analysis plan is registered here: https://www.socialscienceregistry.org/trials/67, and was registered on September 6th 2013. The complete set of results are available upon request.

Year 1; November in Year 2) and because there is year-to-year variation in when markets hit their peak and trough, this measure may fail to capture the full effect of treatment on prices. Moreover, this measure is statistically underpowered, ignoring 77% of our monthly data by focusing solely on the price gap between two months, rather than exploiting the full nine months of data collected over the season.

Therefore, in our primary specifications, we relax our attachment to this underpowered and perhaps misspecified measure November-June price gap, instead analyze the non-parametric effect of treatment on the evolution of monthly prices, as well as a level and time trend effect. Appendix I.3 presents the pre-specified November-June effect. For all analyses, we maintain our original hypothesis that effect of high-density treatment on prices will be initially positive if receipt of the loan allows farmers to pull grain off the market in the post-harvest surplus period and later negative as stored grain is released onto the market.

In two other instances we add to the PAP. First, in addition to the regression results specified in the PAP, we also present graphical results for many of the outcomes. These results are based on non-parametric estimates of the parametric regressions specified in the PAP, and are included because they clearly summarize how treatment effects evolve over time, but since they were not explicitly specified in the PAP we mention them here. Second, we failed to include in the PAP the (ex post obvious) regressions in which the individual-level treatment effect is allowed to vary by the sublocation-level treatment intensity, and present these below.

3.2   Estimation of treatment effects

In all analyses, we present results separately by year and pooled across years. Because the Year 2 replication produced results that are quantitatively quite similar to the Year 1 results for most outcomes, we rely on the pooled results as our specification of primary interest. However, for the sake of transparency and comparison, we report both.

There are three main outcomes of interest: inventories, maize net revenues, and consumption. Inventories are the number of 90kg bags of maize the household had in their maize store at the time of the each survey. This amount is visually verified by our enumeration team, and so is likely

16

to be measured with minimal error. We define maize net revenues as the value of all maize sales minus the value of all maize purchases, and minus any additional interest payments made on the loan for individuals in the treatment group. We call this "net revenues" rather than "profits" since we likely do not observe all costs; nevertheless, costs are likely to be very similar across treatment groups (fixed costs of storing at home were already paid, and variable costs of storage are very low). The values of sales and purchases were based on recall data over the period between each survey round. Finally, we define consumption as the log of total household expenditure over the 30 days prior to each survey. For each of these variables we trim the top and bottom 0.5% of observations, as specified in the pre-analysis plan.

Letting T_{jy} be an an indicator for whether group j was assigned to treatment in year y, and Y_{ijry} as the outcome of interest for individual i in group j in round $r \in (1, 2, 3)$ in year y. The main specification pools data across follow-up rounds 1-3 (and for the pooled specification, across years):

$$Y_{ijry} = \alpha + \beta T_{jy} + \eta_{ry} + d_t + \varepsilon_{ijry} \tag{1}$$

The coefficient β estimates the Intent-to-Treat and, with round-year fixed effects η_{ry}, is identified from within-round variation between treatment and control groups. β can be interpreted as the average effect of being offered the loan product across follow-up rounds, though as we detail below, loan take-up was high. To absorb additional variation in the outcomes of interest, we also control for survey date (d_t) in the regressions. Each follow-up round spanned over three months, meaning that there could be (for instance) substantial within-round drawdown of inventories. Inclusion of this covariate should help to make our estimates more precise without biasing point estimates. Standard errors are clustered at the loan group level.

The assumption in Equation 8 is that treatment effects are constant across rounds. In our setting, there are reasons why this might not be the case. In particular, if treatment encourages storage, one might expect maize revenues to be *lower* for the treated group immediately following harvest, as they hold off selling, and *greater* later on during the lean season, when they release their stored grain. To explore whether treatment effects are constant across rounds, we estimate:

$$Y_{ijry} = \sum_{r=1}^{3} \beta_r T_{jy} + \eta_{ry} + d_t + \varepsilon_{ijry} \tag{2}$$

and test whether the β_r are the same across rounds (as estimated by interacting the treatment indictor with round dummies). Unless otherwise indicated, we estimate both (8) and (9) for each of the hypotheses below.

To explore heterogeneity in treatment effects, we estimate:

$$Y_{ijry} = \alpha + \beta_1 T_{jy} + \beta_2 Z_{i0} + \beta_2 T_{jy} * Z_{i0} + \eta_{ry} + d_t + \varepsilon_{ijry} \tag{3}$$

where Z_{i0} is the standardized variable by which we explore heterogeneity, as measured at baseline. As pre-specified, we explore heterogeneity by impatience (as measured in standard time preference questions), the number of school-aged children, the initial liquid wealth level, the percent of baseline sales sold early (prior to January 1), and the seasonal price increase expected between September 2012 and June 2013. Because a baseline was only run prior to Year 1, we are only able to present these specifications for the Year 1 intervention.

To quantify market level effects of the loan intervention, we tracked market prices at 52 market points throughout our study region, and we assign these markets to the nearest sublocation. To estimate price effects we begin by estimating the following linear model:

$$p_{msty} = \alpha + \beta_1 H_s + \beta_2 month_t + \beta_3 (H_s * month_t) + \varepsilon_{mst} \tag{4}$$

where p_{mst} represents the maize sales price at market m in sublocation s in month t in year y.[27] H_s is a binary variable indicating whether sublocation s is a high-intensity sublocation, and $month_t$ is a time trend (in each year, Nov = 0, Dec = 1, etc). If access to the storage loan allowed farmers to shift purchases to earlier in the season or sales to later in the season, and if this shift in marketing behavior was enough to alter supply and demand in local markets, then our prediction is that $\beta_1 > 0$ and $\beta_3 < 0$, i.e. that prices in areas with more treated farmers are higher after harvest

---

[27]Prices are normalized to 100 among the "low" intensity markets in the first month ($H_s = 0$, $month_t = 0$). Therefore, price effects can be interpreted as a percentage change from control market post-harvest prices.

but lower closer to the lean season.

While H_s is randomly assigned, and thus the number of treated farmers in each sublocation should be orthogonal to other location-specific characteristics that might also affect prices (e.g. the size of each market's catchment), we have only 17 sublocations. This relatively small number of clusters could present problems for inference (Cameron et al., 2008). We begin by clustering errors at the sublocation level when estimating Equation 4. We also report standard errors estimated using both the wild bootstrap technique described in Cameron et al. (2008) and the randomization inference technique used in Cohen and Dupas (2010).

To understand how treatment density affects individual-level treatment effects, we estimate Equations 8 and 9, interacting the individual-level treatment indicator with the treatment density dummy. The pooled equation is thus:

$$Y_{ijsry} = \alpha + \beta_1 T_{jy} + \beta_2 H_s + \beta_3 (T_{jy} * H_s) + \eta_{ry} + d_t + \varepsilon_{ijsry} \tag{5}$$

If the intervention produces sufficient individual level behavior to generate market-level effects, we predict that $\beta_3 < 0$ and perhaps that $\beta_2 > 0$ - i.e. treated individual in high-density areas do worse than in low density areas, and control individuals in high density areas do better than control individuals in low density areas. As in Equation 4, we report results with errors clustered at the sublocation level.

For long-run effects, we first estimate the following regression for each year separately:

$$Y_{ij} = \alpha + \beta T_{jy} + \varepsilon_{ij} \tag{6}$$

in which Y_{ij} is the outcome of interest for individual i in group j. The sample is restricted to those who were in the Year y study.

We further also estimate the following specification:

$$Y_{ij} = \alpha + \beta_1 T_{j1} + \beta_2 T_{j2} + \beta_3 T_{j1} * T_{j2} + \varepsilon_{ij} \tag{7}$$

in which T_{j1} is an indicator for being an in treated group in year 1, T_{j12} is an indicator for being in a treated group in year 2, and $T_{j1} * T_{j2}$ is an interaction term for being in a group that was treated in both years. The sample is restricted to those who were in the study for both years. Because of this sample restriction, and because decision to return to the sample from the Year 1 to Year 2 study was differential based on treatment status (see Appendix L), this last specification is open to endogeneity concerns and therefore should not be interpreted causally. For the sake of transparency, we present it regardless, but with the aforementioned caveat.

4 Individual level results

4.1 Harvest loan take up

Take-up of the loan treatments was quite high. Of the 954 individuals in the Year 1 treatment group, 610 (64%) applied and qualified for the loan. In Year 2, 324 out of the 522 treated individuals (62%) qualified for and took up the loan.[28]

Unconditional loan sizes in the two treatment groups were 4,817 Ksh and 6,679 Ksh, or about $57 and $79 USD, respectively. The average loan sizes conditional on take-up were 7,533 Ksh (or about $89 USD) for Year 1 and 10,548 Ksh (or $124) for Year 2.[29] Default rates were extremely low, at less than 2%.

---

[28]Relative to many other credit-market interventions in low-income settings in which documented take-up rates range from 2-55% of the surveyed population (Karlan et al., 2010), the 60-65% take-up rates of our loan product were very high. This is perhaps not surprising given that our loan product was offered as a top-up for individuals who were already clients of an MFI. Nevertheless, OAF estimates that about 30% of farmers in a given village in our study area enroll in OAF, which implies that even if *no* non-OAF farmers were to adopt the loan if offered it, population-wide take-up rates of our loan product would still exceed 15%.

[29]Recall in Year 1 there were two versions of the loan, one offered in October and the other in January. Of the 474 individuals in the 77 groups assigned to the October loan treatment (T1), 329 (69%) applied and qualified for the loan. For the January loan treatment (T2), 281 out of the 480 (59%) qualified for and took up the loan. Unconditional loan sizes in the two treatment groups were 5,294 Ksh and 4,345 Ksh (or about $62 and $51 USD) for T1 and T2, respectively, and we can reject at 99% confidence that the loan sizes were the same between groups. The average loan sizes conditional on take-up were 7,627Ksh (or about $90 USD) for T1 and 7,423Ksh (or $87) for T2, and in this case we cannot reject that conditional loan sizes were the same between groups.

4.2 Primary effects of the loan offer

We begin by estimating treatment effects in the standard fashion, assuming that there could be within-randomization-unit spillovers (in our case, the group), but that there are no cross-group spillovers. In all tables and figures, we report results broken down by each year and pooled. As explained in Section 3, the Year 2 replication produced results that are quantitatively quite similar to the Year 1 results for most outcomes, and as such, we report in the text the pooled results, unless otherwise noted.

Tables 2-4 and Figure 5 present the results of estimating Equations 8 and 9 on the pooled treatment indicator, either parametrically (in the table) or non-parametrically (in the figure). The top panels in Figure 5 show the means in the treatment group (broken down by year and then pooled, in the final panel) over time for our three main outcomes of interest (as estimated with Fan regressions). The bottom panels present the difference in treatment minus control over time, with the 95% confidence interval calculated by bootstrapping the Fan regression 1000 times.

Farmers responded to the intervention as anticipated. They held significantly more inventories for much of the year, on average about 25% more than the control group mean (Column 6 in Table 2). Inventory effects are remarkably similar across both years of the experiment.

Net revenues[30] are significantly lower immediately post harvest and significantly higher later in the year (Column 6 in Table 3). The middle panel of Figure 5 presents the time trend of net revenue effects, which suggest that treated farmers purchase more maize in the immediate post-harvest period, when prices are low (as represented by more negative net revenues November to February) and sell more later in the lean season, when prices are high (as represented by more positive revenues May to July). The net effect on revenues averaged across the year is positive in both years of the experiment, and is significant in the Year 2 and the pooled data (see Columns 1, 3, and 5 in Table 3). Breaking down Year 1 results by the timing of loan suggest that the reason results in Year 1 are not significant is that the later loan, offered in January to half of the treatment group, was less effective than the October loan. Table D.1 presents results for the Year 1 loan, broken down by loan timing. We see in Column 5 that the October loan (T1) produced revenue

---

[30]From which loan interest rates were subtracted for those who took out a loan.

effects that are more similar in magnitude (and now significant, at 5%) to those of the Year 2 loan (which was offered almost at the same time). The January loan (T2) had no significant effect on revenues. Appendix Section D explores the effects of loan timing in greater detail.

The total effect on net revenues across the year can be calculated by adding up the coefficients in Column 6 of Table 3,which yields an estimate of 1,548 Ksh, or about $18 at the prevailing exchange rate at the time of the study. Given the unconditional average loan size of 5,476 Ksh in the pooled data, this is equivalent to a 28% return (net of loan and interest repayment), which we consider large.

The final panel of Figure 5 and Table 4 present the consumption effects (as measured by logged total household consumption). While point estimates are positive in both years, they are not significant at traditional confidence levels when pooled (in Year 2, treatment is associated with a 7 percentage point increase in consumption, significant at 10%, but in Year 1, estimated effects are only slightly greater than zero and are not significant).[31]

Table 5 presents effects on the pattern of net sales (quantity sold - quantity purchased), as well as prices paid and received. We see that in the immediate post-harvest period, net sales are significantly lower among the treated group, as sales decrease/purchases increase. Later in the season, this trend reverses, as net sales significantly expand among the treated. As a result of this shifted timing of sales and purchases, treated individuals enjoy significantly lower purchase prices (as prices are shifted to earlier in the season, when prices tend to be lower) and receive significantly higher sales prices (as sales are shifted to later in the season, when prices tend to be higher). The total impact on net sales is a weakly positive effect, which – off of a negative average net sales amount – means that households are slightly less in deficit.[32,33]

---

[31]Because the consumption measure includes expenditure on maize, in Appendix E.1 we also estimate effects on consumption excluding maize and consumption excluding all food. Results are similar using these measures.

[32]Unlike the impact on net sales *per round*, on which we have strong theoretical predictions, the impact on total net sales is ex-ante ambiguous, from a theoretical perspective. In practice, the total effect on net sales will be a combined response of the increase in purchases in response to lower effective purchase prices and increases in sales in response to higher effective sales prices. Valuing the increase in net sales, we estimate that 48% of the increase in revenues among treated individuals is driven by an increase in net sales and 52% by a shifting in the timing of sales.

[33]From where is the increase in net sales drawn? We assume net sales = amount harvested - post-harvest losses - amount consumed - amount transferred and decompose the treatment effect on each component part. We see a marginally significant (at 10%) increase in amount that treated households transfer to others by 0.2 bags. We are unable to identify with precision any effects on the other components of net sales (results available upon request).

4.3 Heterogeneity

Tables C.1-C.3 in Appendix C present the pre-specified dimensions of heterogeneity in treatment effects on inventories, revenues and log household consumption. Because the pre-specified specification is an intention-to-treat estimation, we also present a regression of take-up on the standardized variable of heterogeneity. While we see greater take-up of the loan by impatient households and households with more school-aged children, we see no significant heterogeneity in treatment effects by these dimensions. We observe slightly larger treatment effects among wealthy households (marginally significant for revenue outcomes, but not significant for inventories or consumption). Interestingly, we see significantly and large increases in the estimated treatment effects for households with a larger percentage of early sales at baseline (that is, those who were less likely to store at baseline). It may be that these households have the greatest room for movement in storage behavior and/or that these households were most constrained at baseline. For inventories and revenues, treatment appears to cut in half the gap between the baseline storers and non-storers. Expectations regarding the impending seasonal price increase does not appear to be related to take-up or treatment effects.

4.4 Secondary effects of the loan offer

Appendix Section E presents outcomes on potential secondary outcomes of interest. We find no significant effects on profits earned from and hours worked at non-farm household-run businesses (Tables E.1-E.2), nor on wages earned from and hours worked in salaried employment (Tables E.3 and E.4). We also find no significant effects on schools fees paid (the primary expenditure that households say constrain them to sell their maize stocks early; see Table E.7), nor do we find significant effects on food expenditure (Table E.5). We do in Year 1 find a significant 0.07 point increase on a happiness index (an index for the following question: "Taking everything together, would you say you are very happy (3), somewhat happy (2), or not happy (1)"). However, we find no significant increase in this measure in Year 2.

4.5 Long-run effects

Appendix Section F presents the long-run effects of the loan, as measured in the Long-Run Follow-Up (LRFU) survey conducted November-December 2015, which measures outcomes one to two years after completion of the intervention (for the Year 2 and Year 1 loan respectively). In this section, we primarily focus on the effects of each year of the study as estimated separately, as these results can be interpreted causally. For the sake of transparency, we also present a specification with the two treatment years interacted, but with the aforementioned caveats described in Section 3.

We first explore outcomes for the 2014 long-rains harvest, the season immediately following the completion of the Year 2 study. If farmers are able to use revenues from the one- (sometimes two-) time loan to "save their way" out of this credit constraint, we should expect to see sustained shifts in the timing of sales, as well as long-run revenue effects. However, in Table F.1 we observe no statistically significant differences in the timing of transactions (neither in terms of the percent of purchases conducted in the low-price harvest season nor the percent of sales conducted in the high-price lean season). We also see no statistically significant difference in long-run net revenues (though due to the imprecision of these estimates, we cannot rule out large, positive effects).[34] We also see no long-run effect on amount and value sold or purchased (Tables F.2-F.4), though again estimates are relatively noisy.

We are able to ask more detailed questions about the subsequent season (the 2015 long-rains harvest), which occurred immediately prior to the LRFU survey and therefore required shorter recall. Measuring impacts on input usage and harvest levels, we test the hypothesis that loan

---

[34]While we see no significant changes in sales timing or revenue in among the pooled treatment group, we see when breaking these results down by treatment status some interesting heterogeneity (see Table F.10. Point estimates suggest (and are significant in Year 2) that the percent sold in the lean season and the percent purchased in the harvest season are higher in low-saturation areas. In high saturation areas, the negative interaction terms cancels this effect out (see Table F.10). This is consistent with the idea that in low intensity areas, the lack of effect on prices means storage is highly profitable, encouraging individuals to purchase more in the post-harvest period and sell more in the lean season. In contrast, in high intensity areas, price effects dampen the returns to arbitrage, and there is lower incentive to store. However, we see that control individuals in high intensity areas may be storing more, buying more (significant among Year 2 individuals) in the harvest period, when prices are low. As a result, we see cannot rule out sizable increases in revenues for control individuals in high-intensity areas; though this effect is measured with considerable noise, it is consistent with the idea that control individuals may benefit from the loan. See Appendix F for greater discussion of this heterogeneity.

access produced long-run increases in on-farm investment.[35] However, Table F.5 suggests little movement on this margin. We estimate fairly precise null effects on labor inputs, non-labor inputs, and 2015 long-rains harvest levels.

We also explore other outcomes for the 2015 year. We find no significant effects on a variety of outcomes, including: maize eaten, food expenditures, consumption, educational expenditure, school attendance, non-farm enterprise profits, hours worked in non-farm enterprises, and hours worked in salaried positions (Tables F.6 - F.8). Point estimates on wages in salaried positions are positive, but is only significant in Year 2. Finally, we see slight increases in self-reported happiness, but only among the Year 1 treated sample.

In summary, while we cannot rule out potentially large long-run effects on revenues, we find no significant evidence that the loan permanently alters farmers' timing of sales or a variety of other household-level economic outcomes. Consistent with this, we find no long-run effects on local market prices (though effects are in the same direction as the short-run effects, but are much muted; see Table F.9). We therefore find little evidence that a one-time injection of credit can permanently ameliorate the underlying constraints limiting arbitrage.

4.6 Temptation and kin tax

To test whether self-control issues or social pressure to share with others limits storage, we test the impact of laminated tags that brand the maize as committed to OAF. Estimates are shown in Table 13. We find no significant difference in inventories, revenues, or consumption for individuals who receive only the tags (without the loan), and point estimates are small. Therefore, the tags do not appear to have any effect on storage behavior. However, this may simply be because tags are a weak form of commitment, either to one's self or to others.

---

[35]This could occur if revenues from the loan relaxed credit constraints that previously restricted farmers' ability to invest in inputs. Alternatively, if the loan led to long-run improvements in the price farmers receive for their crops, this increased output price could increase incentives to invest in production-enhancing inputs. An improved price could be attained either in the lean season, if the farmer in question himself stores, or at harvest time, if other farmers are arbitraging and producing lower overall season price fluctuations (though note in Tables F.1 and F.9 we see no evidence of such long-run shifts in either sales timing or prices).

4.7 Savings one's way out of the credit constraint

How long might it take for a farmer to "save his way out" of this credit constraint? In Appendix H, we present various estimates suggesting that it would take the farmer 3-6 years to self-finance the loan, if he were to save the full returns from his investment, but 34 years if he saved at a more standard savings rate of 10%. Therefore, low savings rates are important to understanding why credit constraints persist in the presence of high return, divisible investment opportunities.

In order to test the importance of savings constraints, we examine the impact of the lockbox, as well as its interaction with the loan. Table H.1 presents these results. We observe no significant effects of the lockbox on inventories, revenues, or consumption in the overall sample. Interestingly, when interacted with the loan, we see that receiving the lockbox alone is associated with significantly *lower* inventories; perhaps the lockbox serves as a substitute savings mechanism, rather than grain. However, receiving both the lockbox and the loan is associated with a reversal of this pattern. We see no such heterogeneity on revenues. Interestingly, the point estimates on consumption are negative (though not significant) for the lockbox and loan when received separately; however, the interaction of the two is positive (and significant, at 95%), canceling out this effect.

5 General equilibrium effects

Because the loan resulted in greater storage, which shifted supply across time, and given the high transport costs common in the region, we might expect this intervention to affect the trajectory of local market prices. By shifting sales out of a relative period of abundance, we would expect the loan to result in higher prices immediately following harvest. Conversely, by shifting sales into a period of relative scarcity, we would expect the loan to result in lower prices later in the lean season. Note, however, that these effects will only be discernible if (1) the treatment generates a substantial shock to the local supply of maize; and (2) local markets are somewhat isolated, such that local prices are at least partially determined by local supply.

5.1 Market level effects

To understand the effect of our loan intervention on local maize prices, we identified 52 local market points spread throughout our study area where OAF staff indicated their clients typically bought and sold maize, and our enumerators tracked monthly maize prices at these market points. We then match these market points to the OAF sublocation in which they fall.

A note on the matching process: "sublocation" is an OAF administrative unit that is well-defined in terms of client composition (i.e. OAF divides its farmer groups into sublocations based on geographic proximity), but which is less well-defined in terms of precise geographic boundaries (that is, no shape file of sublocations exists). Given this, we use GPS data on both the market location and the location of farmers in our study sample to calculate the "most likely" sublocation, based on the designated sublocation to which the modal study farmer falling within a 3km radius belongs.[36] This procedure, include the radius to be used, was pre-specified. As was also pre-specified, we test robustness to alternative radii of 1km and 5km.

We then utilize the sublocation-level randomization in treatment intensity to identify market-level effects of our intervention, estimating Equation 4 and clustering standard errors at the sublocation level. Regression results are shown in Table 6 and plotted non-parametrically in Figure 6. In each year, we explore the price changes from the period following loan disbursal (November in Year 1, December in Year 2) until the beginning of the subsequent harvest (August in both years). In Figure 6, which presents the pooled data, we see prices in high-intensity markets on average start out almost 4% higher in the immediate post-harvest months. As the season progresses, prices in high-density markets begin to converge and then dip below those low-density markets, ending almost 2% lower in high-density areas compared to low-density. Table 6 presents these results according to the empirical specifically outlined in Section 3. In line with the graphic results visible in Figure 6, here we see the interaction term on "Hi" treatment intensity is positive (and significant at 5%), while the interaction term between the monthly time trend and the high intensity dummy is negative (though not significant). Columns 4-5 display robustness to alternative radii; we find

---

[36]Because we draw twice the sample from high-intensity areas compared to low (in accordance with our randomized intensity), we weight the low-intensity observations by two to generate a pool reflective of the true underlying OAF population. From this pool, we identify the modal farmer sublocation.

similar point estimates.

The overall picture painted by the market price data is consistent with the individual-level results presented above. Price effects are most pronounced (and statistically significant) early on in the season. This is when we observe the largest and most concentrated shock to the supply on the market (note in Table 2 that the greatest shift in inventories is seen in Round 1). Sensibly, treatment effects are most concentrated around the time of the loan disbursal, which represents a common shock affecting all those taking out the loan; this produces a simultaneous inward shift in supply in the post-harvest period. In contrast, the release of this grain onto the market in the lean period appears to happen with more diffuse timing among those the treatment group (as can be seen in Figure 5, in which we note a gradual reduction in the treatment-control gap in inventories, rather than the sharp drop we would expect if all treated individuals sold at the same time). Anecdotally, farmers report that the timing of sales is often driven by idiosyncratic shocks to the household's need for cash, such as the illness of a family member, which may explain the observed heterogeneity in timing in which the treatment group releases its stores. Perhaps as a result of these more diffuse treatment effects in the lean season, price effects are smaller and measured with larger standard errors in the second half of the year.

Are the size of these observed price effects plausible? A back-the-envelope calibration exercise suggests yes. OAF works with about 30% of farmers in the region. Of these farmers, 80% were enrolled in the study in high density areas, while 40% were enrolled in low-density areas. About 58% of those enrolled received the loan offer [37] Together, this implies that about 14% of the population was offered treatment in high-intensity sublocations and 7% in low-intensity areas, such that the treatment was offered to 7 percentage points more of the population in high-density areas. Table 2 suggests that treated individuals experienced average increases in inventory (i.e. inward supply shifts) of 24.5%. Taken together, this suggests a contraction in total quantity available in the high-density markets by 1.7%. Experiments conducted in the same region in Kenya suggest an average demand elasticity of -1.1 (Bergquist, 2017). This would imply that we should expect

---

[37]In Year 1, 66% of the sample received the loan offer (1/3 received the offer in October, 1/3 received the loan offer in January, and 1/3 served as control). In Year 2, 50% of the sample received the loan offer (1/2 received the offer in November and 1/2 served as control). In this calibration exercise, we use the average of the two years' rates.

to see an overall price increase of 1.5%. In the period immediately following harvest, when the inventory effects are most concentrated – during which time inventories are 47.7% higher among treatment individuals – we see an inward supply shift of 3.3%, and should therefore expect to see a 3.0% increase in price.[38] This is quite close to what we observe in Figure 6. We see a jump in price of about 2.5% during this period,[39] which then peters out to a slightly negative (though not significant) effect towards the end of the season.

5.2   Robustness checks

We check the robustness of the regression results to functional form assumptions. Table I.1 presents a binary version of Equation 4, replacing $month_t$ with an indicator $lean_t$ for being in the lean season (defined as April-August) and the interaction term with $lean_t * H_s$. Results suggest similar significant increases in price post-harvest in high-intensity markets. The lean season interaction term suggests that prices in high-intensity markets are lower overall in the lean season, although the point estimate on the interaction term is only slightly larger in absolute value than the the main H_s treatment coefficient, such that the combined effect of treatment in the lean season is to lower prices in high-intensity markets only slightly below those in low-intensity overall. Comparing these effects to Figure 6, we observe this is because at the beginning of the lean season prices are still higher in high intensity markets, with a cross-over mid-lean season as prices in high-intensity markets drop below those low-intensity markets. However, the 1km and 5km specifications shown in the right panel in Figure 6 shows suggest that this crossover occurs closer to the transition from the harvest to lean season; therefore the 1km and 5km specification of the binary specification, shown in Columns 4-5 of Table I.1, estimate a more substantial decrease in price for the full lean season.

We also check the robustness of these results to a more continuous measure of treatment at

---

[38]Note this exercise assumes no trade across sublocations. On the opposite extreme, the case of perfect market integration with zero transaction costs would imply perfect smoothing of any localized supply shock, and we would therefore observe no change in price. We therefore view the range of 0-3% as the extreme bounds of what price changes we should expect to observe.

[39]We measure shifts in post-harvest inventories in Round 1 of the survey, which conducted roughly January-February for the average respondent. We therefore estimate the change in price change in January-February from Table 6 to be $3.97 + 2.5 * (-0.57) = 2.5$.

the market-level, following the technique described in Miguel and Kremer (2004). We construct an estimate of the ratio of total treated farmers to the total farmers in our sample within a 3km radius around each market.[40] We re-estimate an equation identical to Equation 4 with H_s replaced with $ratio_m$, the aforementioned ratio. Results are presented in Table I.2. We also present non-parametric estimates of this specification in Figure I.1, displaying average prices in markets with above- vs. below-median ratios. While results are somewhat less precisely estimated in this specification, the broad patterns remain consistent: prices are higher in the post-harvest period and lower in the lean period in markets with a greater proportion of treated individuals in the area.

We also check robustness to small cluster standard error adjustments. These market-level price results rely on the treatment saturation randomization being conducted at the sublocation level (a higher level than the group-level randomization employed in the individual-level results). While we cluster standard errors at the sublocation level,[41] one might be concerned due to the small number of sublocations – of which we have 17 – that asymptotic properties may not apply to our market-level analyses and that our standard errors may therefore be understated. We run several robustness checks to address these small sample concerns. In Appendix I, we use a nonparametric randomization inference approach employed by Bloom et al. (2013) and Cohen and Dupas (2010) to draw causal inferences in the presence of small samples. Results using these alternative approaches are broadly consistent with those from the primary specifications. We also check the robustness of our results by conducting the wild bootstrap procedure proposed by Cameron et al. (2008). While we do see some decrease in statistical precision, these adjustments are small. Finally, to ensure that results are not sensitive to a single outlier sublocation, we drop each sublocation one-by-one and re-run our analysis; the pattern observed in the full data is generally robust to this outlier analysis. See Appendix I for further details.

---

[40]Because we draw twice the sample from high-intensity areas compared to low (in accordance with our randomized intensity), for the total farmer count, we weight the low-intensity observations by two to generate a count reflective of the true underlying OAF population.

[41]For all analyses in this paper, we cluster our standard errors at the level of randomization. For the individual results shown in Section 4, this is at the group level. For the results presented in this section, which relying on the sublocation-level randomized saturation, we cluster at the sublocation level.

5.3 Related Outcomes

We also check whether treatment intensity affected other outcomes of interest related to market price. First, we check whether treatment effects can be seen in farmgate prices (see Table I.5). We see similar patterns in these prices as well. We also explore whether trader movement responds to treatment. We see some evidence that fewer traders enter high-intensity treated markets in the immediate post-harvest period in Year 2 (see Table I.6), a sensible demand response to the increase in price observed during a time when traders are typically purchasing.[42]

6 Individual results with spillovers

Mass storage appears to raise prices at harvest time and lower price in the lean season, thereby smoothing out seasonal price fluctuations. What effect does this have on the individual profitability of the loan, which is designed to help farmers to take advantage of these price variations? That is, how do the individual-level returns to arbitrage vary with the number of arbitrageurs?[43]

To answer this question, we revisit the individual results, re-estimating them to account for the variation in treatment density across sublocations. Table 7 and Figure 7 display how our main outcomes respond in high versus low density areas for treated and control individuals. We find that inventory treatment effects do not significantly differ as a function of treatment intensity for the pooled treatment.

Effects on net revenues, however, paint a different picture. Treatment effects in low-intensity areas are much larger – roughly double — those estimated in the pooled specification. This is because most of the revenue effects seen in the pooled specification are concentrated among treated individuals in low-intensity sublocations. In contrast, revenue effects for treated individuals in

---

[42] This, along with the overall weaker treatment intensity in Year 2, may contribute to the smaller price effects observed in Year 2. In terms of weaker treatment intensity, note that the sample size in Year 2 is only about 65% that of Year 1. As a result, the intensity in Year 2 is only about 65% what it was in Year 1. Note that the point estimate on "Hi" in column 2 (Y2) of Table 6 is almost exactly 65% of the coefficient on column 1 (Y1). The coefficient on "Hi Intensity * Month" in column 2 (Y2) is close to (a bit more than) 65% of the coefficient on column 1 (Y1).

[43] Local market effects may not be the only channel through which treatment density affected individual-level results. For example, sharing of maize or informal lending between households could also be affected by the density of loan recipients. Appendix K explores these alternative channels and presents evidence suggesting that the individual-level spillover results are most consistent with spillovers through effects on local markets. However, we cannot rule out that other mechanisms could also be at play.

high-intensity sublocations are substantially lower (and, in fact, are statistically indistinguishable from zero in the pooled results presented Column 6 of Table 7).[44],[45] Therefore, while individuals in both high and low-intensity sublocations store significantly more as a result of treatment, only treated individuals in low-intensity sublocations earn significantly higher revenues. As with earlier estimates, estimates for consumption remain relatively imprecisely estimated.[46],[47]

Why might loan profitability be lower in high treatment density areas? Intuitively, arbitrage – the exploitation of price differentials – is most profitable to an individual when she is the only one arbitraging. As others begin to arbitrage as well, general equilibrium effects drive down these differentials and therefore diminish the direct returns to arbitrage. Conversely, for those who do *not* engage in arbitrage, these spillovers may be positive. Though the timing of their sales will not change, they may benefit from relatively higher sale prices at harvest-time and relatively lower purchase prices during the lean season.

We see some evidence of these positive spillovers to control group revenues in high-intensity treatment areas (see middle panel of Figure 7 and the estimate on the Hi dummy in Column 6 of Table 7). However, it should be noted that this effect is measured with considerable noise and and thus remains more speculative.[48] Given the diffuse nature of spillover effects, it is perhaps unsurprising that identifying these small effects with statistical precision is challenging.[49] However, they are suggestive of important distributional dynamics for welfare, which we explore below.

---

[44]Table 7 displays "p-val T+TH=0," which indicates the joint significance of $\beta_1 + \beta_3$ from Equation 5; this represents the full effect of treatment for individuals in high-intensity sublocations.

[45]While the interaction term "Treat*Hi" is only significant at traditional levels in Year 1, we attribute at least some of the weakened Year 2 interaction term to the lower treatment intensity in Year 2. Recall that the sample size in Year 2 is only about 65% that of Year 1. As a result, the intensity in Year 2 is only about 65% what it was in Year 1. If we scale the coefficient on "Treat*Hi" in Year 2 (column 2) to account for this difference (i.e. divide by 0.65), we get an estimate much closer to the Year 1 estimate. In addition, any trader movement that dampened Year 2 market-level effects may have further contributed to this weaker Year 2 effect.

[46]Interestingly, they are strongly positive for treated individuals in the high-intensity areas in Year 2. However, because there is no clear pattern across years, we avoid speculating or over-interpreting this fragile result.

[47]Because the consumption measure includes expenditure on maize, in Appendix E.1 we also estimate effects on consumption excluding maize and consumption excluding all food. Results are similar using these measures.

[48]And even goes in the opposite direction in the Year 2 results alone; see Column 5 of Table 7.

[49]Simple power calculations suggest that if the point estimate of 165 is the true effect, a sample size of 218,668 – more than 32 times our current sample size – would be necessary to detect this effect with 95% confidence.

6.1 Discussion

The randomized saturation design allows us to capture how both direct and indirect treatment effects vary with saturation level. Table 8 breaks down the distribution of welfare gains from the loan, based on saturation rate and revenue effects drawn from the pooled results. While this exercise takes all point estimates as given, note that some are less precisely measured than others.[50] As a result, there are likely large standard errors around some of the figures presented in Table 8. This exercise should therefore be interpreted as an illustration of how general equilibrium effects can shape the distribution of welfare gains in isolated markets, rather than precise quantitative estimates.

In the first row, we present the direct gains per household, representing the increase in revenues driven by treatment for those who are treated (specifically calculated as the coefficient on the "Treat" dummy in low saturation areas and as the coefficient on the "Treat" dummy plus the coefficient on the "Treat*Hi" interaction term in high saturation areas). As discussed above, we see that the direct treatment effects are greater for those in low saturation sublocations, where treated individuals are closer to "being the only one arbitraging," than in high saturation areas.

The second row presents the indirect gains per household. This is estimated as zero in low saturation areas and as the coefficient on "Hi" in high saturation areas.[51] We see in row 3 that, in the high saturation areas, the indirect gains are 58% the size of the direct gains. When we account for the much larger size of the total population relative to that of just the direct beneficiaries (presented in rows 5 and 4 respectively), we find that the total size of the indirect gains swamp that of the direct gains in high saturation areas (rows 7 and 6 respectively).

These findings have two implications. First, the total gains from the intervention (presented in row 7) are much higher in high saturation areas than they are in low saturation areas. While the direct gains to the treatment group are lower in areas of high saturation, the small per-household indirect gains observed in these areas accrue to a large number of untreated individuals, resulting in an overall increase in total gains (note that although there is a large degree of imprecision in our

---

[50]For example, the point estimate on "Treat*Hi" is not quite significant at traditional levels, while the point estimate on "Hi" is measured with large noise.

[51]Though note that low-intensity treatment areas may also experience GE effects which we are unable to detect. We are only able to detect *relative* differences in prices across low- and high-intensity areas.

estimate of the indirect gains, the qualitative result that higher saturation produces larger gains than low saturation holds even at indirect gains as low as 38Ksh/individual, less than $0.5 USD).[52] High saturation offers greater relaxation of a barrier to intertemporal trade (credit constraints) and thereby produces larger aggregate gains.

Second, the distribution of gains shifts in the presence of general equilibrium effects. While in low saturation areas all of the gains appear to come from direct gains, in high saturation areas, 81% of the total gains are indirect gains (row 9).[53] General equilibrium effects therefore more evenly distribute gains across the entire population, reducing the proportion of the gains that direct beneficiaries exclusively receive and increasing the share enjoyed by the full population.

This redistribution of gains has implications for private sector investment in arbitrage. Row 10 presents the per-household private gains accruing to arbitragers, as estimated by the coefficient on the "Treat" indicator in low saturation areas and by the coefficient on the "Treat" dummy plus the coefficient on the "Treat*Hi" interaction term plus the coefficient on the "Hi" interaction term in high saturation areas. This represents the per-household gains accruing to treated farmers in our sample, under each level of saturation. It also represents the most that private sector banks or other financial institutions could hope to extract from each farmer to whom they might provide loans for storage. Row 11 presents the total private gains, multiplying the per-household gains by the number of treated individuals. Despite the fact that high saturation areas have two times the number of treated farmers, the total private gains are still lower in these areas compared to low saturation areas.

These calculations suggest that private sector financial institutions may face incentives that result in the under-provision of finance for arbitrage. Although overall social gains are higher at greater levels of saturation (row 8), because much of these gains are indirect, private sector institutions will not be able to capture them. For private sector institutions, the available gains

---

[52]Also contributing is the fact that although the direct benefits/household are only a quarter of the size in high areas, there are twice the number of beneficiaries, which makes up some of the gap in terms of total direct gains.

[53]It is possible that there are general equilibrium effects – and therefore indirect gains – occurring in the low saturation areas that we simply cannot detect in the absence of a pure control group. If this is the case, it would mean that our current estimates underestimate the total gains, as well as the percentage of gains coming from indirect gains, in low saturation areas. However, it would also mean that we are underestimating these figures in the high intensity areas as well.

for capture are actually lower at high levels of saturation (row 11). Row 12 attempts to quantify this disincentive. At low levels of saturation, private sector institutions could fully internalize all gains, capturing up to 100% of the total revenue increases generated by the product (under our assumption of no indirect gains in the low saturation case). However, at high saturation rates, only 31% of the total gains are private. Financial institutions therefore will fail to internalize 69% of the gains at these higher saturation levels, which will likely result in under-provision of financial products, compared to the socially optimal level.

7   Conclusion

Large and regular increases in the price of maize between the harvest and the lean season offer farmers substantial arbitrage opportunities. However, smallholder farmers appear unable to arbitrage these price fluctuations due to high harvest-time expenditure needs and an inability to access credit markets, necessitating high harvest-time sales of maize.

We study the effect of offering Kenyan maize farmers access to a loan during the harvest period. We find that access to this perhaps counter-intuitively timed credit "frees up" farmers to use storage to arbitrage these price movements. Farmers offered the loan shift maize purchases into the period of low prices, put more maize in storage, and sell maize at higher prices later in the season, increasing farm revenue. Using experimentally-induced variation in the density of treatment farmers across locations, we document that this change in storage and marketing behavior aggregated across treatment farmers also affects local maize prices: post harvest prices are significantly higher in high-density areas, consistent with more supply having been taken off the market in that period, and are lower later in the season (though not significantly so). These general equilibrium effects feed back to our profitability estimates, with treatment farmers in low-density areas – where price differentials were higher and thus arbitrage opportunities greater – differentially benefiting.

The findings make a number of contributions. First, our results are among the few experimental results to find a positive and significant effect of microcredit on the profits of microenterprises (farms in our case). This is also to our knowledge one of the first experimental studies to directly account for general equilibrium effects in this literature. At least in our particular setting, failing to

account for these GE effects substantially alters the conclusions drawn about the average benefits of improved credit access.

This has methodological implications for a broader set of interventions that may shift local supply – such as agricultural technologies that increase local food supply or vocational training programs that increase local skilled labor supply – in the presence of thin or imperfectly integrated markets. Our results suggest that, when implemented in rural or fragmented markets, these interventions may lead local prices to respond substantially enough to alter the profitability of these interventions for direct beneficiaries and to impact the welfare of non-beneficiaries. Explicit attention to GE effects in future evaluations is likely warranted.

Finally, we show how the absence of financial intermediation can be doubly painful for poor households in rural areas. Lack of access to formal credit causes households to turn to much more expensive ways of moving consumption around in time, and aggregated across households this behavior generates a large scale price phenomenon that further lowers farm income and increases what most households must pay for food. The results suggest that expanding access to affordable credit could reduce this price variability and thus have benefits for recipient and non-recipient households alike. Welfare estimates in our setting suggest that a large portion of the benefits of expanded loan access could accrue indirectly to non-borrowers. Under such a distribution of welfare gains, private sector financial institutions may undersupply credit relative to the social optimum, raising the possibility that public credit programs could raise aggregate welfare.

What our results do not address is why wealthy local actors – e.g. large-scale private traders – have not stepped in to bid away these arbitrage opportunities. Traders do exist in the area and can commonly be found in local markets. In a panel survey of local traders in the area, we record data on the timing of their marketing activities and storage behavior. But we find little evidence of long-run storage. When asked to explain this limited storage, many traders report being able to make even higher total profits by engaging in spatial arbitrage across markets (relative to temporal arbitrage). Nevertheless, this does not explain why the scale or number of traders engaging in both spatial and intertemporal arbitrage has not expanded; imperfect competition among traders may play a role (Bergquist, 2017).

References

Acemoglu, Daron, "Theory, general equilibrium and political economy in development economics," *Journal of Economic Perspectives*, 2010, *24* (3), 17–32.

Aggarwal, Shilpa, Eilin Francis, and Jonathan Robinson, "Grain Today, Gain Tomorrow: Evidence from a Storage Experiment with Savings Clubs in Kenya," *Working Paper*, 2017.

Aker, Jenny C, "Rainfall shocks, markets and food crises: the effect of drought on grain markets in Niger," *Center for Global Development, working paper*, 2012.

Ambler, Kate, Alan De Brauw, and Susan Godlonton, "Measuring postharvest losses at the farm level in Malawi," *Australian Journal of Agricultural and Resource Economics*, 2018, *62*.

Anderson, Michael L, "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484).

Andreoni, James and Charles Sprenger, "Estimating Time Preferences from Convex Budgets," *American Economic Review*, 2012, *102* (7), 3333–56.

Angelucci, Manuela, Dean Karlan, and Jonathan Zinman, "Win some lose some? Evidence from a randomized microcredit program placement experiment by Compartamos Banco," Technical Report, National Bureau of Economic Research 2013.

Attanasio, Orazio, Britta Augsburg, Ralph De Haas, Emla Fitzsimons, and Heike Harmgart, "Group lending or individual lending? Evidence from a randomised field experiment in Mongolia," 2011.

Baland, Jean-Marie, Catherine Guirkinger, and Charlotte Mali, "Pretending to be poor: Borrowing to escape forced solidarity in Cameroon," *Economic Development and Cultural Change*, 2011, *60* (1), 1–16.

Banerjee, Abhijit V and Andrew F Newman, "Occupational choice and the process of development," *Journal of political economy*, 1993, pp. 274–298.

_ **and Esther Duflo**, "Giving credit where it is due," *The Journal of Economic Perspectives*, 2010, *24* (3), 61–79.

Banerjee, Abhijit Vinayak, "Microcredit Under the Microscope: What Have We Learned in the Past Two Decades, and What Do We Need to Know?," *Annual Review of Economics*, 2013, (0).

Banerjee, Ahbijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan, "The Miracle of Microfinance?: Evidence from a Randomized Evaluation," *working paper, MIT*, 2013.

Barrett, Christopher, "Displaced distortions: Financial market failures and seemingly inefficient resource allocation in low-income rural communities," in Erwin Bulte and Ruerd Ruben, eds., *Development Economics Between Markets and Institutions: Incentives for growth, food security and sustainable use of the environment*, Wageningen Academic Publishers, 2007.

__ **and Paul Dorosh**, "Farmers' Welfare and Changing Food Prices: Nonparametric Evidence From Rice In Madagascar," *American Journal of Agricultural Economics*, 1996, *78* (3).

Basu, Karna and Maisy Wong, "Evaluating Seasonal Food Security Programs in East Indonesia," *working paper*, 2012.

Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry, "Self-Selection into Credit Markets: Evidence from Agriculture in Mali," *Working paper*, 2014.

Berge, Lars Ivar, Kjetil Bjorvatn, and Bertil Tungodden, "Human and financial capital for microenterprise development: Evidence from a field and lab experiment," *NHH Dept. of Economics Discussion Paper*, 2011, (1).

Bergquist, Lauren Falcao, "Pass-through, Competition, and Entry in Agricultural Markets: Experimental Evidence from Kenya," *Working Paper*, 2017.

Blattman, Christopher, Nathan Fiala, and Sebastian Martinez, "Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda," *Quarterly Journal of Economics*, 2014, *129* (2), 697–752.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, "Does management matter? Evidence from India," *The Quarterly Journal of Economics*, 2013, *128* (1), 1–51.

Breza, Emily and Cynthia Kinnan, "Measuring the Equilibrium Impacts of Credit: Evidence from the Indian Microfinance Crissi," *Working paper*, 2018.

Bruhn, Miriam and David McKenzie, "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 2009, pp. 200–232.

__ **, Dean S Karlan, and Antoinette Schoar**, "The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in mexico," *Yale University Economic Growth Center Discussion Paper*, 2012, (1010).

Brune, Lasse, Xavier Giné, Jessica Goldberg, and Dean Yang, "Commitments to save: A field experiment in rural malawi," *University of Michigan, May (mimeograph)*, 2011.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller, "Bootstrap-based improvements for inference with clustered errors," *The Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Casey, Katherine, Rachel Glennerster, and Edward Miguel, "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan*," *The Quarterly Journal of Economics*, 2012, *127* (4), 1755–1812.

Cohen, Jessica and Pascaline Dupas, "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment," *Quarterly Journal of Economics*, 2010.

Conley, Timothy G, "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45.

Crepon, Bruno, Florencia Devoto, Esther Duflo, and William Pariente, "Impact of microcredit in rural areas of Morocco: Evidence from a Randomized Evaluation," *working paper, MIT*, 2011.

Dillion, Brian, "Selling Crops Early to Pay for School: A Large-scale Natural Experiment in Malawi," *Working Paper*, 2017.

Dupas, Pascaline and Jonathan Robinson, "Why Don't the Poor Save More? Evidence from Health Savings Experiments," *American Economic Review, forthcoming*, 2013.

Fafchamps, Marcel, "Cash crop production, food price volatility, and rural market integration in the third world," *American Journal of Agricultural Economics*, 1992, *74* (1).

__ , **David McKenzie, Simon Quinn, and Christopher Woodruff**, "Microenterprise Growth and the Flypaper Effect: Evidence from a Randomized Experiment in Ghana," *Journal of Development Economics*, 2013.

Field, Erica, Rohini Pande, John Papp, and Natalia Rigol, "Does the Classic Microfinance Model Discourage Entrepreneurship Among the Poor? Experimental Evidence from India," *American Economic Review*, 2012.

Fink, Gunther, Kelsey Jack, and Felix Masiye, "Seasonal credit constraints and agricultural labor supply: Evidence from Zambia," *NBER Working Paper*, 2014, (20218).

Galor, Oded and Joseph Zeira, "Income distribution and macroeconomics," *The review of economic studies*, 1993, *60* (1), 35–52.

Jonathan, Luc Christiaensen Kaminski and Christopher L Gilbert, "The End of Seasonality? New Insights from Sub-Saharan Africa," *World Bank Policy Research Working Paper*, 2014, (6907).

Kaboski, Joseph P and Robert M Townsend, "The impact of credit on village economies," *American economic journal. Applied economics*, 2012, *4* (2), 98.

Kaminski, Jonathan and Luc Christiaensen, "Post-Harvest Loss in Sub-Saharan Africa - What Do Farmers Say," *ELSEVIER*, 2014.

Karlan, D., J. Morduch, and S. Mullainathan, "Take up: Why microfinance take-up rates are low and why it matters," Technical Report, Financial Access Initiative 2010.

Karlan, Dean and Jonathan Morduch, "Access to Finance," *Handbook of Development Economics, Volume 5*, 2009, (Chapter 2).

__ **and Jonathan Zinman**, "Microcredit in theory and practice: using randomized credit scoring for impact evaluation," *Science*, 2011, *332* (6035), 1278–1284.

__ , **Ryan Knight, and Christopher Udry**, "Hoping to win, expected to lose: Theory and lessons on micro enterprise development," Technical Report, National Bureau of Economic Research 2012.

Lee, David S, "Training, wages, and sample selection: Estimating sharp bounds on treatment effects," *The Review of Economic Studies*, 2009, *76* (3), 1071–1102.

McCloskey, Donald and John Nash, "Corn at interest: The extent and cost of grain storage in Medieval Englang," *American Economic Review*, 1984, *74* (1).

McKenzie, David, "Beyond baseline and follow-up: the case for more T in experiments," *Journal of Development Economics*, 2012.

_ **and Christopher Woodruff**, "Experimental evidence on returns to capital and access to finance in Mexico," *The World Bank Economic Review*, 2008, *22* (3), 457–482.

Meier, Stephan and Charles Sprenger, "Stability of time preferences," *IZA Discussion paper*, 2010.

Mel, Suresh De, David McKenzie, and Christopher Woodruff, "Returns to capital in microenterprises: evidence from a field experiment," *The Quarterly Journal of Economics*, 2008, *123* (4), 1329–1372.

_ , _ , **and** _ , "Are women more credit constrained? Experimental evidence on gender and microenterprise returns," *American Economic Journal: Applied Economics*, 2009, pp. 1–32.

Miguel, Edward and Michael Kremer, "Worms: identifying impacts on education and health in the presence of treatment externalities," *Econometrica*, 2004, *72* (1), 159–217.

Minten, Bart and Steven Kyle, "The effect of distance and road quality on food collection, marketing margins, and traders wages: evidence from the former Zaire," *Journal of Development Economics*, 1999, *60* (2).

Park, Albert, "Risk and household grain management in developing countries," *The Economic Journal*, 2006, *116* (514), 1088–1115.

Saha, Atanu and Janice Stroud, "A household model of on-farm storage under price risk," *American Journal of Agricultural Economics*, 1994, *76* (3), 522–534.

Stephens, Emma C. and Christopher B Barrett, "Incomplete credit markets and commodity marketing behaviour," *Journal of Agricultural Economics*, 2011, *62* (1), 1–24.

World Bank, "Malawi Poverty and Vulnerability Assessment: Investing in our Future," 2006.

Figure 1: **Monthly average maize prices**, shown at East African sites for which long-term data exist, 1994-2011. Data are from the Regional Agricultural Trade Intelligence Network, and prices are normalized such that the minimum monthly price = 100. Our study site in western Kenya is shown in green, and the blue squares represent an independent estimate of the months of the main harvest season in the given location. Price fluctuations for maize (corn) in the US are shown in the lower left for comparison

41
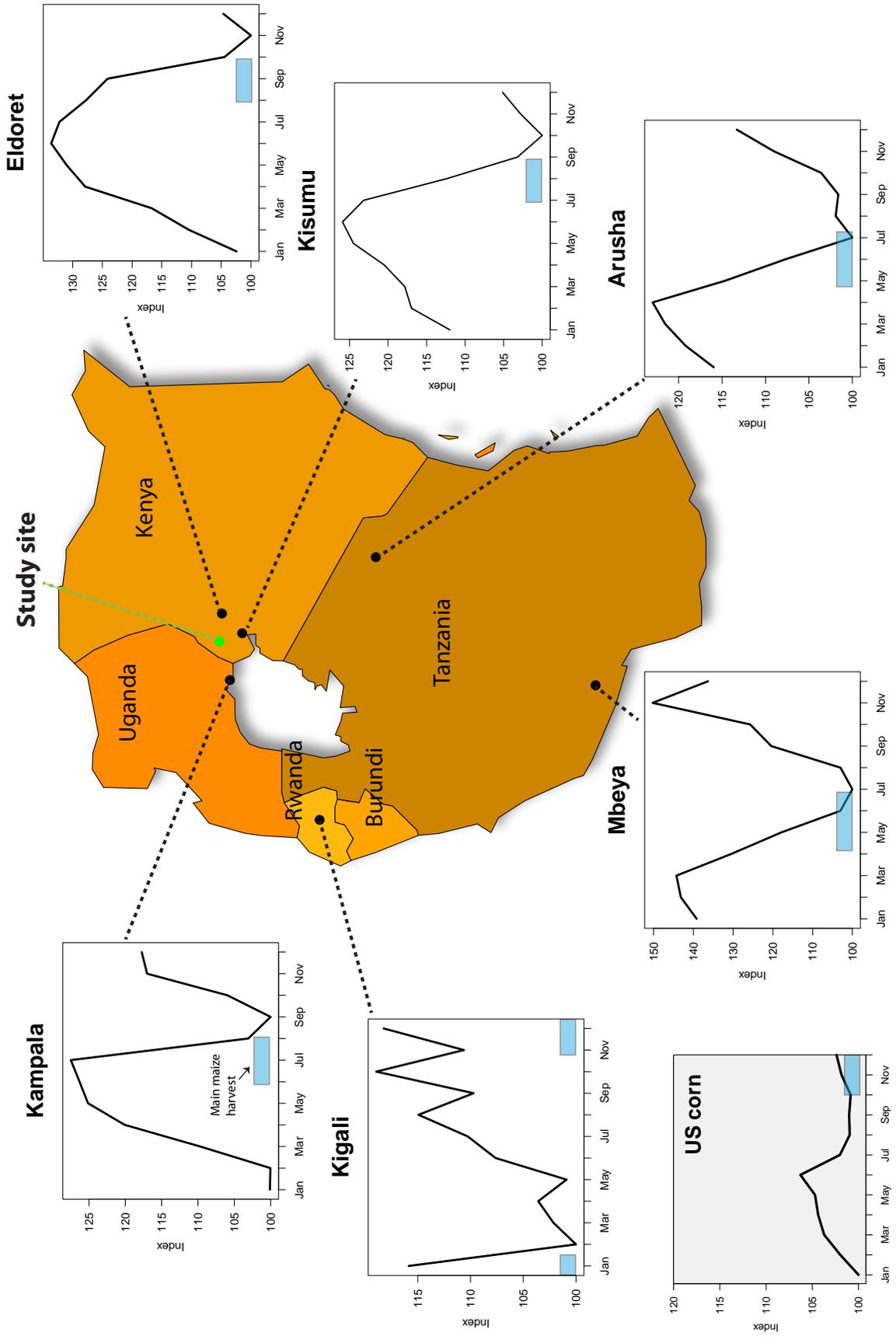
Figure 2: **Study design.** Randomization occurs at three levels. First, treatment intensity was randomized across 17 sublocations (top level, each box represents a sublocation). This randomization was held constant across the two years. Second, treatment was randomized at the group level within sublocations (second level, each box representing a group in a given sublocation). In Year 1, treatment groups were further divided into October and January loans. In Year 2, only one timing of the loan was offered (in November). Finally, in Year 1, there was a third level of randomization at the individual level, in which the tags and lockbox were cross-randomized (bottom level). In Year 2, no individual level treatments were offered. Total numbers of randomized units in each bin are given on the left.

42
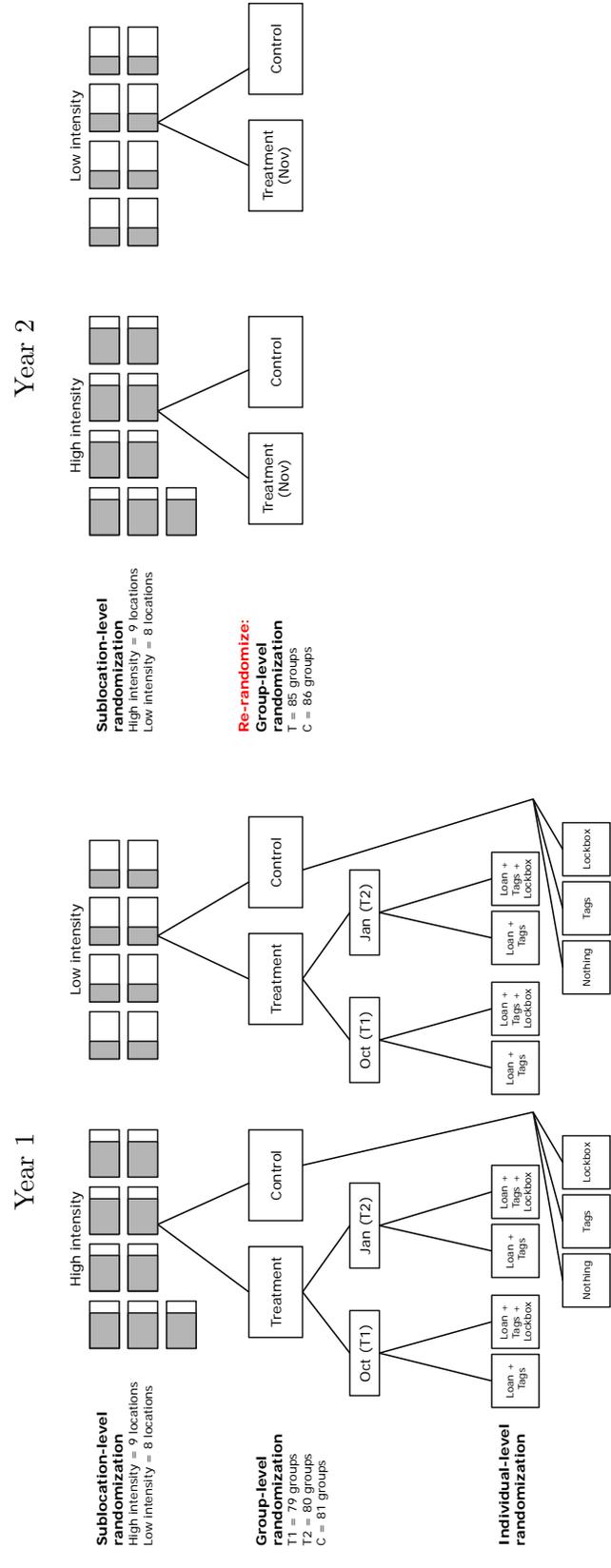
Figure 3: **Study timeline**. The timing of the interventions and data collection are show. Year 1 spanned 2012-2013, Year 2 spanned 2013-2014, and the long-run follow-up data collection occurred 2014-2015. The market survey, shown in red, ran from November 2012-December 2015. The baseline was run August-September 2012. Three rounds of household data were collected on a rolling basis in each year of the main study. A long-run follow-up survey was run September-December 2015. Light blue arrows show the timing of the loan announcement (immediately as harvest was ending), while dark blue arrows display the date of loan disbursal (October and January in Year 1, November in Year 2). Harvest time is highlighted in grey and occurs in September-October each year.
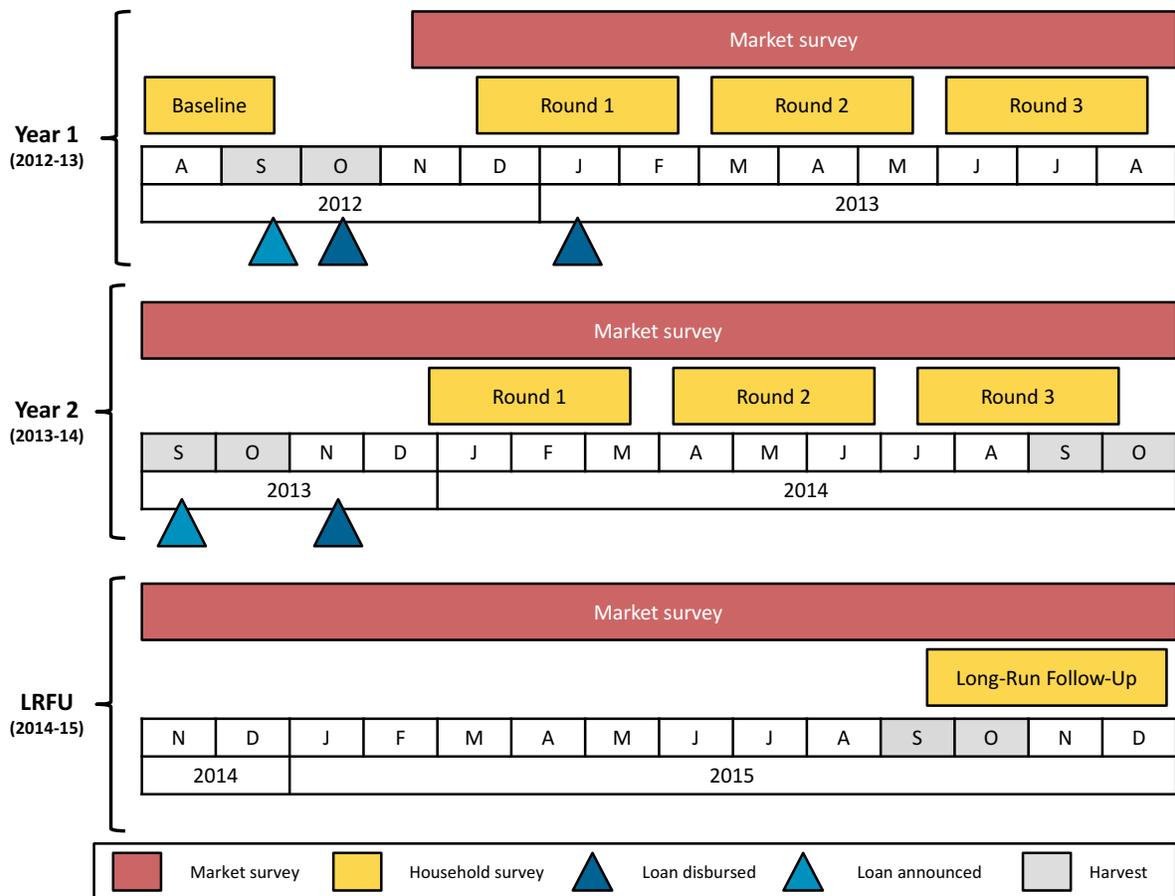
Figure 4: **Panel A: Maize price trends (pre-study period)**. Farmer-reported average monthly maize prices for the period 2007-2012, averaged over all farmers in our sample. Prices are in Kenyan shillings per goro (2.2kg). **Panel B: Maize price trends (study period & post-study period)**. Author-collected average monthly maize prices for the period 2012-2014 (study period) and 2014-2015 (post study period), averaged over all markets in our sample. Prices are in Kenyan shillings per goro (2.2kg).
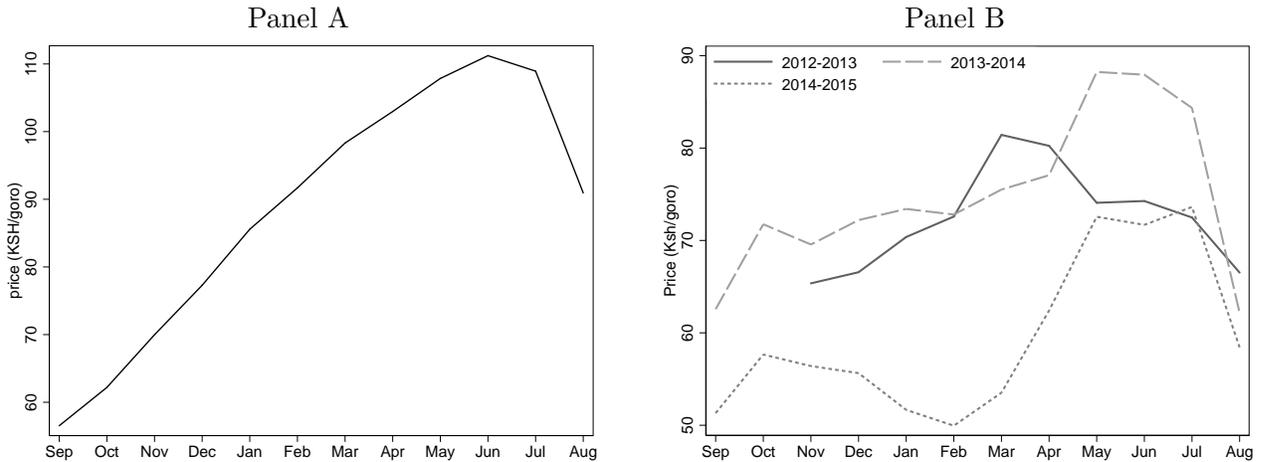
Figure 5: **Pooled Treatment effects**. The top row of plots shows how average inventories, net revenues, and log household consumption evolve from December to August in Y1 and Y2 (pooled) in the treatment group versus the control group, as estimated with fan regressions. The bottom row shows the difference between the treatment and control, with the bootstrapped 95% confidence interval shown in grey (100 replications drawing groups with replacement).

44
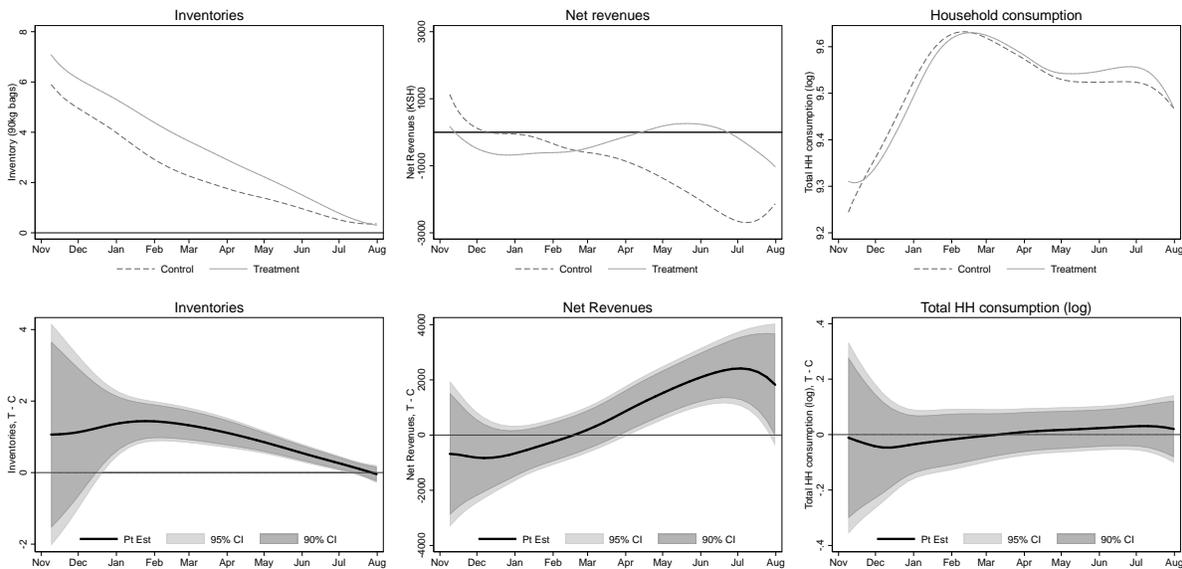
Figure 6: **Pooled market prices for maize as a function of local treatment intensity.**
Markets matched to treatment intensity using sublocation of the modal farmer within 3km of each
market. The left panel shows the average sales price in markets in high-intensity areas (solid line)
versus in low-intensity areas (dashed line) over the study period. The middle panel shows the aver-
age difference in log price between high- and low-intensity areas over time, with the bootstrapped
95% confidence interval shown in light grey and the 90% confidence interval shown in dark grey.
The right panel shows the robustness of results to alternative radii (1km, 3km, and 5km)
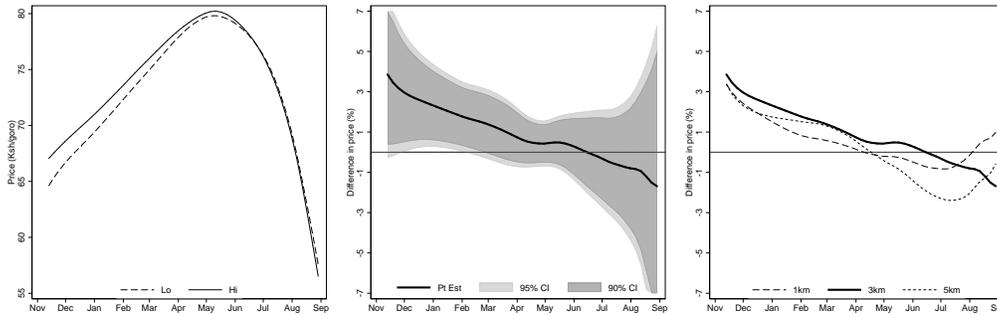
Figure 7: **Pooled Treatment effects by treatment intensity**. Average inventories, net rev-
enues, and log HH consumption over the study period in the treatment group versus the control
group, split apart by high intensity areas (orange lines) and low-intensity areas (black lines).
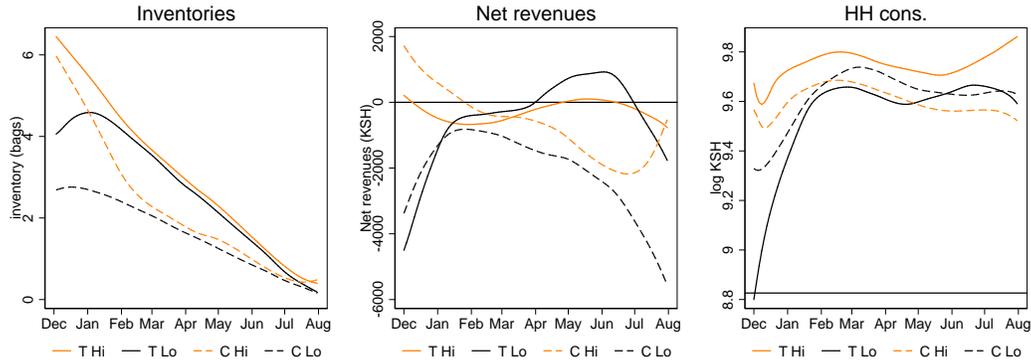
Table 1: **Summary statistics and balance among baseline covariates.** Balance table for the Y1 study (restricted to the Y1 sample, for which we have baseline characteristics. The first two columns give the means in each treatment arm. The 3rd column gives the total number of observations across the two groups. The last two columns give differences in means normalized by the Control sd, with the corresponding p-value on the test of equality.

Baseline characteristic	Treat	Control	Obs	T - C std diff	p-val
Male	0.30	0.33	1,589	-0.08	0.11
Number of adults	3.00	3.20	1,510	-0.09	0.06
Kids in school	3.00	3.07	1,589	-0.04	0.46
Finished primary	0.72	0.77	1,490	-0.13	0.02
Finished secondary	0.25	0.27	1,490	-0.04	0.46
Total cropland (acres)	2.44	2.40	1,512	0.01	0.79
Number of rooms in hhold	3.07	3.25	1,511	-0.05	0.17
Total school fees (1000 Ksh)	27.24	29.81	1,589	-0.06	0.18
Average monthly cons (Ksh)	14,970.86	15,371.38	1,437	-0.03	0.55
Avg monthly cons./cap (log Ksh)	7.97	7.96	1,434	0.02	0.72
Total cash savings (KSH)	5,157.40	8,021.50	1,572	-0.09	0.01
Total cash savings (trim)	4,731.62	5,389.84	1,572	-0.05	0.33
Has bank savings acct	0.42	0.43	1,589	-0.01	0.82
Taken bank loan	0.08	0.08	1,589	-0.02	0.73
Taken informal loan	0.24	0.25	1,589	-0.01	0.84
Liquid wealth	93,878.93	97,280.92	1,491	-0.03	0.55
Off-farm wages (Ksh)	3,916.82	3,797.48	1,589	0.01	0.85
Business profit (Ksh)	2,302.59	1,801.69	1,589	0.08	0.32
Avg %Δ price Sep-Jun	133.49	133.18	1,504	0.00	0.94
Expect 2011 LR harvest (bags)	9.36	9.03	1,511	0.02	0.67
Net revenue 2011	-3,303.69	-4,088.62	1,428	0.03	0.75
Net seller 2011	0.32	0.30	1,428	0.05	0.39
Autarkic 2011	0.07	0.06	1,589	0.03	0.51
% maize lost 2011	0.02	0.01	1,428	0.03	0.57
2012 LR harvest (bags)	11.18	11.03	1,484	0.02	0.74
Calculated interest correctly	0.71	0.73	1,580	-0.03	0.50
Digit span recall	4.57	4.58	1,504	-0.01	0.89
Maize giver	0.26	0.26	1,589	-0.00	0.99

"Liquid wealth" is the sum of cash savings and assets that could be easily sold (e.g. livestock). Off-farm wages and business profit refer to values over the previous month. Net revenue, net seller, and autarkic refer to the household's maize marketing position. "Maize giver" is whether the household reported giving away more maize in gifts than it received over the previous 3 months.

Table 2: **Inventory Effects, Individual Level.** Regressions include round-year fixed effects, with errors clustered at the group level.

	Y1		Y2		Pooled	
	(1) Overall	(2) By rd	(3) Overall	(4) By rd	(5) Overall	(6) By rd
Treat	0.52*** (0.16)		0.50*** (0.14)		0.53*** (0.12)	
Treat - R1		0.82*** (0.31)		1.21*** (0.24)		1.03*** (0.20)
Treat - R2		0.71*** (0.19)		0.24 (0.15)		0.52*** (0.12)
Treat - R3		0.06 (0.07)		0.04 (0.37)		0.07 (0.19)
Observations	3836	3836	2944	2944	6780	6780
Mean DV	2.67	2.67	1.68	1.68	2.16	2.16
R squared	0.35	0.35	0.18	0.19	0.29	0.30

Table 3: **Net Revenue Effects, Individual Level.** Regressions include round-year fixed effects, with errors clustered at the group level.

	Y1		Y2		Pooled	
	(1) Overall	(2) By rd	(3) Overall	(4) By rd	(5) Overall	(6) By rd
Treat	279.78 (292.16)		800.24** (330.63)		524.66** (220.25)	
Treat - R1		-1146.56*** (325.13)		-23.71 (478.41)		-608.68** (285.70)
Treat - R2		534.85 (485.80)		1917.28*** (532.81)		1170.71*** (359.84)
Treat - R3		1371.95*** (436.12)		520.76 (403.27)		985.79*** (302.09)
Observations	3795	3795	2935	2935	6730	6730
Mean DV	334.41	334.41	-3434.38	-3434.38	-1616.12	-1616.12
R squared	0.01	0.01	0.04	0.05	0.09	0.09

Table 4: **HH Consumption (log) Effects, Individual Level.** Regressions include round-year fixed effects, with errors clustered at the group level.

	Y1		Y2		Pooled	
	(1) Overall	(2) By rd	(3) Overall	(4) By rd	(5) Overall	(6) By rd
Treat	0.00		0.07*		0.04	
	(0.03)		(0.04)		(0.03)	
Treat - R1		-0.04		0.07		0.01
		(0.05)		(0.05)		(0.03)
Treat - R2		0.02		0.08*		0.05
		(0.04)		(0.05)		(0.03)
Treat - R3		0.03		0.06		0.04
		(0.05)		(0.05)		(0.03)
Observations	3792	3792	2944	2944	6736	6736
Mean DV	9.48	9.48	9.61	9.61	9.55	9.55
R squared	0.00	0.00	0.01	0.01	0.02	0.02

Table 5: **Net Sales and Effective Prices, Individual Level.** Columns 1-2 regressions on net sales (quantity sold minus quantity purchased) include round-year fixed effects, with errors clustered at the group level. Columns 3-4 include only one observation per individual (per year). Round fixed effects are omitted in these specifications in order to estimate the effect of treatment on prices paid and received, which change because of shifts in the timing of transactions; therefore round controls are not appropriate. "Effective purchase price" is constructed by the dividing the total value of all purchases over the full year (summed across rounds) by the total quantity of all purchases over the full year. "Effective sales price" is constructed similarly.

	Net Sales		Effective Price	
	Overall	By rd	Purchase	Sales
Treat	0.12*		-104.94***	131.70***
	(0.06)		(31.57)	(40.85)
Treat - R1		-0.26**		
		(0.10)		
Treat - R2		0.27***		
		(0.10)		
Treat - R3		0.29***		
		(0.09)		
Observations	6108	6108	2014	1428
Mean DV	-0.62	-0.62	3084.78	2809.76
R squared	0.16	0.16	0.01	0.01

Table 6: **Market prices for maize as a function of local treatment intensity.** "Hi" intensity is a dummy for a sublocation randomly assigned a high number of treatment groups. "Month" is a linear month time trend (beginning in Nov at 0 in each year). Standard errors are clustered at the sublocation level. Prices measured monthly following loan disbursal (Nov-Aug in Y1; Dec-Aug in Y2). Price normalized to 100 in Nov in low-intensity sublocations.

	Main Specification (3km)			Robustness (Pooled)	
	Y1	Y2	Pooled	1km	5km
Hi	4.41*	2.85	3.97**	2.79	3.77*
	(2.09)	(1.99)	(1.82)	(1.72)	(1.82)
Month	1.19***	1.22***	1.36***	1.33***	1.54***
	(0.36)	(0.38)	(0.35)	(0.34)	(0.29)
Hi Intensity * Month	-0.57	-0.48	-0.57	-0.52	-0.83**
	(0.42)	(0.46)	(0.39)	(0.39)	(0.37)
Observations	491	381	872	872	872
R squared	0.08	0.03	0.06	0.06	0.06

Table 7: **Inventory, Net Revenues, and HH Consumption (log) Effects, Accounting for Treatment Intensity.** Regressions include round-year fixed effects with errors clustered at the sublocation level. P-values on the test that the sum of the Treat and Treat*Hi equal zero are provided in the bottom rows of the table.

	Inventory			Net Revenues			Consumption		
	(1) Y1	(2) Y2	(3) Pooled	(4) Y1	(5) Y2	(6) Pooled	(7) Y1	(8) Y2	(9) Pooled
Treat	0.76***	0.55***	0.74***	1059.60**	1193.77	1101.39**	0.01	-0.05	-0.01
	(0.19)	(0.18)	(0.15)	(437.73)	(685.05)	(430.09)	(0.04)	(0.04)	(0.02)
Hi	0.12	-0.03	0.02	533.90	-152.60	164.94	-0.00	-0.08	-0.05
	(0.36)	(0.22)	(0.24)	(551.18)	(558.95)	(479.68)	(0.05)	(0.05)	(0.04)
Treat*Hi	-0.33	-0.07	-0.29	-1114.63*	-555.21	-816.77	-0.01	0.17***	0.07*
	(0.23)	(0.25)	(0.19)	(535.59)	(804.86)	(520.04)	(0.05)	(0.06)	(0.04)
Observations	3836	2944	6780	3795	2935	6730	3792	2944	6736
Mean DV	2.74	1.38	2.04	-253.51	-3620.40	-1980.02	9.47	9.65	9.56
R squared	0.35	0.18	0.29	0.01	0.04	0.09	0.00	0.02	0.03
p-val T+TH=0	0.01	0.02	0.01	0.86	0.15	0.41	0.97	0.01	0.08

Table 8: **Distribution of gains in the presence of general equilibrium effects** Calculations employ per-round point estimates on revenues β_1, β_2, and β_2 (estimated in Ksh) from Column 6 of Table 7 (multiplied by three to get the annual revenue gains). Direct gains per household (row 1) are calculated as the coefficient on the "Treat" dummy in low saturation areas and as the coefficient on the "Treat" dummy plus the coefficient on the "Treat*Hi" interaction term in high saturation areas). Indirect gains per household (row 2) are estimated as zero in low saturation areas and as the coefficient on "Hi" in high saturation areas. The total gains from the intervention (row 7) include the direct gains that accrue to borrowers (row 1) and the indirect gains generated by GE effects (row 2). In high saturation areas, 81% of the total gains are indirect gains (row 9). The private gains per household are estimated by the coefficient on the "Treat" indicator in low saturation areas and by the coefficient on the "Treat" dummy plus the coefficient on the "Treat*Hi" interaction term plus the coefficient on the "Hi" interaction term in high saturation areas. Row 11 presents the total private gains, multiplying the per-person gains by the number of treated individuals. Additional assumptions and calculation details are laid out in Appendix M. Note that while the private gains are greater at low saturation, the total gains are greater at high saturation.

	Low Saturation	High Saturation
1. **Direct gains/HH**	3,304	854
2. **Indirect gains/HH**	0	495
3. **Ratio of indirect to direct gains**	0.00	0.58
4. **Direct beneficiary population (HH)**	247	495
5. **Total local population (HH)**	3,553	3,553
6. **Total direct gains**	816,984	422,248
7. **Total indirect gains**	0	1,757,880
8. **Total gains (direct + indirect)**	816,984	2,180,128
9. **Fraction of gains indirect**	0.00	0.81
10. **Private gains/HH**	3,304	1,349
11. **Total private gains**	816,984	666,945
12. **Fraction of gains private**	1.00	0.31