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

MARSHALL BURKE

LAUREN FALCAO BERGQUIST

EDWARD MIGUEL

Large and regular seasonal price fluctuations in local grain markets appear to offer African farmers substantial intertemporal 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 intertemporally. Providing timely access to credit allows farmers to buy at lower prices and sell at higher prices, increasing farm revenues and generating a return on investment of 29%. To understand general equilibrium (GE) 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. *JEL* Codes: D21, D51, G21, O13, O16, Q12.

I. 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). Price fluctuations over time are particularly pronounced, with grain prices in major markets regularly rising by 25–40% between the harvest and lean seasons, and

often by 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, many farmers tend to sell their crops immediately after harvest and then, several months later during the lean season, return to the market as consumers once prices have risen.

In this article, 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 market 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 partly due to limited borrowing opportunities. Lacking access to credit or savings, farmers report selling their grain at low postharvest prices to meet urgent cash needs (e.g., to pay school fees). To meet consumption needs later in the year, many 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 nongovernmental organization (NGO), we offered 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 loan sold significantly less and purchased significantly more maize in the period immediately following harvest, and this pattern reversed during the period of higher prices six to nine months later. This change in their marketing behavior resulted in a statistically significant increase in revenues (net of loan interest) of 1,573 Ksh (roughly US $18), a return on investment of 29% over a roughly nine-month period. We replicated the experiment in back-to-back years to test the robustness of these results and find remarkably similar results on primary outcomes in both years.

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 conducted a long-run follow-up survey with respondents one to two years after the harvest-time credit intervention had been discontinued. We find no evidence of sustained shifts in the timing of farm sales in subsequent seasons, nor any long-run effect on sales or revenues in subsequent years, though we do find some evidence of heterogeneity by treatment saturation.

To explore whether this shift in sales behavior by individual farmers had an effect on market-level prices, we experimentally varied the density of treated farmers across locations and tracked market prices at 52 local market points. We find that increased grain storage at the market level (induced by the credit intervention) led to significantly higher prices immediately after harvest and 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 altered the profitability of the loan. By dampening the arbitrage opportunity posed by seasonal price fluctuations, treated individuals in areas highly saturated with loans showed diminished revenue gains relative to farmers in lower saturation areas. We find that although treated farmers in high-saturation areas stored significantly more than their control counterparts, doing so was not significantly more profitable; the reduction in seasonal price dispersion in these areas appeared to have reduced the benefits of loan adoption. In contrast, treated farmers in low-density areas had 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.[1] In terms of policy, the general equilibrium (GE) effects shaped the distribution of the welfare gains of the harvest-time loan in our setting: whereas recipients gained relatively less than they would in the absence of such effects, we find suggestive evidence that nonrecipients benefited from smoother prices, even though their storage behavior remained unchanged. Though estimated effects on untreated individuals are relatively imprecisely estimated, a welfare calculation taking the point

1. Because we do not examine prices of other goods beyond maize, some scholars might not consider these estimates to represent the full general equilibrium effect.

estimates at face value suggests that 81% of overall gains in high treatment intensity areas were due to these spillover effects. These indirect gains, 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 is violated (Rubin 1980), the evaluation of an individually randomized loan product may conclude that there is a null effect even when there are large positive social welfare impacts. Although 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 local supply, and credit-induced expansion may therefore be less profitable to borrowers than it would be in a more integrated market or 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 heterogeneous effects of expanded credit access. Experimental evaluations have generally found that small enterprises randomly given 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 constraints. Why do we find positive effects on firm profitability when many other

---

2. Experimental evaluations of microcredit include Karlan and Zinman (2011), Angelucci, Karlan, and Zinman (2015), Banerjee et al. (2013), Attanasio et al. (2015), and Crépon et al. (2015), among others. See Banerjee (2013) and Karlan and Morduch (2010) for 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, McKenzie, and Woodruff (2008), McKenzie and Woodruff (2008), Fafchamps et al. (2013), and Blattman, Fiala, and Martinez (2014). Studies finding much more limited returns include Karlan, Knight, and Udry (2012) and Berge, Bjorvatn, and Tungodden (2014).

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 shown to be important for encouraging investment (Field et al. 2013). Finally, as described, the GE 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, Knight, and Udry 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 article is closest to a number of recent papers that

---

3. Existing studies have concluded that many small businesses or potential microentrepreneurs simply might not possess profitable investment opportunities (Karlan, Knight, and Udry 2012; Banerjee 2013; Banerjee et al. 2013; Fafchamps et al. 2013) or may lack the managerial skill or ability to channel capital toward these investments (Berge, Bjorvatn, and Tungodden 2014; Bruhn, Karlan, and Schoar 2018).

4. For instance, Karlan, Knight, and Udry (2012, 22) conclude by stating, "Few if any studies have satisfactorily tackled the impact of improving one set of firms' performance on GE 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, for example, 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 that directly compete for a limited supply of inputs to production. For example, see De Mel, McKenzie, and Woodruff (2008) and their discussion of returns to capital for bamboo sector firms, which must compete over a limited supply of bamboo.

examine the role of borrowing constraints in households' storage decisions and seasonal consumption patterns.[6] Using secondary data from Kenya, Stephens and Barrett (2011) argue that credit constraints substantially alter smallholder farmers' marketing and storage decisions; Basu and Wong (2015) show that allowing farmers to borrow against future harvests can substantially increase lean-season consumption. Similarly, Dillon (2017) finds in Malawi that an administrative change in the school calendar that shifted the timing of school fee payments to earlier in the year forced credit-constrained households with school-aged children to sell their crops earlier and at a lower price. Fink, Jack, and Masiye (2018) show that agricultural loans aimed at alleviating seasonal labor shortages can improve household welfare in Zambia, while Beaman et al. (2015) 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 postharvest cash needs cause farmers to store less than they would in an unconstrained world. Taken together, the body of evidence suggests that farmers are credit constrained at multiple points in the year and that alleviating these constraints can have important effects on production decisions, consumption outcomes, and local prices.

The remainder of the article proceeds as follows. Section II describes the setting and the experiment. Section III describes our data, estimation strategy, and preanalysis plan. Section IV presents baseline estimates, ignoring the role of GE effects. Section V presents the market-level effects of the intervention. Section VI shows how these market-level effects shape the individual-level returns to the loan. Section VII concludes.

II. Setting and Experimental Design

II.A. Arbitrage Opportunities in Rural Grain Markets

Seasonal fluctuations in prices for staple grains appear to offer substantial intertemporal arbitrage opportunities, in our study region of East Africa and in other parts of the developing world.

---

6. In an early contribution, McCloskey and Nash (1984) attribute the dramatic reduction in seasonal grain price fluctuations observed in England between the fourteenth and seventeenth centuries to a reduction in interest rates.
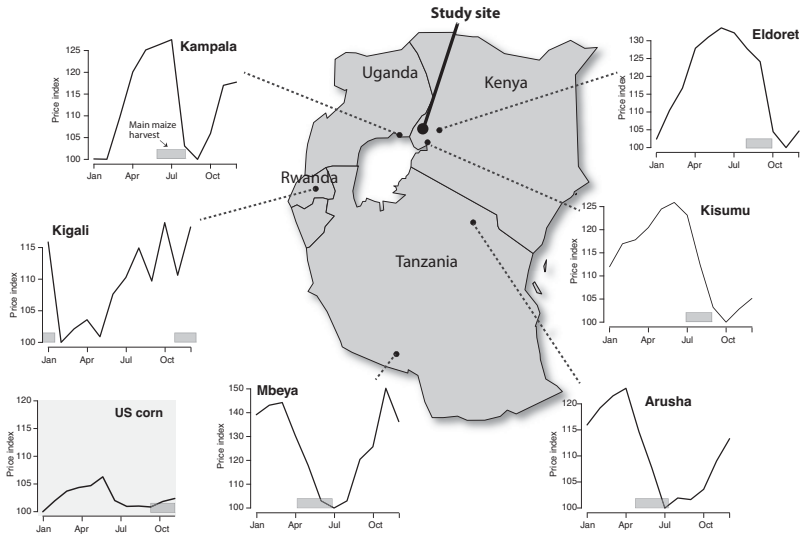
FIGURE I

Monthly Average Maize Prices

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 indicated, and the gray squares represent an independent estimate of the months of the main harvest season in the given location. Price fluctuations for maize (corn) in the United States are shown in the lower left for comparison.

Although 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 these markets are shown in Figure I. Increases in maize prices in the six to eight months following harvest average roughly 25–40% in these markets; price increases reported elsewhere in Africa are consistent with these figures, if not higher.[7]

These increases also appear to be a lower bound on typical increases observed in the smaller markets in our study area, which (relative to these much larger markets) are characterized

7. For instance, Barrett (2007) reports seasonal rice price variation in Madagascar of 80%, World Bank (2007) reports seasonal maize price variation of about 70% in rural Malawi, and Aker (2008) reports seasonal variation in millet prices in Niger of 40%.
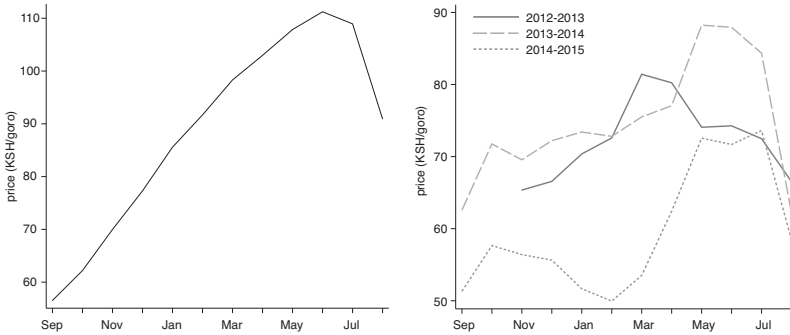
FIGURE II

Maize Price Trends

Left panel: maize price trends in the prestudy period, computed from 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 (2.2 kg). Right panel: maize price trends (study period and poststudy period). Average monthly maize prices for the period 2012–2014 (study period) and 2014–2015 (poststudy period), averaged over all markets in our sample (data from market survey). Prices are in Kenyan shillings per goro. The exchange rate during the study period ranged from 80 to 90 Kenyan shillings per US\$.

by smaller catchment areas and less outside trade. We asked farmers at baseline to estimate the average monthly prices of maize at their local market point over the five years prior to our experiment. As shown in Figure II, 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, and for a year after the intervention ended (more on this data collection below).[9] Figure II also 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

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, to align with the long-run follow-up data collection on the farmer side.

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 nonfood consumption dropping noticeably during the lean season (Kaminski, Christiaensen, and Gilbert 2014; Basu and Wong 2015). 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 suggest the cost of these storage materials is low, at around 3.5% of the value of the crop at harvest time. Postharvest losses also appear minimal in this setting, with an average of 2.5% of the crop lost over a six- to nine-month storage period (see Online 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. Online Appendix Figure A.1 shows data from two earlier pilot studies conducted either by our NGO partner (in 2010–2011, with 225 farmers) or in conjunction with our partner (in 2011–2012, with a different sample of 700 farmers). These studies tracked maize inventories, purchases, and

---

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%.

sales for farmers in our study region. In both years, the median farmer exhausted her inventories about five months after harvest and switched from being a net seller of maize to a net purchaser, as shown in the right panels of the figure.[11] 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. 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 postharvest period. In particular, because nearly all of these farm households have school-aged children, and a large percentage of a child's school fees are typically due in the few months after harvest in January, many farmers report selling much of their harvest to pay these fees, which account for 17% of the value of harvest and 37% of harvest income.[12] Indeed, many schools in the area will accept in-kind payment in maize during this period. Farmers report having to pay other bills that accumulated throughout the year during the postharvest period. Finally, many farmers spend more on discretionary expenditures during this harvest period, which may be reflective of high levels of impatience or present-biased preferences. Regardless of the source, harvest is a time of large expenditures; we estimate that 43% of farmers' expenditures occur in the three months after harvest (round 1 in our survey).

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. Although storage would be highly profitable at even the relatively high interest rates charged by formal banking

---

11. More than half of the farmers in our sample report having some form of nonfarm income from either nonagricultural self-employment or salaried employment.

12. The percent of harvest income is larger than the percent of harvest value because not all of the harvest is sold.

institutions in Kenya (around 20% annually, compared to the typically greater than 40% price increases regularly observed over the nine-month postharvest period),[13] very few of the smallholder farmers in our sample have access to these formal loans; only 8% of households reported having taking a loan from a commercial bank or lender in the year prior to the baseline survey. Informal credit markets also appear relatively thin, with less than 25% of farmers reporting having given or received a loan from a moneylender, family member, or friend in the three months before the baseline. Furthermore, these loans often carry much higher interest rates. For example, the median household in our sample that took out a loan from a moneylender reported paying interest rates of 25% a month. Given such a high rate, it would not make sense for farmers to borrow informally from moneylenders for the purpose of storage (and given the aggregate nature of the harvest season, which affects all households at the same time, respondents in focus groups reported that borrowing from friends, family members, or other informal sources for the purpose of storage was challenging). Finally, although all of our study farmers at baseline are receiving in-kind provision of fertilizer and seeds on credit at planting from our partner organization One Acre Fund (more on this below), OAF had not provided cash loans to its clients, nor any sort of loan other than the in-kind input loan at planting, prior to our intervention.

Absent other means of borrowing, and given the high expenditure needs they report in the postharvest period, farmers end up liquidating grain rather than storing it. 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 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). Forty percent of our sample reported only purchasing maize over this period, and the median farmer in this group made all of their purchases after

---

13. This is true even after accounting for storage costs and losses (e.g., due to spoilage), which we estimate to be around 6%. Moreover, as noted, these fluctuations are often much greater than 40% in rural areas such as the one in which our study takes place.

January. Stephens and Barrett (2011) report similar patterns for other households in western Kenya during an earlier period.

It is worth noting that other factors besides credit constraints may be at play in restricting farmers' ability to store. However, based on pilot and baseline data collected for this project, it appears unlikely that storage is constrained by either the fixed or marginal costs of storing additional bags, nor by grain losses due to moisture or pests when grain is stored for many months; under the technology they currently use, farmers estimate they lose only 2.5% of their grain to rotting or pests when it is stored for six to nine months. Online Appendix B explores these other factors and other interventions in the literature that have attempted to address them, for example, by distributing storage equipment or encouraging communal storage (Basu and Wong 2015; Aggarwal, Francis, and Robinson 2018). For the most part, these factors are outside of the scope of this article, but the GE mechanisms explored in Sections V and VI of this article are in principle relevant for any intervention that succeeds in improving farmers' ability to store.

II.B. 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 a harvest-time cash loan. OAF is an agricultural NGO that provides fertilizer and seeds to groups of farmers on credit, as well as input delivery and training on improved farming techniques. Prior to this intervention, OAF had provided cash loans to its clients; their existing product is given in-kind and only at planting time.[14] 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.[15] OAF typically serves about 30% of farmers in a given sublocation.

The study sample is drawn from existing groups of OAF farmers in Webuye and Matete districts in western Kenya. The year 1 sample consists of 240 existing OAF farmer groups drawn from 17

14. The timing of farmer repayment for the in-kind product is spaced throughout the year. The average value of the in-kind products is $89.

15. A sublocation is a group of four to five villages, with a typical population of 400–500 people.

different sublocations, 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.[16] Ultimately, of the 1,019 individuals in our year 2 sample, 602 are drawn from the year 1 sample and 417 are new.

Figure III 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), and in low-intensity sites, we only enrolled 40% of OAF groups in the sample (for a sample of 69 groups). Within each sublocation, groups were randomized into treatment or control. In year 1, two-thirds of the groups in each sublocation were randomized into treatment (more on this below) and one-third into control. In year 2, half of the groups 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 groups 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,[17] but rerandomized groups into treatment and

16. 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 Online Appendix Table L.2), 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.

17. Such that, for example, if a sublocation was a high-intensity sublocation in year 1 it remained a high-intensity sublocation in year 2. Although we would have liked to rerandomize the intensity across sublocations, during the study design we saw no easy way to both stratify individual-level treatments and rerandomize treatment intensity, given how we had initially randomized treatment intensity (which required sampling more groups in the high-intensity areas at baseline)
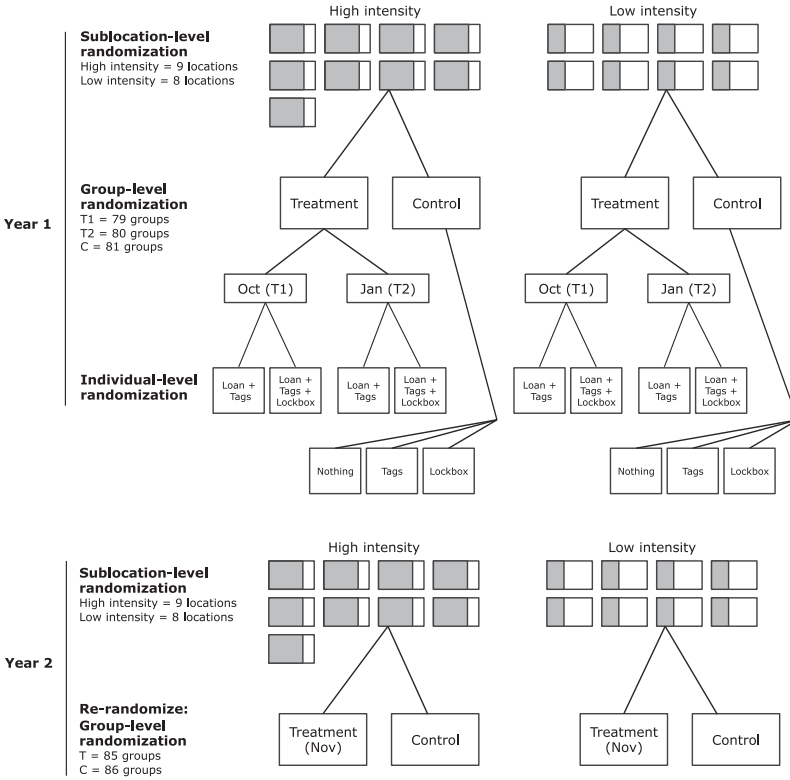
FIGURE III

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 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. Numbers of randomized units given on the left.

and given our original interest in estimating multiyear treatment effects (which we pursued by stratifying Y2 treatment on Y1 treatment status). For this reason, treatment intensity across sublocations was retained in both years of the study.

control, stratifying on their treatment status from year 1.[18] Given the roughly 35% reduction in overall sample size in year 2, treatment saturation rates (the number of treated farmers per sublocation) were effectively 35% lower in year 2 compared with 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). 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), and the other half received the loan in January (immediately before school fees are due, although still several months before the local lean season). As described in Section IV, 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 six farmers in each loan group, and a randomly selected eight farmers in each of the 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.[19] In year 1, there was a cap of seven bags for which farmers could be eligible; in year 2, the cap was five bags. In year 1, to account for the expected price increase, October bags were valued at 1,500 Ksh and January bags at 2,000 Ksh. In year 2, bags were valued at 2,500 Ksh. Each loan carried with it

18. 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.

19. 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.

a "flat" interest rate of 10%, with full repayment due after nine months.[20] 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.[21] These tags also represent a "nudge" or encouragement to store from OAF. Because tags could represent a meaningful treatment in their own right, in the year 1 study we offered a separate treatment arm in which groups received only the tags.[22] This allows us to separate the effect of an OAF nudge to store from the role of credit per se.

Finally, because self-control or other 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 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. Although 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.

20. For example, a farmer who committed five bags when offered the October loan in year 1 would receive 5 * 1,500 = 7,500 Ksh in cash in October (∼US $90 at current exchange rates), and would be required to repay 8,250 Ksh by the end of July. Annualized, this interest rate is slightly lower than the 16–18% APR charged on loans at Equity Bank, the main rural lender in Kenya.

21. Such behavior is consistent with evidence from elsewhere in Africa that individuals may take out loans or use commitment savings accounts as a way to demonstrate that they have little to share with others (Baland, Guirkinger, and Mali 2011; Brune et al. 2016).

22. 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.
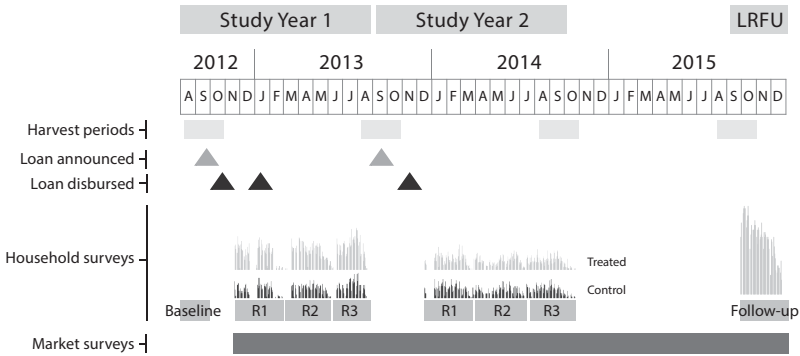
FIGURE IV

Timing of the Main Harvest Periods, Study Interventions, and Data Collection Periods

Rectangles indicate the main maize harvest period. Light gray and dark gray histograms depict the timing of individual household surveys for treated (light) and control (dark) households, and for households in the long-run follow up (LRFU; gray). Light gray boxes labeled R1, R2, and R3 depict the main survey rounds as described in the text. Dark gray box depicts the timing of the market price surveys, which occurred throughout the study period.

The tags and lockbox treatments were randomized at the individual level during year 1. These treatments were not included in year 2 because of minimal treatment effects in year 1 (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. 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 (similarly for the men and for the January loan). In the control groups, in which we were following eight 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.

III. DATA AND ESTIMATION

The timing of the study activities is shown in Figure IV. 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 nonfood consumption expenditure; household borrowing, lending, and saving behavior; household transfers with other family members and neighbors; sources of nonfarm income; time and risk preferences; and digit span recall.

We undertook three follow-up rounds over the ensuing nine months, spanning the "long rains" postharvest period, the "long rains" spring planting period, and concluding just prior to the following year's "long rains" harvest season.[23] The multiple follow-up rounds were motivated by three factors. First, a simple intertemporal model of storage and consumption decisions suggests that although 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) have 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 same sample as year 1, a second baseline was not run prior to year 2. In practice, due to the administrative shifts in farmer group composition described in greater detail in Section II, 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).[24] A similar schedule of three

---

23. The "long rains" season is the primary growing season.

24. Because the loan offer was rerandomized in year 2, however, this should not affect inference regarding the impacts of the loan. We can also run balance tables for year 2 farmers if we restrict to the sample that was also present in year 1. Farmers appear balanced on most outcomes, though there does appear to be some

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%, 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).

Table I 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. Online Appendix Table K.1 shows the analogous table comparing individuals in the high- and low-treatment-density areas; we find balance on all variables except two: the average price increase from September to June and the percent of farmers who can correctly calculate the interest rate. Although this is in line with the number of variables one would expect to be imbalanced due to chance given the number of variables tested, the former difference is potentially important and suggests that at baseline, areas of high treatment saturation may historically have seen greater price swings than low saturation areas. Given that we find lower price swings in high-intensity areas after treatment, this suggests that our estimated treatment effect may be, if anything, an underestimate of the true impact of treatment. Moreover, we show that all results are robust to controlling for any baseline imbalances in Online Appendix J and K.[25]

Year 1 treatment status is predictive of year 2 re-enrollment in the study (treated individuals were more likely to reregister for OAF in the second year, perhaps reflecting a positive appraisal of the value of the loan). However, because year 2 treatment status was rerandomized and stratified by year 1 treatment status, this does not alter the internal validity of the year 2 results.[26]

---

imbalance in harvest levels at baseline (two years prior) among this subsample. While we lack baseline controls to adjust for this in our full sample, we can control for self-reported harvest in year 2; Online Appendix Table F.2 demonstrates that results are robust to controlling for this potential imbalance.

25. See Online Appendix Table J.5 and Table K.2.

26. This does, however, mean that we cannot exploit the rerandomization 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. This also slightly alters the composition of the year 2 sample, relevant to external validity. Online Appendix L explores this further.

TABLE I

SUMMARY STATISTICS AND BALANCE AMONG BASELINE COVARIATES.

Baseline characteristic	Treat	Control	Obs.	T – C Std. diff.	T – C *p*-value
Male	0.30	0.33	1,589	− 0.08	.11
Number of adults	3.00	3.20	1,510	− 0.09	.06
Children in school	3.00	3.07	1,589	− 0.04	.46
Finished primary school	0.72	0.77	1,490	− 0.13	.02
Finished secondary school	0.25	0.27	1,490	− 0.04	.46
Total cropland (acres)	2.44	2.40	1,512	0.01	.79
Number of rooms in household	3.07	3.25	1,511	− 0.05	.17
Total school fees	27,240	29,814	1,589	− 0.06	.18
Average monthly consumption (Ksh)	14,971	15,371	1,437	− 0.03	.55
Average monthly consumption/capita (log)	7.97	7.96	1,434	0.02	.72
Total cash savings (Ksh)	5,157	8,021	1,572	− 0.09	.01
Total cash savings (trim)	4,732	5,390	1,572	− 0.05	.33
Has bank savings acct	0.42	0.43	1,589	− 0.01	.82
Taken bank loan	0.08	0.08	1,589	− 0.02	.73
Taken informal loan	0.24	0.25	1,589	− 0.01	.84
Liquid wealth (Ksh)	93,879	97,281	1,491	− 0.03	.55
Off-farm wages (Ksh)	3,917	3,797	1,589	0.01	.85
Business profit (Ksh)	2,303	1,802	1,589	0.08	.32
Avg %∆ price Sep–Jun	133.49	133.18	1,504	0.00	.94
Expect %∆ price Sep 12–Jun 13	124.68	117.26	1,510	0.14	.15
2011 LR harvest (bags)	9.36	9.03	1,511	0.02	.67
Net revenue 2011 (Ksh)	−3,304	−4,089	1,428	0.03	.75
Net seller 2011	0.32	0.30	1,428	0.05	.39
Autarkic 2011	0.07	0.06	1,589	0.03	.51
% maize lost 2011	0.02	0.01	1,428	0.03	.57
2012 LR harvest (bags)	11.18	11.03	1,484	0.02	.74
Calculated interest correctly	0.71	0.73	1,580	− 0.03	.50
Digit span recall	4.57	4.58	1,504	− 0.01	.89
Maize giver	0.26	0.26	1,589	− 0.00	.99

Balance table for the Y1 sample (for which we have baseline characteristics). The first two columns give the means in each treatment arm. The third column gives the total number of observations across the two groups. The last two columns give differences in means normalized by the standard deviation in the control group, with the corresponding *p*-value on the test of equality.

Notes. "Total school fees" are the total school fees paid by the household in the past 12 months. "Taken bank loan" is whether anyone in the household has taken any loans from a commercial bank or commercial lender in the past 12 months. "Taken informal loan" is whether anyone in the household has taken any loans from a moneylender or someone else outside the household in the past 12 months. "Liquid wealth" is the sum of cash savings and assets that could be easily sold (e.g., livestock). "Off-farm wages" is the total amount earned by anyone in the household who worked in a job for cash in the past month. "Business profits" are the total profits earned from all business run by anyone in the household. "Avg %∆ price Sep–Jun" is the percentage difference between the (self-reported) average market price for maize in September and June over the past five years. "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 three months.

To explore the long-run effects of the loan, we also ran a long-run follow-up (LRFU) survey from November to 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 (1,019 individuals) and a representative subset of the year 1 only sample (another 481 individuals), for a total sample of 1,500 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, nonfarm 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.[27] Online Appendix L provides further discussion.

In addition to farmer-level surveys, we 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. Each month, enumerators visited each market and collected prices from three traders in the market; we average these three prices to get the price for that market in that month. Finally, we use administrative data on loan repayment that was generously shared by OAF.

III.A. Preanalysis Plan

To limit both risks and perceptions of data mining and specification search (Casey, Glennerster, and Miguel 2012), we registered a preanalysis plan (PAP) for year 1 prior to the analysis of

---

27. 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 farmers 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.

any follow-up data.[28] The year 2 analysis follows a near identical plan. The PAP can be found in Online 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. Because in practice the loan was offered at slightly different points in time (October and January in 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, the November–June price gap, and instead analyze the nonparametric effect of treatment on the evolution of monthly prices, as well as a level and time trend effect. Online Appendix J.VI presents the prespecified November–June effect. For all analyses, we maintain our original hypothesis that the 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 postharvest surplus period and later negative as stored grain is released onto the market.

In two other cases we add to the PAP. First, in addition to the regression results specified in the PAP, we present graphical results for many of the outcomes. These results are based on nonparametric estimates of the parametric regressions specified in the PAP and are included because they clearly summarize how treatment effects evolve over time, but because 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.

III.B. Estimation of Treatment Effects

In all analyses, we present results both separately by year and pooled across years. Because the year 2 replication produced results that are quantitatively quite similar to the year 1 results

---

28. The PAP is registered at: https://www.socialscienceregistry.org/trials/67, and was registered on September 6, 2013. The complete set of results are available on request.

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 90-kg bags of maize the household had in their maize store at the time of each survey. This amount is visually verified by our enumeration team and is likely 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" because we probably 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 variable we trim the top and bottom 0.5% of observations, as specified in the PAP.

T_{jy} is an indicator for whether group j was assigned to treatment in year y, and Y_{ijry} is 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):

$$(1) \qquad Y_{ijry} = \alpha + \beta T_{jy} + \eta_{ry} + d_t + \gamma_s + \varepsilon_{ijry}.$$

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 later, loan take-up was high. To absorb additional variation in the outcomes of interest, we also control for survey date (d_t), as prespecified. Each follow-up round spanned 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. Finally, we follow Bruhn and McKenzie (2009) and control for stratification dummies (γ_s), again

as prespecified.[29] Standard errors are clustered at the loan group level. Finally we also present family error-wise corrected *p*-values for our main family of outcomes.

The assumption in equation (1) 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:

$$(2) \qquad Y_{ijry} = \sum_{r=1}^{3} \beta_r T_{jy} + \eta_{ry} + d_t + \gamma_s + \varepsilon_{ijry},$$

and test whether the β_r are the same across rounds (as estimated by interacting the treatment indicator with round dummies). Unless otherwise indicated, we estimate equations (1) and (2) for each of the hypotheses below.

To explore heterogeneity in treatment effects, we estimate:

$$(3) \quad Y_{ijry} = \alpha + \beta_1 T_{jy} + \beta_2 Z_{i0} + \beta_2 T_{jy} * Z_{i0} + \eta_{ry} + d_t + \gamma_s + \varepsilon_{ijry},$$

where Z_{i0} is the standardized variable by which we explore heterogeneity, as measured at baseline. As prespecified, 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

---

29. We also prespecified controlling for baseline values of outcome variables, but because we lack a baseline for our year 2 data, we omit these in the main specification. In Online Appendix F, we show results with baseline controls for the year 1 sample and the subset of the year 2 sample that was present in year 1 and therefore has baseline measures.

region, and we assign these markets to the nearest sublocation. To estimate price effects we begin by estimating the following linear model:

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

where p_{msty} represents the maize sales price at market m in sublocation s in month t in year y.[30] 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 $\beta_3 < 0$, that is, that prices in areas with more treated farmers are higher after harvest 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, Gelbach, and Miller 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, Gelbach, and Miller (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 (1) and (2), interacting the individual-level treatment indicator with the treatment density dummy. The pooled equation is thus:

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

If the intervention produces sufficient individual-level behavior to generate market-level effects, we predict that $\beta_3 < 0$ and perhaps that $\beta_2 > 0$—that is, treated individuals 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

---

30. 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 postharvest prices.

areas. As in equation (4), we report results with errors clustered at the sublocation level.[31]

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

$$(6) \qquad Y_{ij} = \alpha + \beta T_{jy} + d_t + \varepsilon_{ij},$$

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 also estimate the following specification:

$$(7) \qquad Y_{ij} = \alpha + \beta_1 T_{j1} + \beta_2 T_{j2} + \beta_3 T_{j1} * T_{j2} + d_t + \varepsilon_{ij},$$

in which T_{j1} is an indicator for being in a treated group in year 1, T_{j2} 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 the decision to return to the sample from the year 1 to year 2 study was differential based on treatment status (see Online 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.

IV. INDIVIDUAL-LEVEL RESULTS

IV.A. Harvest Loan Take-up

Take-up of the loan treatments was quite high. Of the 954 individuals in the year 1 treatment group, 617 (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.[32]

---

31. Note that we cannot include controls for stratification dummies in this specification, as treatment was stratified on sublocation, the level of the randomized saturation treatment.

32. Relative to many other credit market interventions in low-income settings in which documented take-up rates range from 2% to 55% of the surveyed population (Karlan, Morduch, and Mullainathan 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 microfinance institutions. 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

Unconditional loan sizes in the two treatment groups were 4,817 Ksh and 6,679 Ksh, or about US \$57 and US \$79, respectively. The average loan sizes conditional on take-up were 7,533 Ksh (about US \$89) for year 1 and 10,548 Ksh (or \$124) for year 2.[33] This is 43% of the value of harvest (valued at harvest time prices). Of those who take out loans, 24% take out the maximum loan size. Default rates were extremely low, at less than 2%.

IV.B. 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 both broken down by each year and pooled. As explained in Section III, 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 II–IV and Figure V present the results of estimating equations (1) and (2) on the pooled treatment indicator, either parametrically (in the table) or nonparametrically (in the figure). The left column in Figure V shows 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 right column presents the difference in treatment minus control over time, with the 95% confidence interval calculated by bootstrapping the Fan regression 1,000 times.[34]

Farmers respond to the intervention as anticipated. They hold significantly more inventories for much of the year, on average about 25% more than the control group mean (Table II, column (6)). Inventory effects are remarkably similar across both years of

---

non-OAF farmers were to adopt the loan if offered it, population-wide take-up rates of our loan product would still exceed 15%.

33. Recall in year 1 there were two versions of the loan, one offered in October and the other in January. Of the 474 individuals assigned to the October loan treatment (T1), 333 (71%) applied and qualified for the loan. For the January loan treatment (T2), 284 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 (about US \$62 and US \$51) for T1 and T2, respectively, and we can reject with 99% confidence that the loan sizes were the same between groups. The average loan sizes conditional on take-up were 7,627 Ksh (about US \$90) for T1 and 7,423 Ksh (US \$87) for T2, and in this case we cannot reject that conditional loan sizes were the same between groups.

34. In Online Appendix F we check the robustness of these nonparametric results to the choice of bandwidth size.

TABLE II

INVENTORY EFFECTS, INDIVIDUAL LEVEL

	Y1		Y2		Pooled	
	Overall (1)	By rd (2)	Overall (3)	By rd (4)	Overall (5)	By rd (6)
Treat	0.57***		0.55***		0.56***	
	(0.14)		(0.13)		(0.10)	
Treat - R1		0.87***		1.24***		1.05***
		(0.28)		(0.24)		(0.18)
Treat - R2		0.75***		0.30*		0.55***
		(0.17)		(0.17)		(0.12)
Treat - R3		0.11		0.08		0.09
		(0.08)		(0.34)		(0.16)
Observations	3,836	3,836	2,944	2,944	6,780	6,780
Mean DV	2.67	2.67	1.68	1.68	2.16	2.16
Std. dev. DV	3.51	3.51	2.87	2.87	3.23	3.23
R squared	0.37	0.37	0.21	0.21	0.33	0.33
p-val Treat	<.01		<.01		<.01	
p-val Treat FWER	<.01		<.01		<.01	
p-val Treat - R1		<.01		<.01		<.01
p-val Treat - R1 FWER		<.01		<.01		<.01
p-val Treat - R2		<.01		.07		<.01
p-val Treat - R2 FWER		<.01		.17		<.01
p-val Treat - R3		.18		.81		.56
p-val Treat - R3 FWER		.33		.91		.63

Notes. The dependent variable is inventories, as measured by the number of 90-kg bags of maize held by the household at the time of survey. "Treat" is an indicator for being in a treatment group. "Treat - Rx is an interaction between an indicator for being in a treatment group and an indicator for being in round x. Regressions include round-year fixed effects, strata dummies, and controls for survey date, with errors clustered at the group level. "Mean DV" and "Std. dev. DV" are the mean and standard deviation of the dependent variable among the control group. Standard and family-wise error rate (FWER) p-values are presented (family of outcomes is inventories, net revenues, consumption, and effective prices, as prespecified). Significant at 90% (*), 99% (***) confidence.

the experiment. The size of the inventory effect in round 1 suggests that 50% of the loan was "spent" on a reduction in net maize sales (or conversely, an increase in maize inventories).[35] It is possible

35. The increase in inventories in round 1 (pooled specification, Table II) is 1.05 bags. Given the average value of a bag of maize in round 1 is 2,625 Ksh, this is valued at 2,756 Ksh. The average loan size (unconditional on take-up, which is appropriate because the estimated treatment effects are intention-to-treat effects) was 5,500 Ksh. This suggest that 50% of the loan was "spent" on a reduction in net maize sales. Note also that the round 1 survey occurs after the October and November loans were disbursed in years 1 and 2, respectively (see Figure IV). Therefore, farmers may have sold part of their inventory that they used for OAF loan eligibility by the time we measure it in round 1.

TABLE III

NET REVENUE EFFECTS, INDIVIDUAL LEVEL

	Y1		Y2		Pooled	
	Overall (1)	By rd (2)	Overall (3)	By rd (4)	Overall (5)	By rd (6)
Treat	265		855***		533***	
	(257)		(302)		(195)	
Treat - R1		−1,165***		16		−614**
		(323)		(445)		(272)
Treat - R2		510		1,995***		1,188***
		(447)		(504)		(337)
Treat - R3		1,370***		565		999***
		(413)		(403)		(291)
Observations	3,795	3,795	2,935	2,935	6,730	6,730
Mean DV	334	334	−3,434	−3,434	−1,616	−1,616
Std. dev. DV	6,055	6,055	6,093	6,093	6,359	6,359
R squared	0.03	0.04	0.07	0.08	0.12	0.12
p-val Treat	.30		.01		.01	
p-val Treat FWER	.38		.01		.01	
p-val Treat - R1		<.01		.97		.02
p-val Treat - R1 FWER		<.01		.97		.04
p-val Treat - R2		.26		<.01		<.01
p-val Treat - R2 FWER		.38		<.01		<.01
p-val Treat - R3		<.01		.16		<.01
p-val Treat - R3 FWER		<.01		.26		<.01

Notes. The dependent variable is net revenues, as measured by the value (in Ksh) of maize sales minus the value of maize purchases that round. The exchange rate during the study period ranged from 80 to 90 Kenyan shillings per US$. "Treat" is an indicator for being in a treatment group. "Treat - Rx is an interaction between an indicator for being in a treatment group and an indicator for being in round x. Regressions include round-year fixed effects, strata dummies, and controls for survey date, with errors clustered at the group level. "Mean DV" and "Std. dev. DV" are the mean and standard deviation of the dependent variable among the control group. Standard and family-wise error rate (FWER) p-values are presented (family of outcomes is inventories, net revenues, consumption, and effective prices, as prespecified). Significant at 95% (**), 99% (***) confidence.

that some of the loan was used for immediate consumption, as one would expect if households are smoothing consumption, given that the return from the loan is not realized until later in the season.[36]

Net revenues[37] are significantly lower immediately postharvest and significantly higher later in the year (Table III, column (6)). The middle part of Figure V presents the time trend of net

36. The positive (though not significant) effect on consumption, in conjunction with a negative effect on net revenues, seen in round 1 are consistent with such an explanation.

37. From which loan interest rates were subtracted for those who took out a loan.
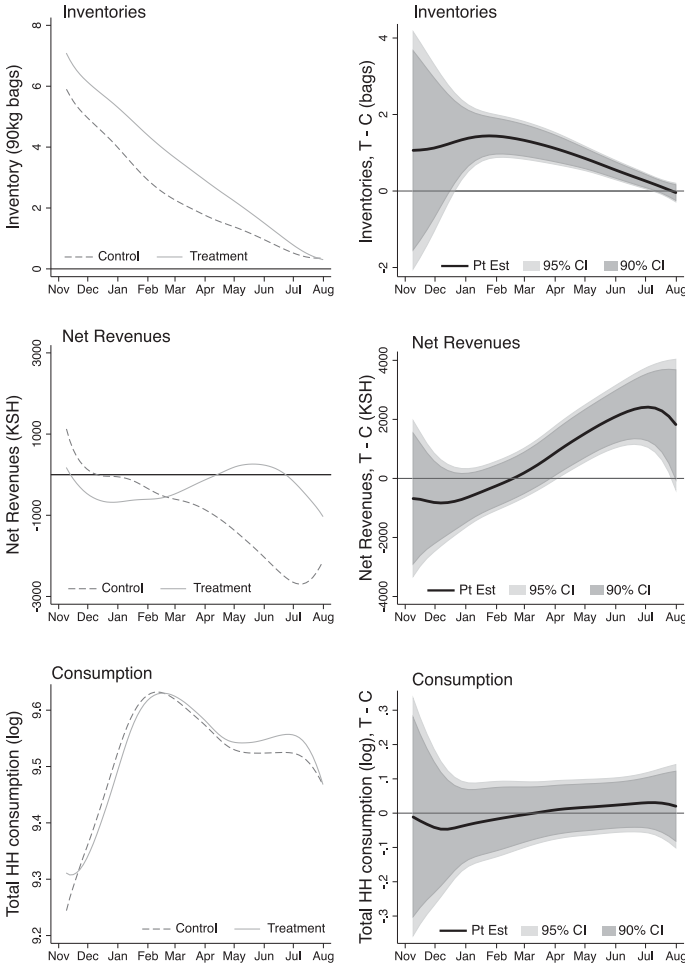
FIGURE V

Pooled Treatment Effects

The left column of plots shows how average inventories, net revenues, and log household consumption evolve from November to August in Y1 and Y2 (pooled) in the treatment group versus the control group, as estimated with Fan regressions. The right column shows the difference between the treatment and control, with the bootstrapped 95% confidence interval shown in gray (100 replications drawing groups with replacement). Inventories are measured by the number of 90-kg bags of maize held by the household. Net revenues are the value (in Ksh) of maize sales minus the value of maize purchases. HH consumption (measured in log Ksh) is aggregated from a detailed 30-day recall consumption module. The exchange rate during the study period ranged from 80 to 90 Ksh per US$.

TABLE IV

HH Consumption (Log) Effects, Individual Level

	Y1		Y2		Pooled	
	Overall (1)	By rd (2)	Overall (3)	By rd (4)	Overall (5)	By rd (6)
Treat	0.01		0.06*		0.04	
	(0.03)		(0.04)		(0.02)	
Treat - R1		− 0.03		0.06		0.01
		(0.05)		(0.05)		(0.03)
Treat - R2		0.03		0.08*		0.05*
		(0.04)		(0.04)		(0.03)
Treat - R3		0.04		0.05		0.04
		(0.04)		(0.05)		(0.03)
Observations	3,792	3,792	2,944	2,944	6,736	6,736
Mean DV	9.48	9.48	9.61	9.61	9.55	9.55
Std. dev. DV	0.63	0.63	0.63	0.63	0.64	0.64
R squared	0.03	0.03	0.05	0.05	0.06	0.06
p-val Treat	.68		.08		.13	
p-val Treat FWER	.69		.10		.13	
p-val Treat - R1		.49		.17		.69
p-val Treat - R1 FWER		.49		.26		.69
p-val Treat - R2		.48		.08		.09
p-val Treat - R2 FWER		.49		.17		.13
p-val Treat - R3		.36		.27		.16
p-val Treat - R3 FWER		.47		.35		.21

Notes. The dependent variable is log HH consumption (measured in logged Ksh), aggregated from a detailed 30-day recall consumption module. "Treat" is an indicator for being in a treatment group. "Treat - Rx is an interaction between an indicator for being in a treatment group and an indicator for being in round x. Regressions include round-year fixed effects, strata dummies, and controls for survey date, with errors clustered at the group level. "Mean DV" and "Std. dev. DV" are the mean and standard deviation of the dependent variable among the control group. Standard and family-wise error rate (FWER) p-values are presented (family of outcomes is inventories, net revenues, consumption, and effective prices, as prespecified). Significant at 90% (*) confidence.

revenue effects, which suggest that treated farmers purchase more maize in the immediate postharvest 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 both the year 2 and the pooled data (see Table III, columns (1), (3), and (5)). Breaking down year 1 results by the timing of the 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. Online Appendix Table D.1

presents results for the year 1 loan, broken down by loan timing. We see in column (6) that the October loan (T1) produced revenue 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. Online Appendix Section D.I 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 Table III column (6), which yields an estimate of 1,573 Ksh, or about US $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 29% return (net of loan and interest repayment), which we consider large.

The bottom row of Figure V and Table IV present the consumption effects (as measured by logged total household consumption). Although point estimates are positive in both years, and marginally significant in year 2, they are not significant at traditional confidence levels when pooled.[38] It is worth noting, however, that the magnitude of the point estimates suggests that much of the increase in net revenues may have gone to consumption, though we lack the statistical precision to say so with certainty.[39]

Table V presents effects on the pattern of net sales and prices paid and received. We see that in the immediate postharvest 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

38. While the round 2 coefficient in the pooled specification is significant at 10% using traditional *p*-values, this effect does not survive the family-wise error rate correction. Because the consumption measure includes expenditure on maize, in Online Appendix E.I we also estimate effects on consumption excluding maize and consumption excluding all food. Results are similar using these measures.

39. Taken literally, the pooled point estimates on net revenues in Table III, column (6), suggests that revenue increased by 1,573 Ksh in total. This is a little less than 3% of the total consumption measured by our survey. The point estimate on consumption effects from Table IV, column (5) is a 4% increase, quite close to this predicted increase of 3% if all additional revenue went to consumption. The lack of treatment effect on cash savings observed (results available on request) is also consistent with the interpretation that most of the increase in net revenue went to consumption.

TABLE V

NET SALES AND EFFECTIVE PRICES, INDIVIDUAL LEVEL

	Net sales		Effective price	
	Overall	By rd	Purchase	Sales
Treat	0.19***		− 57.45**	145.51***
	(0.06)		(27.16)	(41.77)
Treat - R1		− 0.21**		
		(0.10)		
Treat - R2		0.38***		
		(0.10)		
Treat - R3		0.37***		
		(0.09)		
Observations	6,740	6,740	2,014	1,428
Mean DV	− 0.41	− 0.41	3,084.78	2,809.76
Std. dev. DV	2.04	2.04	534.45	504.82
R squared	0.10	0.10	0.09	0.07
p-val Treat			.03	<.01
p-val Treat FWER			.04	<.01

Notes. The dependent variable in columns (1)–(2) is net sales (quantity sold minus quantity purchased, measured in 90-kg bags of maize) that round. Columns (1)–(2) include round-year fixed effects, strata dummies, and controls for survey date, with errors clustered at the group level. The dependent variable in column (3) is "effective purchase price," which is constructed by 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. The dependent variable in column (4) is "effective sales price," which is constructed similarly. Columns (3)–(4) include only one observation per individual (per year). Round fixed effects are omitted in these specifications 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. Instead we include year fixed effects and strata dummies. In all columns, "Treat" is an indicator for being in a treatment group. "Treat - Rx" is an interaction between an indicator for being in a treatment group and an indicator for being in Round x. "Mean DV" and "Std. dev. DV" are the mean and standard deviation of the dependent variable among the control group. Standard and family-wise error rate (FWER) *p*-values are presented for effective prices (family of outcomes is inventories, net revenues, consumption, and effective prices, as prespecified). FWER *p*-values are not presented for net sales, which was not included in the prespecified main family of outcomes. Significant at 95% (**), 99% (***) confidence.

(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 small positive effect, which—off of a negative average net sales amount—means that households are slightly less in deficit.[40] We see

40. 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. From where is the increase in net sales drawn? We assume net sales = amount harvested−postharvest losses−amount consumed−amount transferred and decompose the treatment

in Online Appendix Figure F.2, which shows a more flexible estimate of treatment effects by date, rather than round, that effect estimates in year 1 and year 2 have a similar shape for most of the season.

As a result of these findings, OAF has begun scaling this loan product in Kenya (following a brief hiatus, during which the long-run follow-up study was completed). Given the finding that the timing of credit is important, the product being scaled is the earlier loan, akin to the year 1 October loan and the year 2 loan. In Online Appendix D.II, we pool these two treatments to estimate the likely impact of the policy-relevant program.[41] Because the earlier loan is more effective, we estimate even larger impacts from this product, including a marginally significant increase in consumption of 5% (see Online Appendix Table D.4). Online Appendix F includes several robustness checks for these results.

IV.C. Heterogeneity

Online Appendix Tables C.1–C.3 present the prespecified dimensions of heterogeneity in treatment effects on inventories, revenues, and log household consumption. Because the prespecified 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 somewhat larger treatment effects among wealthy households (significant for revenue outcomes, but not significant for inventories or consumption). Interestingly, we see significant 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

---

effect on each component part. We see a marginally significant (at 10%) increase in the amount that treated households transfer to others by 0.02 bag. We are unable to identify with precision any effects on the other components of net sales (results available on request). Online Appendix F presents effects on net sales and prices broken down by year. It appears that the overall increase in net sales observed in the pooled data is stronger in year 2. In addition, we do not observe the decrease in net sales in year 2, round 1 that we see in year 1, round 1. This may be partially due to slight differences in the timing of the survey rounds across years (see Figure IV for exact survey timing). In particular, round 1 survey collection occurred a bit later in year 2 than in year 1.

41. We thank a referee for suggesting this specification.

that these households have the greatest room for adjustment 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 nonstorers. Expectations regarding the impending seasonal price increase do not appear to be related to take-up or treatment effects. Online Appendix Table C.4 presents heterogeneity by baseline credit access, which was not prespecified. Interestingly, we see that the percentage of households that take up any loans is higher among those who were borrowing at baseline.[42] We see no significant heterogeneity in effects on inventories, net revenue, or consumption by baseline credit access, though these effects are noisily estimated.

IV.D. Secondary Effects of the Loan Offer

Online Appendix E presents effects on potential secondary outcomes of interest. We find no significant effects on food expenditure, calories consumed, or maize eaten (Tables E.1–E.3). We also find no significant effects on school fees paid (the primary expenditure that households say constrains them to sell their maize stocks early; see Table E.4), though effects are generally positive and are marginally significant at 10% for year 1. We find no effect on labor or nonlabor inputs used in the subsequent planting period (Table E.5). We also find no significant effects on nonagricultural business outcomes, including profits earned from and hours worked at nonfarm household-run businesses (Tables E.6 and E.7) and wages earned from and hours worked in salaried employment (Tables E.8 and E.9). We find in Table E.10 a marginally significant decrease in the percentage of households that borrow anything from other sources (both formal and informal) by 2 percentage points, off of a base of 22% borrowing; there may therefore be some offsetting effect on outside loan sources. However, this effect is quite small in magnitude. Moreover, we see no effect on the (unconditional) amount borrowed (Table E.11). We see a significant increase in self-reported happiness by 0.04 points on a 3-point scale (an index for the following question:

---

42. This may be the result of the requirement that farmers have at least something in storage to be eligible for the loan. We do see that the size of the loan taken out is smaller among those who were borrowing at baseline (point estimate of -377.353, with a standard error of 217.129).

"Taking everything together, would you say you are very happy (3), somewhat happy (2), or not happy (1)") (Table E.12). This represents an increase of 0.08 of a standard deviation.

IV.E. Nudges, Temptation, and Social Pressure

The foregoing results suggest that well-timed loans can enable farmers to engage in greater arbitrage of seasonal price differentials and earn higher annual revenues. We interpret this as primarily resulting from relaxing postharvest credit constraints; however, the structure of the loan—the amount of which was a function of the number of bags in storage at the time of loan disbursal—may have also generated a nudge for farmers to store. Similarly, it is possible that the group loan structure may have spurred group-monitoring dynamics. Although we cannot unbundle these alternative possible mechanisms in our main treatment, the tag treatment—in which bags of stored maize were given laminated tags branding the maize as committed to OAF—allows us to explicitly test the impact of a product that nudges farmers to store more grain and generates social awareness of the intent to store, but which, crucially, does not provide liquidity.

We find no effect of this "nudge only" treatment on storage behavior. Estimates are displayed in Online Appendix Table H.1. We see no significant difference in inventories, revenues, or consumption, and point estimates are small. This suggests that credit per se is important in generating the effects seen from the main loan product.[43]

Several other pieces of evidence suggest that relaxing credit constraints was a crucial mechanism. First, the "nudge" to store only lasted until the loan was disbursed; there was no further requirement that farmers store beyond the date of loan disbursal. Yet we see persistent effects on inventories long past the removal of the nudge (see Figure V). The loan-timing results provide further evidence that relaxing liquidity constraints is important. In year 1 of the experiment, both the October and January loans were announced (and the link to stored bags fully explained) in September. If the observed effects are solely driven by a nudge from OAF or by group-monitoring dynamics, we should expect to see similar results for these interventions (in fact, we might expect results to be stronger with the January loan, as the inventory check for

---

43. This also suggests that the tags did not generate significant change to margins related to self-control or kin tax.

loan disbursal occurs later in the season, and therefore the nudge lasts for longer). However, we find instead much stronger results from the October loan, suggesting that the primary lever at play is receiving credit at the right time in the season.

IV.F. Long-Run Effects

Online Appendix G presents the long-run effects of the loan, as measured in the 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 present a specification with the two treatment years interacted, but with the aforementioned caveats described in Section III.

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 Online Appendix Table G.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 in the percent of sales conducted in the high-price lean season). We also see no statistically significant difference in long-run net revenues (although due to the imprecision of these estimates, we cannot rule out large, positive effects).[44] We also see no long-run effect

---

44. Although we see no significant changes in sales timing or revenue in the pooled treatment group, we see when breaking these results down by treatment status some interesting heterogeneity (see Online Appendix Table G.10). Point estimates suggest (and are significant in year 2) that the percentage sold in the lean season and the percentage purchased in the harvest season are higher in low-saturation areas. In high-saturation areas, the negative interaction term cancels this effect out (see Online Appendix Table G.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 postharvest 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, and buying more (significant among year 2 individuals) in the harvest period, when prices are low. As a result, we cannot rule out sizable increases in revenues for control individuals in high-intensity areas; though this effect is measured with considerable noise, it

on amount and value sold or purchased (Online Appendix Tables G.2–G.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 access produced long-run increases in on-farm investment.[45] However, Online Appendix Table G.5 suggests little movement on this margin. We estimate fairly precise null effects on labor inputs, nonlabor inputs, and 2015 long-rains harvest levels.

We also explore other outcomes for 2015. We find no significant effects on a variety of outcomes, including maize eaten, food expenditures, consumption, educational expenditure, school attendance, nonfarm enterprise profits, hours worked in nonfarm enterprises, and hours worked in salaried positions (Online Appendix Tables G.6–G.8). Point estimates on wages in salaried positions are positive, but this 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, although 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.[46] We therefore find little evidence that this one-time injection of credit permanently ameliorated the underlying constraints limiting arbitrage. It is possible that larger injections would do more to push households

---

is consistent with the idea that control individuals may benefit from the loan. See Online Appendix G for greater discussion of this heterogeneity.

45. 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 Online Appendix Tables G.1 and G.9 we see no evidence of such long-run shifts in either sales timing or prices).

46. 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 more muted; see Online Appendix Table G.9).

out of a potential poverty trap zone, as found in studies of "graduation programs" (Banerjee et al. 2015; Bandiera et al. 2017).[47]

IV.G. Saving 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 Online Appendix I we present various estimates suggesting that it would take the farmer three to six 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 for understanding why credit constraints persist in the presence of high-return, divisible investment opportunities.

To test the importance of savings constraints, we examine the impact of the lockbox, as well as its interaction with the loan. Online Appendix Table I.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.

V. 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

---

47. The loan studied here is on average about $100 for those who borrow. Other programs offering larger bundles of assets, skills training, and food stipends have shown long-run effects on poverty. For example, Banerjee et al. (2015) study an asset and skills program valued at $1,120 in purchasing power parity (PPP) terms per household, whereas Bandiera et al. (2017), study seven "graduation programs" that provide asset transfers and food stipends valued at $680–$2,048 per household. Both find significant long-run effects.

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 (i) the treatment generates a substantial shock to the local supply of maize; and (ii) local markets are somewhat isolated, such that local prices are at least partially determined by local supply.

V.A. 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 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 the market location and the location of farmers in our study sample to calculate the "most likely" sublocation of each market, based on the designated sublocation to which the modal study farmer falling within a 3-km radius belongs. Because we draw double the sample from high-intensity compared to low-intensity areas (in accordance with our randomized intensity), we weight the low-intensity observations by 2 to generate a pool reflective of the true underlying OAF population (though in Online Appendix J we show that this weighting has little effect on our estimates). From this pool, we identify the modal farmer sublocation. This procedure, including the radius to be used, was prespecified.[48] As was also prespecified, we test robustness to alternative radii of 1 km and 5 km.

We then use 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 VI and plotted nonparametrically in Figure VI. In each year, we explore the price changes from the period following loan disbursal (November in

---

48. With the exception of the weighting procedure, which we show in Online Appendix J has little effect on results.

TABLE VI

MARKET PRICES FOR MAIZE AS A FUNCTION OF LOCAL TREATMENT INTENSITY

	Main specification (3 km)			Robustness (pooled)	
	Y1	Y2	Pooled	1 km	5 km
High	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)
High 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
p-val High	.052	.172	.044	.124	.056
p-val High bootstrap	.096	.196	.084	.152	.112
p-val Month	.005	.005	.001	.001	.000
p-val Month bootstrap	.040	.000	.034	.022	.000
p-val High * Month	.193	.316	.158	.200	.038
p-val High * Month bootstrap	.176	.316	.170	.218	.056

Notes. The dependent variable is price, as measured monthly following loan disbursal (Nov–Aug in Y1; Dec–Aug in Y2) in market surveys. Prices are normalized to 100 in Nov in low-intensity sublocations. "High" intensity is an indicator 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. To check robustness to small cluster standard error adjustments, p-values from the standard specification are compared to p-values drawn from the wild bootstrap procedure proposed by Cameron et al. (2008), clustered at the sublocation level. Significant at 90% (*), 95% (**), 99% (***) confidence.
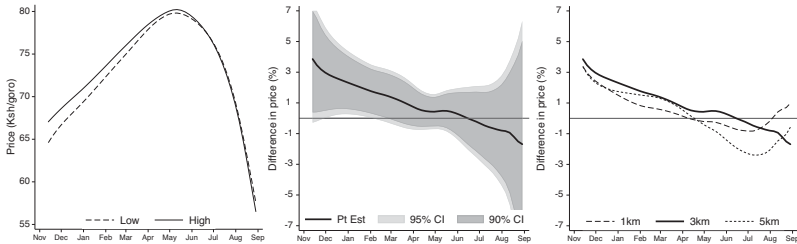

FIGURE VI

Pooled Market Prices for Maize as a Function of Local Treatment Intensity

Markets matched to treatment intensity using sublocation of the modal farmer within 3 km of each market. Left panel: average sales price in markets in high-intensity areas (solid line) versus in low-intensity areas (dashed line) over the study period. Middle panel: average difference in prices between high- and low-intensity areas over time, with the bootstrapped 95% confidence interval shown in light gray and the 90% confidence interval shown in dark gray (prices are normalized to 100 in November in low-intensity sublocations; bootstrap conducted with 1,000 replications drawing groups with replacement). Right panel: robustness of results to alternative radii (1 km, 3 km, and 5 km).

year 1, December in year 2) until the beginning of the subsequent harvest (August in both years). In Figure VI, which presents the pooled data, we see prices in high-intensity markets on average start out almost 4% higher in the immediate postharvest months. As the season progresses, prices in high-intensity 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 VI presents these results according to the empirical specification outlined in Section III. In line with the graphic results in Figure VI, here we see the interaction term on "high" 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 similar point estimates. Online Appendix J presents alternative functional form specifications.

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 II 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 postharvest period. In contrast, the release of this grain onto the market in the lean period appears to happen with more diffuse timing among those in the treatment group (as can be seen in Figure V, 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-of-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.[49] Together, this implies that about 14% of the population was offered treatment in high-intensity sublocations and 7% in low-intensity areas, such that treatment was offered to 7 percentage points more of the population in high-density areas. Table II suggests that treated individuals experienced average increases in inventory (i.e., inward supply shifts) of 25.9%. Taken together, this suggests a contraction in total quantity available in the high-density markets by 1.8%. 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 to see an overall price increase of 1.6%. In the period immediately following harvest, when the inventory effects are most concentrated—during which time inventories are 48.6% higher among treatment individuals—we see an inward supply shift of 3.4%, and should therefore expect to see a 3.1% increase in price.[50] This is quite close to what we observe in Figure VI. We see a jump in price of about 2.5% during this period,[51] which then peters out to a slightly negative (though not significant) effect towards the end of the season.

V.B. Robustness Checks

We check the robustness of the regression results to functional form assumptions. Online Appendix Table J.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 postharvest in high-intensity markets.

49. In year 1, 66% of the sample received the loan offer (one-third received the offer in October, one-third received the loan offer in January, and one-third served as control). In year 2, 50% of the sample received the loan offer (half received the offer in November and half served as control). In this calibration exercise, we use the average of the two years' rates.

50. 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.

51. We measure shifts in postharvest inventories in round 1 of the survey, which was conducted roughly January–February for the average respondent. We therefore estimate the change in price in January–February from Table VI to be $3.97 + 2.5 * (−0.57) = 2.5$.

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 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 markets overall. Comparing these effects to Figure VI, we observe that this is because at the beginning of the lean season prices are still higher in high-intensity markets, with a crossover mid-lean season as prices in high-intensity markets drop below those in low-intensity markets. However, the 1 km and 5 km specifications shown in the right-hand panel in Figure VI suggest that this crossover occurs closer to the transition from the harvest to lean season; therefore the 1 km and 5 km specifications of the binary specification, shown in Online Appendix Table J.1, columns (4)–(5), estimate a more substantial decrease in price for the full lean season.

In another specification check, Online Appendix Table J.2 presents treatment effects by round. We find roughly similarly sized—and in fact, often more precise—effects by round. Of particular interest is whether there is a statistically significant difference between the coefficient estimates on the treatment terms in round 1 and round 3, as this tests whether there is a differential change in prices across high-intensity versus low-intensity areas (this is the analogue of the High * Month interaction term in our primary specification). We find a p-value of .13 for this F-test in our pooled main specification.

We also check the robustness of these results to a more continuous measure of treatment at 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 3-km radius around each market.[52] We reestimate an equation identical to equation (4) with H_s replaced with $ratio_m$, the aforementioned ratio. Results are presented in Online Appendix Table J.3. We also present nonparametric estimates of this specification in Online Appendix Figure J.1, displaying average prices in markets with above- versus

---

52. Because we draw double the sample from high-intensity compared to low-intensity areas (in accordance with our randomized intensity), for the total farmer count, we weight the low-intensity observations by 2 to generate a count reflective of the true underlying OAF population.

below-median ratios. Although results are somewhat less precisely estimated in this specification, the broad patterns remain consistent: prices are higher in the postharvest period and lower in the lean period in markets with a greater proportion of treated individuals in the area.

Finally, we 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,[53] 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 Online Appendix J, 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 are broadly consistent with those from the primary specification, with impacts significant at conventional levels from December to mid-February (p-values are less than .05 for January and less than .1 for December and February). We also check the robustness of our results by conducting the wild bootstrap procedure proposed by Cameron, Gelbach, and Miller (2008) (shown in the notes of Table VI). Although we do see some decrease in statistical precision, these adjustments are small. To ensure that results are not sensitive to a single outlier sublocation, we drop each sublocation one by one and rerun our analysis; the pattern observed in the full data is generally robust to this outlier analysis. Finally, we check the robustness of nonparametric results to the choice of bandwidth size. See Online Appendix J for further details.

V.C. Related Outcomes

We also check whether treatment intensity affected other outcomes of interest related to the market price. First, we check whether treatment effects can be seen in farmgate prices

---

53. For all analyses in this article, we cluster our standard errors at the level of randomization. For the individual results shown in Section IV, this is at the group level. For the results presented in this section, which rely on the sublocation-level randomized saturation, we cluster at the sublocation level.

(as measured by self-reported prices reported by farmers in our household survey, rather than directly from our market surveys; see Online Appendix Table J.7). 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 postharvest period in year 2 (see Online Appendix Table J.8), a sensible demand response to the increase in price observed during a time when traders are typically purchasing.[54]

VI. 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 take advantage of these price variations? That is, how do the individual-level returns to arbitrage vary with the number of arbitrageurs?[55]

To answer this question, we revisit the individual results, re-estimating them to account for the variation in treatment density across sublocations. Table VII and Figure VII 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 (though the point estimate suggests that treated individuals in high-intensity areas may store a bit less than their counterparts in low-intensity areas).

---

54. 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 "High" in Table VI, column (2) (Y2) is almost exactly 65% of the coefficient on column (1) (Y1) (4.41 * 0.65 = 2.87 ≈ 2.85). The coefficient on "High Intensity * Month" in column 2 (Y2) is close to (a bit less than) 65% of the coefficient on column (1) (Y1) ($-0.57 * 0.65 = -0.37 \gtrsim -0.48$).

55. 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. Online 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.

TABLE VII

INVENTORY, NET REVENUES, AND HH CONSUMPTION (LOG) EFFECTS, ACCOUNTING FOR TREATMENT INTENSITY

	Inventory			Net revenues			Consumption		
	Y1 (1)	Y2 (2)	Pooled (3)	Y1 (4)	Y2 (5)	Pooled (6)	Y1 (7)	Y2 (8)	Pooled (9)
Treat	0.76***	0.55***	0.74***	1,060**	1,194	1,101**	0.01	−0.05	−0.01
	(0.19)	(0.18)	(0.15)	(438)	(685)	(430)	(0.04)	(0.04)	(0.02)
High	0.12	−0.03	0.02	534	−153	165	−0.00	−0.08	−0.05
	(0.36)	(0.22)	(0.24)	(551)	(559)	(480)	(0.05)	(0.05)	(0.04)
Treat * High	−0.33	−0.07	−0.29	−1,115*	−555	−817	−0.01	0.17***	0.07*
	(0.23)	(0.25)	(0.19)	(536)	(805)	(520)	(0.05)	(0.06)	(0.04)
Observations	3,836	2,944	6,780	3,795	2,935	6,730	3,792	2,944	6,736
Mean DV	2.74	1.38	2.04	−254	−3,620	−1,980	9.47	9.65	9.56
Std. dev. DV	3.50	2.53	3.12	5,383	6,990	6,478	0.63	0.60	0.62
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	.006	.015	.006	.864	.146	.408	.970	.006	.081
p-val Treat	.001	.010	.000	.028	.102	.021	.767	.228	.627
p-val Treat Boostrap	.000	.000	.000	.062	.120	.062	.768	.226	.616
p-val High	.731	.901	.945	.347	.789	.735	.962	.136	.295
p-val High Boostrap	.740	.900	.956	.336	.818	.706	.960	.120	.294
p-val Treat * High	.165	.802	.149	.054	.501	.136	.802	.007	.091
p-val Treat * High Boostrap	.198	.786	.150	.074	.508	.142	.796	.006	.084

Notes. The dependent variable in columns (1)–(3) is inventories, as measured by the number of 90-kg bags of maize held by the household. The dependent variable in columns (4)–(6) is net revenues, as measured by the value (in Ksh) of maize sales minus the value of maize purchases (the exchange rate during the study period ranged from 80 to 90 Kenyan shillings per US$). The dependent variable in columns (7)–(9) is HH consumption (measured in logged Ksh), aggregated from a detailed 30-day recall consumption module. "Treat" is an indicator for being in a treatment group. "High" intensity is an indicator for residing in a sublocation randomly assigned a high number of treatment groups. Regressions include round-year fixed effects and controls for survey date with errors clustered at the sublocation level. "Mean DV" and "Std. dev. DV" are the mean and standard deviation of the dependent variable among the control group. "*p*-val T + TH = 0" provides the *p*-value from an *F*-test that the sum of the Treat and Treat * High equal 0. To check robustness to small cluster standard error adjustments, *p*-values from the standard specification are compared to *p*-values from the wild bootstrap procedure proposed by Cameron et al. (2008), clustered at the sublocation level. Significant at 90% (*), 95% (**), 99% (***) confidence.
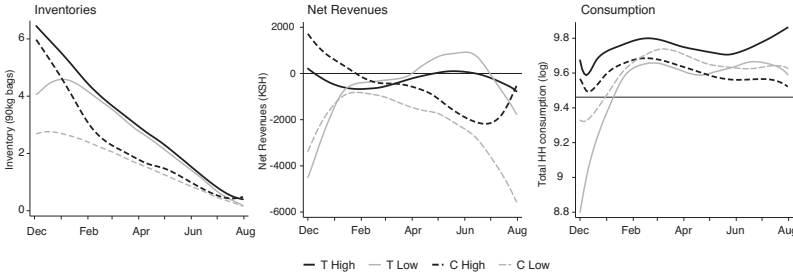
FIGURE VII

Pooled Treatment Effects by Treatment Intensity

Average inventories, net revenues, and log HH consumption over the study period in the treatment group versus the control group, split apart by high-intensity areas (gray lines) and low-intensity areas (black lines). Inventories are measured by the number of 90-kg bags of maize held by the household. Net revenues are the value (in Ksh) of maize sales minus the value of maize purchases. HH consumption (measured in log Ksh) is aggregated from a detailed 30-day recall household consumption module.

Turning to net revenues, we see much starker differences by treatment intensity. We find that treatment effects in low-intensity areas are much larger—roughly double—those estimated in the overall treatment specification in Table III. 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 high-intensity sublocations are substantially lower (and, in fact, are statistically indistinguishable from zero in the pooled results presented in Table VII, column (6)).[56] 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

---

56. Table VII 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. Although the interaction term "Treat * High" 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 * High" in year 2 (column (5)) 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.

earlier estimates, estimates for consumption remain relatively imprecisely estimated.[57]

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, GE effects drive down these differentials and therefore diminish the direct returns to arbitrage (and if this disincentivizes storage among high-intensity treated individuals, this may further diminish their total revenue gains from the program).[58]

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 VII and the estimate on the H_i dummy in Table VII, column (6)). However, it should be noted that this effect is measured with considerable noise and thus remains more speculative.[59] Given the diffuse nature of spillover effects, it is perhaps unsurprising that identifying these small effects with statistical precision is challenging.[60] However, they are suggestive of important distributional dynamics for welfare, which we explore below.

VI.A. Discussion

The randomized saturation design allows us to capture how both direct and indirect treatment effects vary with saturation level. Table VIII breaks down the distribution of welfare gains

---

57. 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 about or overinterpreting this fragile result.

58. In response to these price changes, which dampen the returns to arbitrage, farmers in high-intensity areas may have chosen to store less. The negative point estimate of "Treat * High" on inventory holdings in Table VII, column (3)— though not statistically significant—does suggest that treated individuals in high-intensity areas may store slightly less than their counterparts in low-intensity areas. This would also constitute a GE effect resulting from the price change, but an indirect one involving endogenous responses by farmers.

59. And even goes in the opposite direction in the year 2 results alone; see Table VII, column (5).

60. 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.

TABLE VIII

DISTRIBUTION OF GAINS IN THE PRESENCE OF GENERAL EQUILIBRIUM EFFECTS

	Low saturation	High saturation
1. Direct gains/HH (Ksh)	3,304	854
2. Indirect gains/HH (Ksh)	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 (Ksh)	816,984	422,248
7. Total indirect gains (Ksh)	0	1,757,880
8. Total gains (direct + indirect; Ksh)	816,984	2,180,128
9. Fraction of gains direct	1.00	0.19
10. Fraction of gains indirect	0.00	0.81

Notes. Calculations employ per-round point estimates on revenues β_1, β_2, and β_3 (coefficients on "Treat," "High," and "Treat ∗ High", respectively) from equation (5). These estimates are presented in Table VII, column (6) (in Ksh, multiplied by three to get the annual revenue gains; note the exchange rate during the study period ranged from 80 to 90 Kenyan shillings per US\$). 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 * High" interaction term in high-saturation areas. Indirect gains per household (row 2) are estimated as 0 in low-saturation areas and as the coefficient on "High" in high-saturation areas. The total gains from the intervention (row 8) include the direct gains that accrue to borrowers (row 6) and the indirect gains generated by GE effects (row 7). In high-saturation areas, 81% of the total gains are indirect gains (row 10). Therefore, only 19% of the gains can be captured by the private sector (row 9). Additional assumptions and calculation details are laid out in Online Appendix M. Note that while the private gains are greater at low saturation, the total gains are greater at high saturation.

from the loan, based on saturation rate and revenue effects drawn from the pooled results. Although this exercise takes all point estimates as given, note that some are less precisely measured than others.[61] As a result, there are probably large standard errors around some of the figures presented in Table VIII. This exercise should be interpreted as an illustration of how GE effects can shape the distribution of welfare gains in isolated markets, rather than precise quantitative estimates. Furthermore, we can only speak to the distribution of spillover effects within our sample (see Online Appendix M for additional discussion).

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 * High" interaction term in high-saturation areas). As discussed, we see that the direct treatment effects are greater for those in

---

61. For example, the point estimate on "Treat * High" is not quite significant at traditional levels, whereas the point estimate on "High" is measured with large noise.

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 "High" in high-saturation areas.[62] We see in the third row 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 the fifth and fourth rows, respectively), we find that the total size of the indirect gains swamp that of the direct gains in high-saturation areas (seventh and sixth rows, respectively). Note that this assumes the indirect gains as estimated among control subjects are equivalent to the indirect gains accruing to the rest of the population. While this assumption seems reasonable given that study subjects appear to be representative of households in this community (see Online Appendix Table L.1), even a conservative estimate assuming that zero gains accrue to populations outside of the study suggest that 50% of the gains are indirect (Online Appendix Table M.1).

These findings have two implications. First, the total gains from the intervention (presented in the eighth row) are much higher in high-saturation areas than they are in low-saturation areas. Although 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 our estimates of the indirect gains are imprecisely estimated, the qualitative result that higher saturation produces larger gains than low saturation holds even at indirect gains as low as 114 Ksh/household [US \$1.3], only 23% of the estimated effect).[63] High saturation offers greater relaxation of a barrier to intertemporal trade (credit constraints) and thereby produces larger aggregate gains.

62. 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.

63. 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.

Second, the distribution of gains shifts in the presence of GE effects. While in low-saturation areas all of the gains appear to come from direct gains (ninth row), in high-saturation areas, 81% of the total gains are indirect gains (tenth row).[64] GE 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.[65]

This redistribution of gains has implications for private sector investment in arbitrage. 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 is the direct (excludable) gains presented in the first row. The calculations in the sixth and eighth rows suggest that private sector financial institutions may face incentives that result in the underprovision of finance for arbitrage in these markets. Although overall social gains are higher at greater levels of saturation (eighth row), because much of these gains are indirect, private sector institutions will not be able to capture them. Instead, for private sector institutions, the available gains for capture are actually lower at high levels of saturation (sixth row).

The ninth and tenth rows attempt to quantify this disincentive. At low levels of saturation, private sector institutions can 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 19% of the total gains are direct and therefore excludable. Financial institutions therefore will fail to internalize 81% of the gains at these higher saturation levels, which will likely result in underprovision of financial products, compared to the socially optimal level.

---

64. It is possible that there are GE 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.

65. The spillover effects may not be evenly distributed; those who do little storage at baseline (typically poorer individuals) may benefit more, whereas those who do more storage at baseline (typically wealthier individuals) may be harmed. The spillover effects may therefore be redistributive toward the poor. See Online Appendix M for further discussion.

Given the imprecision in estimates in Table VIII, it is difficult to quantify the role of this mechanism in driving thin credit markets in developing countries. There are many other important factors at play, including the poor's inability to provide collateral—and the resulting difficulties financial institutions face in screening and monitoring these borrowers. However, our results suggest that the presence of positive spillovers may play a contributing role, exacerbating the underprovision of credit in rural and isolated markets.

VII. 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 counterintuitively 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: postharvest 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 GE 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, along with Fink, Jack, and Masiye (2018), our results are among the few experimental results to find a positive and significant effect of microcredit on the revenues of microenterprises; other studies have found either null results (see Banerjee 2013 for a review), or significant effects only among small subsets of the population (Meager 2019; Banerjee et al. 2018). This is also

to our knowledge one of the first experimental studies to directly account for GE effects in the microcredit literature. More broadly, we contribute to a small but growing literature experimentally estimating impacts on market prices (Angelucci and De Giorgi 2009; Imbert and Papp 2015; Muralidharan, Neihaus, and Sukhtankar 2018; Cunha, De Giorgi, and Jayachandran 2019). At least in our particular setting, failing to account for these spillover 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 the interventions for direct beneficiaries and to impact the welfare of nonbeneficiaries. Explicit attention to GE effects in future evaluations is probably 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 nonrecipient households alike. Welfare estimates in our setting suggest that a large portion of the benefits of expanded loan access could accrue indirectly to nonborrowers. 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—for example, 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, we record data on the timing of their marketing activities and storage behavior but 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).

STANFORD UNIVERSITY AND NATIONAL BUREAU OF ECONOMIC RESEARCH
UNIVERSITY OF MICHIGAN
UNIVERSITY OF CALIFORNIA, BERKELEY, AND NATIONAL BUREAU OF ECONOMIC RESEARCH

SUPPLEMENTARY MATERIAL

An Online Appendix for this article can be found at *The Quarterly Journal of Economics* online. Data and code replicating tables and figures in this article can be found in Burke, Bergquist, and Miguel (2018) in the Harvard Dataverse, doi:10.7910/DVN/C8UMQP.

REFERENCES

Acemoglu, Daron, "Theory, General Equilibrium and Political Economy in Development Economics," *Journal of Economic Perspectives*, 24 (2010), 17–32.

Aggarwal, Shilpa, Eilin Francis, and Jonathan Robinson, "Grain Today, Gain Tomorrow: Evidence from a Storage Experiment with Savings Clubs in Kenya," *Journal of Development Economics*, 134 (2018), 1–15.

Aker, Jenny C., "Rainfall Shocks, Markets and Food Crises: The Effect of Drought on Grain Markets in Niger," Center for Global Development Working Paper, 2008.

Angelucci, Manuela, and Giacomo De Giorgi, "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?," *American Economic Review*, 99 (2009), 486–508.

Angelucci, Manuela, Dean Karlan, and Jonathan Zinman, "Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco," *American Economic Journal: Applied Economics*, 7 (2015), 151–182.

Attanasio, Orazio, Britta Augsburg, Ralph De Haas, Emla Fitzsimons, and Heike Harmgart, "The Impacts of Microfinance: Evidence from Joint-Liability Lending in Mongolia," *American Economic Journal: Applied Economics*, 7 (2015), 90–122.

Baland, Jean-Marie, Catherine Guirkinger, and Charlotte Mali, "Pretending to Be Poor: Borrowing to Escape Forced Solidarity in Cameroon," *Economic Development and Cultural Change*, 60 (2011), 1–16.

Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman, "Labor Markets and Poverty in Village Economies," *Quarterly Journal of Economics*, 132 (2017), 811–870.

Banerjee, Abhijit, Emily Breza, Esther Duflo, and Cynthia Kinnan, "Do Credit Constraints Limit Entrepreneurship? Heterogeneity in the Returns to Microfinance," Buffett Institute Global Poverty Research Lab Working Paper No. 17-104, 2018.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan, "The Miracle of Microfinance?: Evidence from a Randomized Evaluation," MIT Working Paper, 2013.

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry, "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries," *Science*, 348 (2015), 761–772.

Banerjee, Abhijit V., and Esther Duflo, "Giving Credit Where It Is Due," *Journal of Economic Perspectives*, 24 (2010), 61–79.

Banerjee, Abhijit V., and Andrew F. Newman, "Occupational Choice and the Process of Development," *Journal of Political Economy*, (1993), 274–298.

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*, 5 (2013), 487–519.

Barrett, Christopher, "Displaced Distortions: Financial Market Failures and Seemingly Inefficient Resource Allocation in Low-Income Rural Communities," in *"Development Economics Between Markets and Institutions: Incentives for Growth, Food Security and Sustainable Use of the Environment,"* Erwin Bulte and Ruerd Ruben, eds. (Wageningen: Wageningen Academic Publishers, 2007).

Barrett, Christopher, and Paul Dorosh, "Farmers' Welfare and Changing Food Prices: Nonparametric Evidence from Rice in Madagascar," *American Journal of Agricultural Economics*, 78 (1996), 656–669.

Basu, Karna, and Maisy Wong, "Evaluating Seasonal Food Storage and Credit Programs in East Indonesia," *Journal of Development Economics*, 115 (2015), 200–216.

Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry, "Self-Selection into Credit Markets: Evidence from Agriculture in Mali," Working Paper, 2015.

Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, and Bertil Tungodden, "Human and Financial Capital for Microenterprise Development: Evidence from a Field and Lab Experiment," *Management Science*, 61 (2014), 707–722.

Bergquist, Lauren Falcao, "Pass-through, Competition, and Entry in Agricultural Markets: Experimental Evidence from Kenya," University of California Berkeley, 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*, 129 (2014), 697–752.

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

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

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar, "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico," *Journal of Political Economy*, 126 (2018), 635–687.

Bruhn, Miriam, and David McKenzie, "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 1 (2009), 200–232.

Brune, Lasse, Xavier Giné, Jessica Goldberg, and Dean Yang, "Facilitating Savings for Agriculture: Field Experimental Evidence from Malawi," *Economic Development and Cultural Change*, 64 (2016), 187–220.

Burke, Marshall, Lauren Falcao Bergquist, and Edward Miguel. "Replication Data for: 'Buy Low and Sell High: Arbitrage and Local Price Effects in Kenyan Markets'," Harvard Dataverse (2018), doi: 10.7910/DVN/C8UMQP.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 90 (2008), 414–427.

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

Cohen, Jessica, and Pascaline Dupas, "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment," *Quarterly Journal of Economics*, 125 (2010), 1–45.

Crépon, Bruno, Florencia Devoto, Esther Duflo, and William Pariente, "Estimating the Impact of Microcredit on Those Who Take It Up: Evidence from a Randomized Experiment in Morocco," *American Economic Journal: Applied Economics*, 7 (2015), 123–150.

Cunha, Jesse M., Giacomo De Giorgi, and Seema Jayachandran, "The Price Effects of Cash Versus In-Kind Transfers," *Review of Economic Studies*, 86 (2019), 240–281.

De Mel, Suresh, David McKenzie, and Christopher Woodruff, "Returns to Capital in Microenterprises: Evidence from a Field Experiment," *Quarterly Journal of Economics*, 123 (2008), 1329–1372.

Dillon, Brian, "Selling Crops Early to Pay for School: A Large-Scale Natural Experiment in Malawi," University of Washington, Working Paper, 2017.

Dupas, Pascaline, and Jonathan Robinson, "Why Don't the Poor Save More? Evidence from Health Savings Experiments," *American Economic Review*, 103 (2013), 1138–1171.

Fafchamps, Marcel, "Cash Crop Production, Food Price Volatility, and Rural Market Integration in the Third World," *American Journal of Agricultural Economics*, 74 (1992), 90–99.

Fafchamps, Marcel, David McKenzie, Simon Quinn, and Christopher Woodruff, "Microenterprise Growth and the Flypaper Effect: Evidence from a Randomized Experiment in Ghana," *Journal of Development Economics*, 106 (2013), 211–226.

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*, 103 (2013), 2196–2226.

Fink, Gunther, Kelsey Jack, and Felix Masiye. "Seasonal Credit Constraints and Agricultural Labor Supply: Evidence from Zambia," NBER Working Paper, 2018.

Galor, Oded, and Joseph Zeira, "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60 (1993), 35–52.

Imbert, Clément, and John Papp, "Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee," *American Economic Journal: Applied Economics*, 7 (2015), 233–263.

Kaboski, Joseph P., and Robert M. Townsend, "The Impact of Credit on Village Economies," *American Economic Journal: Applied Economics*, 4 (2012), 98.

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

Karlan, Dean, Ryan Knight, and Christopher Udry, "Hoping to Win, Expected to Lose: Theory and Lessons on Micro Enterprise Development," NBER Technical Report No. 18325, 2012.

Karlan, Dean, and Jonathan Morduch, "Access to Finance," *Handbook of Development Economics*, 5 (2010), 4703–4784.

Karlan, D., J. Morduch, and S. Mullainathan, "Take Up: Why Microfinance Take-Up Rates are Low and Why it Matters," Financial Access Initiative Technical Report, 2010.

Karlan, Dean, and Jonathan Zinman, "Microcredit in Theory and Practice: Using Randomized Credit Scoring for Impact Evaluation," *Science*, 332 (2011), 1278–1284.

McCloskey, Donald, and John Nash, "Corn at Interest: The Extent and Cost of Grain Storage in Medieval England," *American Economic Review*, 74 (1984), 174–187.

McKenzie, David, "Beyond Baseline and Follow-up: The Case for More T in Experiments," *Journal of Development Economics*, 99 (2012), 210–221.

McKenzie, David, and Christopher Woodruff, "Experimental Evidence on Returns to Capital and Access to Finance in Mexico," *World Bank Economic Review*, 22 (2008), 457–482.

Meager, Rachel, "Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments," *American Economic Journal: Applied Economics*, 11 (2019), 57–91.

Miguel, Edward, and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 72 (2004), 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*, 60 (1999), 467–495.

Muralidharan, Karthik, Paul Neihaus, and Sandip Sukhtankar, "General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India," NBER Working Paper, 2018.

Rubin, Donald B., "Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment," *Journal of the American Statistical Association*, 75 (1980), 591–593.

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

World Bank, "Malawi – Poverty and Vulnerability Assessment: Investing in our Future (Vol. 2): Full Report," 2007, Washington, DC: World Bank.