Selling low and buying high: An arbitrage puzzle in Kenyan villages

Marshall Burke*

March 20, 2014

Abstract

Large and regular seasonal price fluctuations in local grain markets appear to offer African farmers substantial inter-temporal arbitrage opportunities, but these opportunities remain largely unexploited: small-scale farmers are commonly observed to "sell low and buy high" rather than the reverse. In a field experiment in Kenya, we show that credit market imperfections limit farmers' abilities to move grain intertemporally, and that providing timely access to credit allows farmers to purchase at lower prices and sell at higher prices, increasing farm profits. To understand general equilibrium effects of these changes in behavior, we vary the density of loan offers across locations. We document significant effects of the credit intervention on seasonal price dispersion in local grain markets, and show that these GE effects strongly affect our individual level profitability estimates. In contrast to existing experimental work, our results indicate a setting in which microcredit can improve firm profitability, and suggest that GE effects can substantially shape estimates of microcredit's effectiveness.

JEL codes: D21, D51, G21, O13, O16, Q12

Keywords: storage; arbitrage; microcredit; credit constraints; agriculture

*Department of Agricultural and Resource Economics, UC Berkeley. Email: marshall.burke@berkeley.edu. I thank Ted Miguel, Lauren Falcao, Kyle Emerick, Jeremy Magruder, and Chris Barrett for useful discussions, and thank seminar participants at Berkeley, Stanford, Kellogg, and the Pacific Development Conference for useful comments. I also thank Peter LeFrancois and Innovations for Poverty Action for excellent research assistance in the field, and One Acre Fund for partnering with us in the intervention. I gratefully acknowledge funding from the Agricultural Technology Adoption Initiative and an anonymous donor. All errors are my own.

1 Introduction

Imperfections in credit markets are generally considered to play a central role in underdevelopment (Banerjee and Newman, 1993; Galor and Zeira, 1993; Banerjee and Duflo, 2010). These imperfections are thought to be particularly consequential for small and informal firms in the developing world, and for the hundreds of millions of poor people who own and operate them. This thinking has motivated a large-scale effort to expand credit access to existing or would-be microentrepreneurs around the world, and it has also motivated a subsequent attempt on the part of academics to rigorously evaluate the effects of this expansion on the productivity of these microenterprises and on the livelihoods of their owners.

Findings in this rapidly growing literature have been remarkably heterogenous. Studies that provide cash grants to households and to existing small firms suggest high rates of return to capital in some settings but not in others.[1] Further, experimental evaluations of traditional microcredit products (small loans to poor households) have generally found that individuals randomly provided access to these products are subsequently no more productive on average that those not given access, but that subsets of recipients often appear to benefit.[2]

In this paper, I study a unique microcredit product designed to improve the profitability of small farms – a setting that has been outside the focus of most of the experimental literature on credit constraints. Farmers in our setting in Western Kenya, as well as throughout much of the rest of the developing world, face large and regular seasonal fluctuations in grain prices, with increases of 50-100% between post-harvest lows and pre-harvest peaks common in local markets (as described in more detail below). Nevertheless, most of these farmers have difficulty using storage to move grain from times of low prices to times of high prices, and this inability appears at least in part due to limited borrowing opportunities: lacking access to credit or savings, farmers report selling their grain at low post-harvest prices to meet urgent cash needs (e.g., to pay school fees). To meet

[1] Studies finding high returns to cash grants include De Mel, McKenzie, and Woodruff (2008); McKenzie and Woodruff (2008); Fafchamps et al. (2013); Blattman, Fiala, and Martinez (2013). Studies finding much more limited returns include Berge, Bjorvatn, and Tungodden (2011) and Karlan, Knight, and Udry (2012).

[2] Experimental evaluations of microcredit include Attanasio et al. (2011); Crepon et al. (2011); Karlan and Zinman (2011); Banerjee et al. (2013); Angelucci, Karlan, and Zinman (2013). See Banerjee (2013) and Karlan and Morduch (2009) for nice recent reviews of these literatures.

consumption needs later in the year, many then end up buying back grain from the market a few months after selling it, in effect using the maize market as a high-interest lender of last resort (Stephens and Barrett, 2011).

Working with a local agricultural microfinance NGO, I offer randomly selected smallholder maize farmers a loan at harvest, and study whether access to this loan improves their ability to use storage to arbitrage local price fluctuations, relative to a control group. To understand the importance of credit timing in this setting, half of these offers were for a loan immediately after harvest (October), and half for a loan three months later (January). Furthermore, because storage-related changes in behavior could have effects on local prices in a setting of high regional transport costs, I vary the density of treated farmers across locations and track market prices at 50 local market points. Finally, to help bind my hands against data mining (Casey, Glennerster, and Miguel, 2012), I registered a pre-analysis plan prior to the analysis of any follow-up data (see Section 3.1).

Despite a seasonal price rise that was in the left tail of both the historical distribution of local price fluctuations and the distribution (across farmers) of the expected price rise for the study year, I find statistically significant and economically meaningful effects of the loan offer on farm profitability, *but only for farmers in low-treatment-density areas.* On average, farmers offered the loan sold significantly less and purchased significantly more maize in the period immediately following harvest, and this pattern reversed during the period of (typically) high prices 6-9 months later. This change in marketing behavior had discernible effects on prices in local maize markets: prices immediately after harvest were significantly higher in areas with high treatment density, but were lower (although not significantly so) by the end of the study period. Consistent with these differential price effects, I find that while treated farmers in high-density areas stored significantly more than their control counterparts, they were not more profitable; the reduction in seasonal price dispersion in these area reduced the benefits of loan adoption. Conversely, treated farmers in low-density areas have both significantly higher inventories and significantly higher profits relative to control. I find some evidence that the timing of credit matters, with inventories and profits uniformly higher in the treatment group who received the earlier loan, but these results are not always significant.

Why do I find positive effects on firm profitability when other experimental studies on microcredit do not? These studies have offered a number of explanations as to why improved access to capital does not appear beneficial on average. First, many small businesses or potential microentrepreneurs simply might not actually face profitable investment opportunities (Banerjee et al., 2013; Fafchamps et al., 2013; Karlan, Knight, and Udry, 2012; Banerjee, 2013).[3] Second, profitable investment opportunities could exist but established or potential microentrepreneurs might lack either the skills or ability to channel capital towards these investments - e.g. if they lack managerial skills (Berge, Bjorvatn, and Tungodden, 2011; Bruhn, Karlan, and Schoar, 2012), or if they face problems of self-control or external pressure that redirect cash away from investment opportunities (Fafchamps et al., 2013). Third, typical microcredit loan terms require that repayment begin immediately, and this could limit investment in illiquid but high-return business opportunities (Field et al., 2012).

Finally, general equilibrium effects of credit expansion could alter individual-level treatment effect estimates in a number of ways, potentially shaping outcomes for treated individuals (e.g. if microenterprises are dominated by a very small number of occupations and credit-induced expansion of these business bids away profits) as well as for non-recipients (e.g. through increased demand for labor (Buera, Kaboski, and Shin, 2012)). This is a recognized but unresolved problem in the experimental literature on credit, and few experimental studies have been explicitly designed to quantify these effects.[4]

All of these factors likely help explain why our results diverge from existing estimates. Unlike most of the settings examined in the literature, using credit to "free up" storage for price arbitrage

[3]For example, many microenterprises might have low efficient scale and thus little immediate use for additional investment capital, with microentrepreneurs then preferring to channel credit toward consumption instead of investment. Relatedly, marginal returns to investment might be high but total returns low, with the entrepreneur making the similar decision that additional investment is just not worth it.

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

does not require starting or growing a business among this population of farmers, is neutral to the scale of farm output, does not appear to depend on entrepreneurial skill (all farmer have stored before, and all are very familiar with local price movements), and does not require investment in a particularly illiquid asset (inventories are kept in the house and can be easily sold). 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 purchases from times of high prices to times of low prices.

Furthermore, our results also suggest that – at least in our rural setting – treatment density matters and market-level spillovers can substantially shape individual-level treatment effect estimates. Whether these GE also influnced estimated treatment effects in more urban settings is unknown, although there is some evidence that spillovers do matter for microenterprises who directly compete for a limited supply of inputs to production.[5] In any case, my results suggest that explicit attention to GE effects in future evaluations of credit market interventions is likely warranted.

Beyond contributing to the experimental literature on microcredit, my paper is closest to a number of recent papers that examine the role of borrowing constraints in households' storage decisions and seasonal consumption patterns. Using secondary data from Kenya, Stephens and Barrett (2011) also suggest that credit constraints substantially alter smallholder farmers' marketing and storage decisions, and Basu and Wong (2012) show that allowing farmers to borrow against future harvests can substantially increase lean-season consumption. As in these papers, my results show that when borrowing and saving are difficult, households turn to increasingly costly ways to move consumption around in time. In my particular setting, credit constraints combined with post-harvest cash needs cause farmers to store less than they would in an unconstrained world, lowering farm profits even in a year when prices don't rise much. In this setting, even a relatively modest expansion of credit affects local market prices, to the apparent benefit of those with and without access to this credit.

Finally, my results speak to an earlier literature showing how credit market imperfections can combine with other features of economies to generate observed broad-scale economic patterns

[5] See De Mel, McKenzie, and Woodruff (2008) and their discussion of returns to capital for firms in the bamboo sector, all of whom in their setting compete over a limited supply of bamboo.

(Banerjee and Newman, 1993; Galor and Zeira, 1993). These earlier papers showed how missing markets for credit, coupled with an unequal underlying wealth distribution, could generate large-scale patterns of occupational choice. I show that missing markets for credit combined with climate-induced seasonality in rural income can help generate widely-observed seasonal price patterns in rural grain markets, patterns that appear to further worsen poor households' abilities to smooth consumption across seasons. That expansion of credit access appears to help reduce this price dispersion suggests an under-appreciated but likely substantial additional benefit of credit expansion in rural areas.

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

2 Setting and experimental design

2.1 Arbitrage opportunities in rural grain markets

Seasonal fluctuations in prices for staple grains appear to offer substantial intertemporal arbitrage opportunities, both in our study region of East Africa as well as in other parts of Africa and elsewhere in the developing world. While long term price data unfortunately do not exist for the small markets in very rural areas where our experiment takes place, price series are available for major markets throughout the region. Average seasonal price fluctuations for maize in available markets are shown in Figure 1. Increases in maize prices in the six to eight months following harvest average roughly 25-50% in these markets, and these increases appear to be a lower bound on seasonal price increases reported elsewhere in Africa.[6]

These increases also appear to be a lower bound on typical increase observed in the smaller

[6]For instance, Barrett (2008) reports seasonal rice price variation in Madagascar of 80%, World Bank (2006) reports seasonal maize price variation of about 70% in rural Malawi, and Aker (2012) reports seasonal variation in millet prices in Niger of 40%.

markets in our study area, which (relative to these much larger markets) are characterized with much smaller "catchments" and less outside trade. We asked farmers at baseline to estimate average monthly prices for either sales or purchases of maize at their local market point over the last five years, and as shown in the left panel of Figure 3, they reported a typical doubling in price between September (the main harvest month) and the following June. 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% (with a 25th percentile of 60% increase and 75th percentile of 118% - results available on request).

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

Why are farmers not using storage to sell at higher prices and purchase at lower prices? Our experiment will primarily be designed to test the role of credit constraints in shaping storage and marketing decisions, and here we talk through why credit might matter (these explanations will be formalized in a future draft). First, and most simply, 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 were not storing and selling maize at higher prices. In particular, because early all of these farm households have school aged kids, and a large percentage of a child's school fees are typically due in the few months after harvest (prior to January enrollment), many farmers report selling much of their harvest to pay these fees. Indeed, many schools in the area will accept in-kind payment in maize during this period. Farmers also report having to pay other bills they have accumulated throughout the year during the post-harvest period.

Second, as with poor households throughout much of the world, these farmers appear to have very limited access to formal credit. Only eight percent of households in our sample reported having taking a loan from a bank in the year prior to the baseline survey. 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 3 months before the baseline.

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

Nevertheless, there could be other reasons beyond credit constraints why farmer are not taking advantage of apparent arbitrage opportunities. The simplest explanations are that farmers do not know about the price increases, or that it's actually not profitable to store – i.e. arbitrage opportunities are actually much smaller than they appear because storage is costly. These costs could come in the form of losses to pests or moisture-related rotting, or they could come in the form of "network losses" to friends and family, since maize is stored in the home and is visible to friends and family, and there is often community pressure to share a surplus. Third, farmers could be highly impatient and thus unwilling to move consumption to future periods in any scenario. Finally, farmers might view storage as too risky an investment.

Evidence from pilot and baseline data, and from elsewhere in the literature, argues against a few of these possibilities. We can immediately rule out an information story: as shown in Figure 3 and discussed above, all farmers know exactly what prices are doing, and all expect prices to rise substantially throughout the year.[7] Second, pest-related losses appear surprisingly low in our setting, with farmers reporting losses from pests and moisture-related rotting of less than 5% for maize stored for six to nine months. Similarly, the fixed costs associated with storing for these farmers are small and have already been paid: all farmers store at least some grain (note the positive initial inventories in Figure A.1), and grain in simply stored in the household or in small sheds previously built for the purpose. Third, existing literature shows that for households that are both consumers and producers of grain, aversion to price risk should motivate *more* storage rather than less: the worst state of the world for these households is a huge price spike during the lean season, which should motivate "precautionary" storage (Saha and Stroud, 1994; Park, 2006). Fourth, while we cannot rule out impatience as a driver of low storage rates, extremely high discount rates would be needed to rationalize this behavior in light of the expected nine-month doubling of prices. Furthermore, farm households are observed to make many other investments with payouts far in the future (e.g. school fees), meaning that rates of time preference would also have to differ substantially across investments and goods.

Costs associated with network-related losses appear a more likely explanation for an unwillingness to store substantial amounts of grain. Existing literature suggests that community pressure is one explanation for limited informal savings (Dupas and Robinson, 2013; Brune et al., 2011), and in focus groups farmers often told us something similar about stored grain (itself a form of savings). As described below, our main credit intervention might also provide farmers a way to shield stored maize from their network, and we added a small additional treatment arm to determine whether this shielding effect is substantial on its own.

2.2 Experimental design

Our study sample is drawn from existing groups of One Acre Fund (OAF) farmers in Webuye district, Western Province, Kenya. OAF is a microfinance NGO that makes in-kind, joint-liability loans of fertilizer and seed to groups of farmers, as well as providing training on improved farming

[7] The mean across farmers for all three reported prices (the historical purchase price, the historical sales price, and the expected sales price) is a 115-134% increase in prices. For the expected sales price over the ensuing nine months after the September 2012 baseline, the 5th, 10th, and 25th percentiles of the distribution are a 33%, 56%, and 85% increase, respectively, suggesting that nearly all farmers in our sample expect substantial price increases.

techniques. OAF group sizes typically range from 8-12 farmers, and farmer groups are organized into "sublocations" – effectively clusters of villages that can be served by one OAF field officer. OAF typically serves 20-30% of farmers in a given sublocation.

As noted above, extensive focus groups with OAF farmers in the area prior to the experiment suggested that credit constraints likely play a substantial role in smallholder marketing decisions in the region. These interviews also offered three other important pieces of information. First, farmers were split on when exactly credit access would be most useful, with some preferring cash immediately at harvest, and some preferring it a few months later and timed to coincide exactly with when some of them had to pay school fees. This in turn suggested that farmers were sophisticated about potential difficulties in holding on to cash between the time it was disbursed and the time it needed to be spent, and indeed many farmers brought these difficulties up directly in interviews. Third, OAF was willing to offer the loan at harvest if it was collateralized with stored maize, and collateralized bags of maize would be 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 collateral and bag tagging, many farmers (again 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.[8]

We allowed this information to inform the experimental design. First, we offer some randomly selected farmers a loan to be made available in October 2012 (immediately after harvest), and some a loan to be available January 2013. Both loan offers were announced in September 2012. To qualify for the loan, farmers had to commit maize as collateral, and the size of the loan they could qualify for was a linear function of the amount they were willing to collateralize (capped at 7 bags). To account for the expected price increase, October bags were valued at 1500Ksh, and January bags at 2000Ksh. Each loan carried with it a "flat" interest rate of 10%, with full repayment due after nine months.[9] So a farmer who committed 5 bags when offered the October loan would receive

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

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

5*1500 = 7500Ksh in cash in October (∼$90 at current exchange rates), and would be required to repay 8250Ksh by the end of July. 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. As mentioned, each collateralized bag is given a tag with the OAF logo, and is closed with a simple plastic zip-tie by a loan officer, who then disburses the cash.

As discussed above, the tags could represent a meaningful treatment in their own right. To attempt to separate the effect of the credit from any effect of the tag, a separate treatment group received only the tags.[10] Finally, because self- or other-control problems might make it particularly difficult to channel cash toward productive investments in settings where there is a substantial time lag between when the cash is delivered and when the desired investment is made, we 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 were offered a savings lockbox (a simple metal box with a sturdy lock) which they could use as they pleased. While such a savings device could have other effects on household decision making, our thinking was that it would be particularly helpful for loan clients who received the cash before it was needed.

Our sample consists of 240 existing OAF farmer groups drawn from 17 different sublocations in Webuye district, and our total sample size at baseline was 1589 farmers. Figure 2 shows the basic setup of our experiment. There are three levels of randomization. First, we randomly divided the 17 sublocations in our sample into 9 "high" treatment intensity sites and 8 "low" treatment density sites, fixed the "high" treatment density at 80% (meaning 80% of groups in the sublocation would be offered a loan), and then determined the number of groups that would be needed in the "low" treatment sites in order to get our total number of groups to 240 (what the power calculations suggested we needed to be able to discern meaningful impacts at the individual level). This resulted in a treatment intensity of 40% in the "low" treatment-intensity sites, yielding 171 total treated groups in the high intensity areas and 69 treated groups in the low intensity areas.

Second, the October (T1) and January (T2) loan offers were randomized at the group level. The

rural lender in Kenya.

[10]This is of course not perfect – there could be an interaction between the tag and the loan – but we did not think we had the sample size to do the full 2 x 2 design to isolate any interaction effect.

loan treatments were then stratified at the sublocation level and then on group-average OAF loan size in the previous year (using administrative data). Although all farmers in each loan treatment group were offered the loan, we follow only a randomly selected 6 farmers in each loan group, and a randomly selected 8 farmers in each of the control groups (whether or not they actually adopted the loan).

Finally, as shown at the bottom of Figure 2, the tags and lockbox treatments were randomized at the individual level. Using the sample of individuals randomly selected to be followed in each group, we stratified individual level treatments by group treatment assignment and by gender. So, for instance, of all of the women who were offered the October Loan and who were randomly selected to be surveyed, one third of them were randomly offered the lockbox (and similarly for the men and for the January loan). In the control groups, in which we were following 8 farmers, 25% of the men and 25% of the women were randomly offered the lockbox (Cl in Figure 2), with another 25% each being randomly offered the tags (Ct). The study design allows identification of the individual and combined effects of the different treatments, and our approach for estimating these effects is described below.

3 Data and estimation

The timing of the study activities is shown in Figure A.2. We collect 3 types of data. Our main source of data is farmer household surveys. All study participants were baselined in August/September 2012, and we undertook 3 follow-up rounds over the ensuing 12 months, with the last follow-up round concluding August 2013. The multiple follow-up rounds were motivated by three factors. First, a simple inter-temporal model of storage and consumption decisions suggests that while the loan should increase total consumption across all periods, the per-period effects could be ambiguous – meaning that consumption throughout the follow-up period needs to be measured to get at overall effects. Second, because nearly all farmers deplete their inventories before the next harvest, inventories measured at a single follow-up one year after treatment would likely provide very little information on how the loan affected storage and marketing behavior. Finally, as shown in McKenzie (2012), multiple follow-up measurements on noisy outcomes variables (e.g consumption) has the added advantage of increasing power.

The follow-up survey rounds span the spring 2013 "long rains" planting (the primary growing season), and concluded just prior to the 2013 long rains harvest. The baseline survey collected data on farming practices, on storage costs, on maize storage and marketing over the previous crop year, on price expectations for the coming year, on food and non-food consumption expenditure, on household borrowing, lending, and saving behavior, on household transfers with other family members and neighbors, on sources of non-farm income, on time and risk preferences, and on digit span recall. The follow-up surveys collected similar data, tracking 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.

Our two other sources of data are monthly price surveys at 52 market points in the study area (which we began in November 2012 and continued through August 2013), and loan repayment data from OAF administrative records that was generously shared by OAF. 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.

Table 1 shows summary statistics for a range of variables at baseline, and shows balance of these variables across the three main loan treatment groups. Groups are well balanced, as would be expected from randomization. Table A.1 shows the analogous table comparing individuals in the high- and low-treatment-density areas; samples appear balanced on observables here as well. Attrition was also relatively low across our survey rounds: 8% overall, and not significantly different across treatment groups (8% in T1, 9% in T2, 7% in C).

3.1 Pre-analysis plan

To limit both risks and perceptions of data mining and specification search (Casey, Glennerster, and Miguel, 2012), I specified and registered a pre-analysis plan (PAP) prior to the analysis of any follow-up data.[11] Both the PAP and the complete set of results are available upon request.

I deviate significantly from the PAP in one instance: as described below, it became clear that

[11] The pre-analysis plan is registered here: https://www.socialscienceregistry.org/trials/67, and was registered on September 6th 2013.

my method for estimating market-level treatment effects specified in the pre-analysis plan could generate biased estimates, and here I pursue an alternate strategy that more directly relies on the randomization. In two other instances I add to the PAP. First, in addition to the regression results specified in the PAP, I also present graphical results for many of the outcomes. These results are just based on non-parametric estimates of the parametric regressions specified in the PAP, and are included because they clearly summarize how treatment effects evolve over time, but since they were not mentioned in the PAP I mention them here. Second, I failed to include in the PAP the (obvious) regressions in which the individual-level treatment effect is allowed to vary by the sublocation-level treatment intensity. I hope the reader will interpret this oversight, and the subsequent inclusion of these regressions in what follows, as shortsightedness on the part of the author rather than malintent.

3.2 Estimation of treatment effects

We have three main outcomes of interest: inventories, maize net revenues, and consumption. Inventories are the number of bags the household had in their maize store at the time of the each survey. This amount is visually verified by our enumeration team, and so is likely to be measured with very little error. We define maize net revenues as the value of all maize sales minus the value of all maize purchases, and minus any additional interest payments made on the loan for individuals in the treatment group. We call this "net revenues" rather than "profits" since we likely do not observe all costs; nevertheless, costs are likely to be very similar across treatment groups (fixed costs 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 per capita household expenditure over the 30 days prior to each survey. For each of these variables we trim the top and bottom 0.5% of observations, as specified in the pre-analysis plan.

We have one baseline and three follow-up survey rounds, allowing a few different alternatives for estimating treatment effects. Pooling treatments for now, denote T_j as an indicator for whether group j was assigned to treatment, and y_{ijr} as the outcome of interest for individual i in group j

in round $r \in (0, 1, 2, 3)$, with $r = 0$ indicating the baseline. Following McKenzie (2012), our main specification pools data across follow-up rounds 1-3:

$$Y_{ijr} = \alpha + \beta T_j + \phi Y_{ij0} + \eta_r + \varepsilon_{ijr} \tag{1}$$

where Y_{ij0} is the baseline measure of the outcome variable. The coefficient β estimates the Intent-to-Treat and, with round fixed effects η_r, 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. Standard errors will be clustered at the group level.

In terms of additional controls, we follow advice in Bruhn and McKenzie (2009) and include stratification dummies as controls in our main specification. Similarly, controlling linearly for the baseline value of the covariate generally provides maximal power (McKenzie, 2012), but because many of our outcomes are highly time-variant (e.g. inventories) the "baseline" value of these outcomes is somewhat nebulous. As discussed below, for our main outcomes of interest that we know to be highly time varying (inventories and net revenues), we control for the number of bags harvested during the 2012 LR; this harvest occurred pre-treatment, and it will be a primary determinant of initial inventories, sales, and purchases. For other variables like total household consumption expenditure, we control for baseline measure of the variable. Finally, to absorb additional variation in the outcomes of interest, we also control for survey date in the regressions; each follow-up round spanned 3+ months, meaning that there could be (for instance) substantial within-round drawdown of inventories. Inclusion of all of these exogenous controls should help to make our estimates more precise without changing point estimates, but as robustness we will re-estimate our main treatment effects with all controls dropped.

The assumption in (1) is that treatment effects are constant across rounds. In our setting, there are reasons why this might not be the case. In particular, the first follow-up survey began in November 2012 and ended in February 2013, meaning that it spanned the rollout of the January 2013 loan treatment (T2). This means that the loan treatment might not have had a chance to affect outcomes for some of the individuals in the T2 group by the time the first follow-up was conducted (although, to qualify for the T2 loan, households would have needed to hold back inventory, such that inventory effects could have already occurred). Similarly, if the benefits of having more inventory on hand become much larger in the period when prices typically peak (May-July), then treatment effects could be larger in later rounds. To explore whether treatment effects are constant across rounds, we estimate:

$$Y_{ijr} = \sum_{r=1}^{3} \beta_r T_j + \phi Y_{ij0} + \eta_r + \varepsilon_{ijr} \tag{2}$$

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

To quantify market level effects of the loan intervention, we tracked market prices at 52 market points throughout our study region, and we assign these markets to the nearest sublocation. We begin by estimating the following linear model[12]:

$$y_{mst} = \alpha + \beta_1 H_s + \beta_2 month_t + \beta_3 (H_s * month_t) + \varepsilon_{mst} \tag{3}$$

where y_{mst} represents the maize sales price at market m in sublocation s in month t. H_s is a dummy for if sublocation s is a high-intensity sublocation, and $month_t$ is a time trend (Nov = 1, Dec = 2, etc). If access to the storage loan allowed farmers to shift purchases to earlier in the season or sales to later in the season, and if this shift in marketing behavior was enough to alter supply and demand in local markets, then our prediction is that $\beta_1 > 0$ and $\beta_3 < 0$, i.e. that prices in areas with more treated farmers are higher after harvest but lower later in the year.

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 are only randomizing across 17 sublocations. This relatively small number of clusters could present problems for inference (Cameron, Gelbach, and Miller, 2008).

[12] This estimating equation is slightly different than what was proposed in the pre-analysis plan. As was energetically pointed out to the author during a seminar presentation at Berkeley after the pre-analysis plan had been registered, the proposed estimating equation for quantifying market level effects (which relied on counting up the number of treated farmers) could produce biased estimates because we are in practice unable to control for the total number of farmers in the area. Using the randomization dummy avoids this worry.

We begin by clustering errors at the sublocation level when estimating (3). Future versions of the will also report standard errors estimated using both the wild bootstrap technique described in Cameron, Gelbach, and Miller (2008), and the randomization inference technique (e.g. as used by Cohen and Dupas (2010)).

Finally, 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:

$$Y_{ijsr} = \alpha + \beta_1 T_j + \beta_2 H_s + \beta_3 (T_j * H_s) + \phi Y_{ij0} + \eta_r + \varepsilon_{ijsr} \tag{4}$$

If the intervention produces enough individual level behavior to have market effects, we predict that $\beta_3 < 0$ and perhaps that $\beta_2 > 0$ - i.e. treated individual in high-density areas do worse than in low density areas, and control individuals in high density areas do better (due to higher initial prices at which they'll be selling their output). As in Equation 3, we will report results with errors clustered at the sublocation level.

4 Individual level results

4.1 Take up

Take-up of the loan treatments was quite high. Of the 474 individuals in the 77 groups assigned to the October loan treatment (T1), 329 (69%) applied and qualified for the loan. For the January loan treatment (T2), 281 out of the 480 (59%) qualified for and took up the loan. Unconditional loan sizes in the two treatment groups were 5294 Ksh and 4345 Ksh (or about $62 and $51 USD) for T1 and T2, respectively, and we can reject at 99% confidence that the loan sizes were the same between groups. The average loan sizes conditional on take-up were 7627Ksh (or about $90 USD) for T1 and 7423Ksh (or $87) for T2, and in this case we cannot reject that conditional loan sizes were the same between groups.

Relative to many other credit-market interventions in low-income settings in which documented take-up rates range from 1-10% of the surveyed population (Karlan, Morduch, and Mullainathan, 2010), the 60-70% take-up rates of our loan product were extraordinarily high. This is perhaps not surprising given that our loan product was offered as a top-up for individuals who were already clients of an MFI. Nevertheless, OAF estimates that 20-30% of farmers in a given village in our study area enroll in OAF, which implies that even if *no* non-OAF farmers were to adopt the loan if offered it, population-wide take-up rates of our loan product would still exceed 10-20%.

4.2 Overall price increase

I 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. The first thing to note, before turning to these results, is the small average price increase that occurred during our study year, both relative to what farmers (and traders) reported had occurred in the recent past, and relative to what was expected for the study year. As shown in the right panel of Figure 3, farmers had expected a doubling of prices, but prices only increased by 20-30% and peaked 2-3 months earlier than normal. We currently do not know why this is – prices in larger surrounding markets were also flat – but we are currently conducting interviews with local traders to try to understand why this year might have been different. In any case, the rather small price rise is going to substantially shape the returns to holding inventories relative to a more "normal" year.[13]

4.3 Effect of the loan offer

Table 2 and Figure 4 and show the results of estimating Equations 1 and 2 on the pooled treatment indicator, either parametrically (in the table) or non-parametrically (in the figure). The top panels in Figure 4 show the means in each treatment group over time for our three main outcomes of interest (as estimated with fan regressions), and the bottom panels show the difference in treatment minus control over time, with the 95% confidence interval calculated by bootstrapping the fan regression 1000 times.

Farmers responded to the intervention as anticipated. They held significantly more inventories

[13] Consequently, we are running the experiment for another year, hoping to get a more "normal" price draw.

for much of the year, on average about 20% more than the control group mean (Column 1 in Table 2), and net revenues were significantly lower immediately post harvest and significantly higher later in the year (Column 6 in Table 2 and middle panel of Figure 4). The net effect on revenues averaged across the year was positive but not significant (Column 5), and the effect size is rather small: the total effect across the year can be calculated by adding up the coefficients in Column 6, which yields an estimate of 780Ksh, or about $10 at current exchange rates. Given these rather small effects, it is not surprising that the effects on per capita consumption are positive but also small and not significant.

Splitting apart the two loan treatment arms, the results provide some evidence that the timing of the loan affects the returns to capital in this setting. As shown in Figure 5 and Table 3, point estimates suggest that those offered the October loan held more in inventories, reaped more in net revenues, and had higher overall consumption. Overall effects on net revenues are about twice as high as pooled estimates, and are now significant at the 5% level (Column 5 of Table 3), and we can reject that treatment effects are equal for T1 and T2 ($p = 0.04$). Figure 6 shows non-parametric estimates of differences in net revenues over time among the different treatment groups. Seasonal differences are again strong, and particularly strong for T1 versus control.

Why might the October loan have been more effective than the January loan? Note that while we are estimating the intent-to-treat (ITT) and thus that differences in point estimates could in principle be driven by differences in take-up, these latter differences are probably not large enough to explain the differential effects. For instance, "naive" average treatment effect estimates that rescale the ITT coefficients by the take-up rates (70% versus 60%) still suggest substantial differences in effects between T1 and T2. A more likely explanation is that the January loan came too late to be as useful: farmers in the T2 group were forced to liquidate some of their inventories before the arrival of the loan, and thus had less to sell in the months when prices rose. This would explain why inventories began lower, and why T2 farmers appear to be selling more during the immediate post-harvest months than T1 farmers. Nevertheless, they sell less than control farmers during this period and store more, likely because qualifying for the January loan meant carrying sufficient inventory until that point.

Finally, we test whether loan treatment effects are actually being driven by the tags. Estimates are shown in Table A.2. Point estimates are larger across the board for the pooled and T1 groups than for the tags-alone group, but estimates are somewhat noisy, and only for inventories and for T1 revenues can we reject that the effect of the loan was driven by the tags.

5 General equilibrium effects

The experiment was designed to quantify one particular potential general equilibrium effect: the effect of the loan intervention on local maize prices. Such effects appeared plausible for three reasons. First, OAF serves a substantial number of farmers in a given area. In "mature" areas where OAF has been working for a number of years (such as Webuye district where our experiment took place), typically 30% of all farmers sign up for OAF. This means that in high treatment density areas, where 80% of OAF groups were enrolled in the study and 2/3rds of these offered the loan, roughly 10% of the population of farmers took the loan.Second, focus groups had suggested take up of the loan would be quite high, and that farmers did not need to be told that they could make extra money by storing longer. Finally, while we lack long-term price data for local markets in the are, there is some evidence that these markets are not well integrated. In particular, a handful of traders can be found in these markets on the main market day, and in interviews these traders report making substantial profits engaging in spatial arbitrage across these markets, often selling in markets they will later purchase from (and vice versa). This provides some evidence that these markets might be affected by local shifts in supply and demand.[14]

How large might these market price effects be? As a simple calibration, I assume that prices in a given market are set locally – i.e. affected only by local supply and demand. Re-arranging log-log supply and demand equations provides a simple expression for how our treatment-induced change in supply might affect local prices:

$$\%\Delta p_t = \frac{\%\Delta q_t}{\varepsilon_d - \varepsilon_s} \tag{5}$$

[14]Other papers, such as Cunha, De Giorgi, and Jayachandran (2011), find substantial effects of local supply shocks on local prices in settings (in this case, Mexico) where markets are likely much less isolated than ours.

The numerator on the right-hand side is the differential change in total supply between high and low density areas in a given period t. This can be calculated by combining our inventory treatment effect estimates with data on differences in market-level treatment saturation between high- and low-density areas. We calculate a peak inventory effect (i.e. inward supply shift) of about 15% for the December-January months, and estimate that this treatment effect would have been experienced by 5% more of the population in high density areas than in low density areas.[15] Then using estimates of demand and supply elasticities for staple grains in rural Africa derived from the literature ($\varepsilon_d = -0.25$, $\varepsilon_s = 0.1$), we estimate that the peak price difference around December/January would be on the order of 2%.

5.1 Market level effects

To understand the effect of our loan intervention on local maize prices, we identified 52 local market points spread throughout our study area that OAF staff indicated were where their clients typically bought and sold maize, and our enumerators tracked monthly maize prices at these market points. We then match these market points to the OAF sublocation in which they fall. "Sublocations" here are simply OAF administrative units that are well defined in terms of client composition (i.e. which OAF groups are in which sublocation), but less well defined in terms of their exact geographic boundaries. Given this, we match markets to sublocations in two ways: by using administrative estimates of which markets fall in which sublocations (i.e. asking OAF field staff which markets are in their sublocation), and by using GPS data on both the market location and the location of farmers in our study sample to calculate the "most likely" sublocation, based on the designated sublocation to which the majority of nearby farmers belong. In practice, these two methods provided very similar matches, but we show estimates using both approaches for robustness.

We then utilize the sublocation-level randomization in treatment intensity to identify market-level effects of our intervention, estimating Equation 3 and clustering standard errors at the sublo-

[15] I.e. assuming 30% OAF density, 80pct of whom are enrolled in study in high density areas (versus 40% in low density areas), 63% of groups in a given area are in T1 + T2, and 65% who are offered the loan sign up, then differential market-level saturation = 0.30*(0.8-0.4)*0.63*0.65 = 4.9%. Because OAF client farmers are typically higher yielding than other smallholders in the area due to their higher average input use, the average supply effect might be higher – but we do not have the data to verify this.

cation level. Regression results are shown in Table 4 and plotted non-parametrically in Figure 7. Our monthly price data began in November, and we see that prices in high-intensity areas start out about 3% higher in the immediate post-harvest months. As can be seen in Figure 7, prices then converged in the high and low density areas, although the interaction between the monthly time trend and the high intensity dummy is not quite significant at conventional levels. Nevertheless, the overall picture painted by the market price data is remarkably consistent with the individual-level results presented above. Larger inward shifts in supply early on caused prices to start higher in high-intensity areas, and prices equalize at about the time the treated individuals switch from being net buyers to net sellers. Results are similar whether we match markets to sublocations using our own location data, or using OAF estimates of the sublocation into which each market falls (Table 4).

To further check robustness of the price results, we start by dropping sublocations one-by-one and re-estimating prices differences. As shown in the left panel of Figure A.3, differential trends over time in the two areas do not appear to be driven by particular sublocations. Second, building on other experimental work with small numbers of randomization units (Bloom et al., 2013; Cohen and Dupas, 2010), we generate 1000 placebo treatment assignments and compare the estimated price effects under the "true" (original) treatment assignment to estimated effects under each of the placebo assignments.[16] Results are shown in the two right hand panels of Figure A.3. The center panel shows price differences under the actual treatment assignment in black, and the placebo treatment assignments in grey. "Exact" p-values on the test that the price difference is zero are then calculated by summing up, at each point in the support, the number of placebo treatment estimates that exceed the actual treatment estimate and dividing by the total number of placebo treatments (1000 in this case); these are shown in the right-hand panel of the figure. Calculated this way, prices differences are significant at conventional levels for the first 3-4 months post harvest, roughly consistent with the results shown in Figure 7.

[16]With 17 sublocations, 9 of which are "treated" with a high number of treatment farmers, we have 17 choose 9 possible treatment assignments (24,310). We compute treatment effects for a subset of these possible placebo assignments.

5.2 Individual results with spillovers

We now revisit the individual results, re-estimating them to account for the variation in treatment density across sublocations. We note at the outset that while our experiment affected local market prices differentially in high- and low-treatment density areas, changes in treatment density could precipitate other spillovers beyond output price effects. For instance, sharing of maize or informal lending between households could also be affected by having a locally higher density of loan recipients; as an untreated household, your chance of knowing someone who got the loan is higher if you live in a high-treatment-density areas. Nevertheless, these spillovers could be positive or negative – e.g. we don't know *ex ante* whether our treatment would cause individuals to exit informal lending relationships or to expand them, or whether it would allow them to reduce their maize transfers or allow them to give out more maize to untreated households. We attempt to clarify the sign and magnitude of these potential spillovers in what follows.

Table 5 and Figure 8 show how our three main outcomes respond in high versus low density areas for treated and control individuals. Inventory treatment effects do not significantly differ as a function of treatment intensity for the pooled treatment, but differ for T1 (Columns 1 and 2 in Table 5). Nevertheless, in both the high and low intensity areas, inventories are significantly higher for both T1 and the pooled treatment (point estimates are positive for T2 but not significant).

Effects on net revenues paint a different picture. Treatment effects in low intensity areas are now significant for the pooled, T1, and T2 estimates and are much larger than what was estimated earlier. However, point estimates on treatment effects in high-intensity areas are now close to zero and we can never reject that they are different from zero. This suggests that there is something about higher treatment density that erodes the effect of the loan on profitability. There is also some evidence that net revenues were higher in high-intensity control group relative to the low intensity control group (see middle panel of Figure 8 and the estimate on the Hi dummy in Columns 3 and 4 of Table 5), but these effects are not significant. Effects on consumption, as with earlier estimates, remain quite noisy, and we can't rule out reasonably large positive or negative effects for any treatment group.

Could these differential net revenue effects have come through price spillovers alone? Note that we can immediately rule out a few prosaic explanations. First, covariates were balanced at baseline between high- and low-intensity areas (Table A.1), and loan size does not differ systematically across high and low intensity areas. However, we do find that loan take-up was significantly lower in high intensity areas - 13ppt lower on a base of 65% (significant at 1%). We believe that this is likely the result of repayment incentives faced by OAF field staff: our loan intervention represented a substantial increase in the total OAF credit outlay in high-intensity areas, and given contract incentives for OAF field staff that reward a high repayment rate for clients in their purview, these field officers might have more carefully screened potential adopters.[17] This differential take-up could matter for our treatment effects because we estimate the Intent-to-treat, and given a constant treatment-effect-on-the-treated, ITT estimates should be mechanically closer to zero in cases where take-up is lower. Nevertheless, it appears that this differential take-up is unlikely to explain the entire difference in treatment effects between high and low intensity areas: if there are no other spillovers, and treatment-on-treated effects are the same in high and low intensity areas, then ITT estimates in the high intensity ares should be 80% as large (0.52/0.65). However, point estimates on revenue treatment effects are *zero* in the high-intensity areas, which is unlikely explained by differential take-up.

Table A.3 explores other possibilities in more detail, looking at the differential effects over time. First, while differences in inventories do not vary significantly as a function of treatment density, point estimates suggest that inventories were slightly lower for treated individuals in high density areas relative to low density areas, particularly early on. This is consistent with increased transfers from treated to control households in high-intensity areas, but could also be consistent with an equilibrium response to higher prices: more people holding maize off the market post-harvest in these areas caused prices to increase, and in equilibrium this encouraged a little bit more initial selling. However, point estimates also suggest slightly higher inventories for untreated individuals in high relative to low intensity areas early in the period (although estimates are not near significant), which is the opposite of what would be expected if the only spillovers were due to price effects; higher post-harvest prices would presumably encourage more early sales. Given the relatively large

[17] We are exploring this in further discussions with OAF field staff and administration.

standard errors, though, this result is not definitive. The main difference in revenue appears to be because treated farmers in low intensity areas ended up with a little more to sell in the second and third periods, a result of having bought relatively more (at lower prices) in the first period and thus carried more inventory (although again, these estimates are not significant).

As further evidence on the nature of the spillover, we collected data on maize transfers and on household-to-household lending data during our follow-up survey rounds, and can use these data to directly assess whether differential treatment intensity affected these (self-reported) transfers. We find that the amount of cash lent to or borrowed from other households does not appear to respond to either treatment or to treatment intensity, and we similarly find no effect on the amount of transfers made in-kind (results not shown).

Overall, then, the individual-level spillover results are perhaps most consistent with spillovers through market prices. We find no direct evidence of higher transfers in high-intensity areas, and it appears that while treated farmers everywhere stored more, treated farmers in low-intensity areas purchased more maize at low prices early on and carried more inventories into the months of (slightly) higher prices.

We conclude this section by noting that, had we just run the experiment at our high treatment density, we would have found results very similar to what has been found in existing microcredit literature: a significant effect of improved credit access on inventories, but zero effect on revenues. While our rural setting is one in which certain types of spillovers (e.g. through prices) might be more important relative to the more urban settings that typify the existing microcredit experiments, our results do suggest that "headline" estimates of microcredit's impacts could be substantially shaped by the saturation at which the experiment is run.

6 Conclusion

We study the effect of offering Kenyan maize farmers a cash loan at harvest. The timing of this loan is motivated by two facts: the large observed average increase in maize prices between the post harvest season and the lean season six to nine months later, and the inability of most poor farmers appear to successfully arbitrage these prices due to a range of "non-discretionary" consumption expenditures they must make immediately after harvest. Instead of putting maize in storage and selling when the price is higher, farmers are observed to sell much of it immediately, sacrificing potential profits.

We show that access to credit at harvest "frees up" farmers to use storage to arbitrage these prices. 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 profits. Using experimentally-induced variation in the density of treatment farmers across locations, we document that this change in storage and marketing behavior aggregated across treatment farmers also affects local maize prices: post harvest prices are significantly higher in high-density areas, consistent with more supply having been taken off the market in that period, and are lower later in the season (but not significantly so). These general equilibrium effects feed back to our profitability estimates, with farmers in low-density areas – where price differentials were higher and thus arbitrage opportunities greater – differentially benefiting.

Our findings make a number of contributions. First, our results are some of the first experimental results to find a positive and significant effect of microcredit on the profits of microenterprises (farms in our case), and the first experimental study to directly account for general equilibrium effects in this literature. While we cannot claim that these two facts are more generally related, it is the case in our particular setting that failing to account for these GE effects substantially alters the conclusions drawn about the average benefits of improved credit access. This suggests that explicit attention to GE effects in future evaluations of credit market interventions could be warranted.

Second, 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 broad scale price phenomenon that further lowers farm income and increases what these households must pay for food. Our results suggest that in this setting, expanding access to affordable credit could reduce this price variability and thus have benefits for recipient and non-recipient households alike.

What our results do not address is why larger actors – e.g. large-scale private traders – have not stepped in to bid away these arbitrage opportunities. We are exploring this question in follow-up work in the region. Traders do exist in the area and can commonly be found in local markets, and we are repeatedly surveying a sample of these traders to better understand their cost structure and marketing activities. Preliminary findings suggest that, just as high transportation costs appear to affect the temporal dispersion of prices in individual markets by limiting inter-market trade, they also affect the spatial dispersion of prices across markets, and traders report being able to make even higher total profits by engaging in spatial arbitrage (relative to temporal arbitrage). Nevertheless, this does not explain why the scale or number of traders engaging in spatial arbitrage have not expanded, and we hope to better understand this issue in this ongoing work.

References

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

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

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

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

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

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

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

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

Banerjee, A.V., E. Duflo, R. Glennerster, and C. Kinnan. 2013. "The Miracle of Microfinance?: Evidence from a Randomized Evaluation." *working paper, MIT* .

Barrett, C. 2008. "Displaced distortions: Financial market failures and seemingly inefficient resource allocation in low-income rural communities." *working paper, Cornell.* .

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

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

Blattman, Christopher, Nathan Fiala, and Sebastian Martinez. 2013. "Credit Constraints, Occupational Choice, and the Process of Development: Long Run Evidence from Cash Transfers in Uganda." *working paper* .

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

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

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

Brune, L., X. Giné, J. Goldberg, and D. Yang. 2011. "Commitments to save: A field experiment in rural malawi." *University of Michigan, May (mimeograph)* .

Buera, Francisco J, Joseph P Kaboski, and Yongseok Shin. 2012. "The macroeconomics of microfinance." Tech. rep., National Bureau of Economic Research.

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

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

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

Crepon, B., F. Devoto, E. Duflo, and W. Pariente. 2011. "Impact of microcredit in rural areas of Morocco: Evidence from a Randomized Evaluation." *working paper, MIT* .

Cunha, Jesse M, Giacomo De Giorgi, and Seema Jayachandran. 2011. "The price effects of cash versus in-kind transfers." Tech. rep., National Bureau of Economic Research.

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

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

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

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

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

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

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

Karlan, Dean, Ryan Knight, and Christopher Udry. 2012. "Hoping to win, expected to lose: Theory and lessons on micro enterprise development." Tech. rep., National Bureau of Economic Research.

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

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

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

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

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

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

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

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

Tables and Figures

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

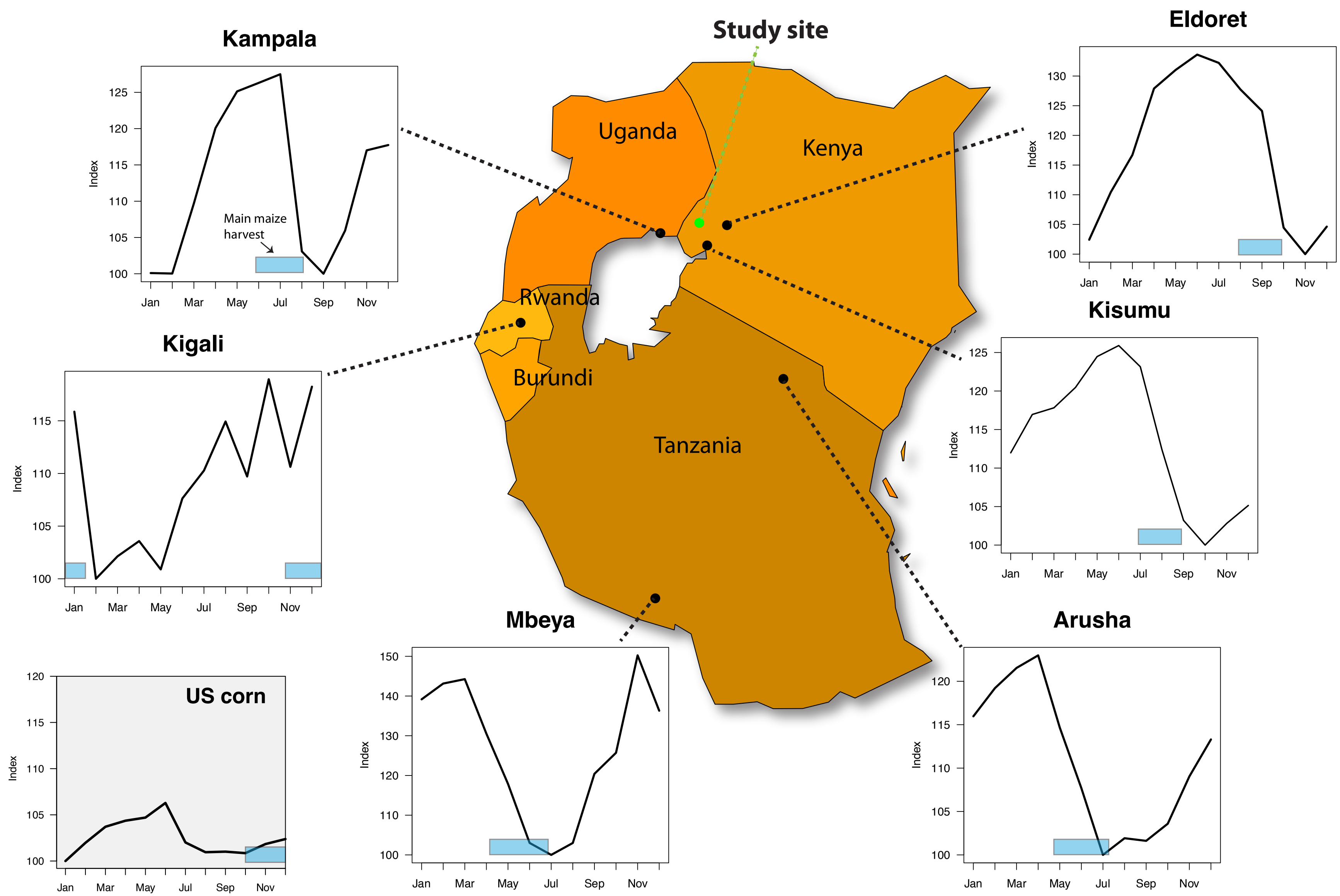


Figure 2: **Study design**. Randomization occurs at three levels. First, treatment intensity was randomized across 17 sublocations (top row, each box represents a sublocation). Second, treatment was randomized at the group level within sublocations (second row, each box representing a group in a given sublocation). Finally, tags and lockbox treatments were cross-randomized at the individual level (bottom row). Total numbers of randomized units in each bin are given on the left.

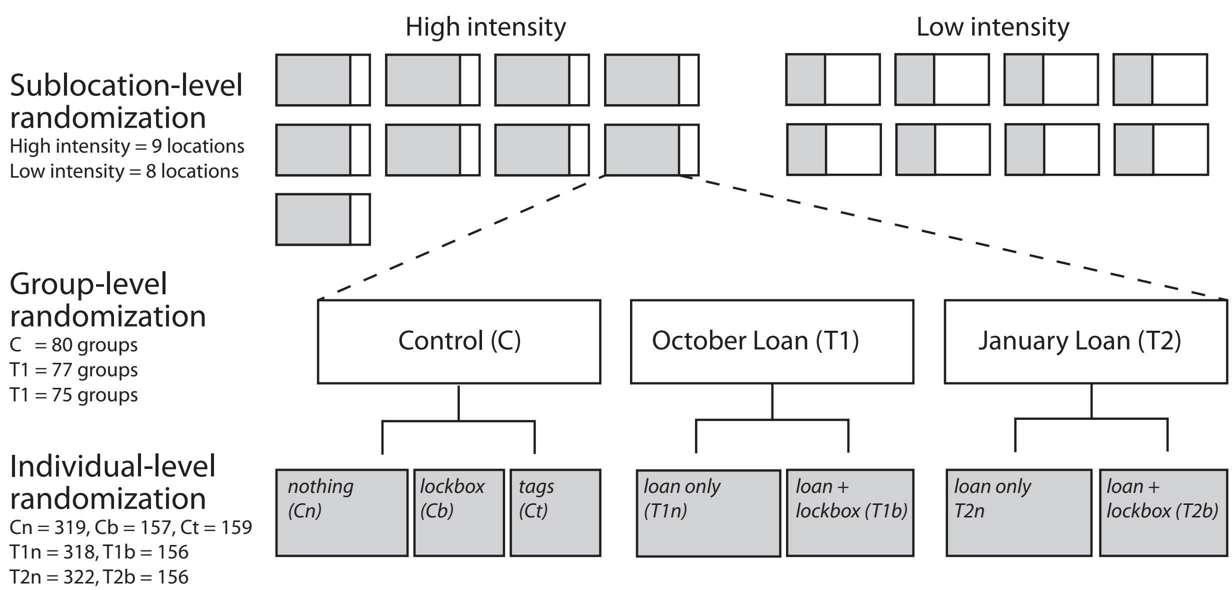
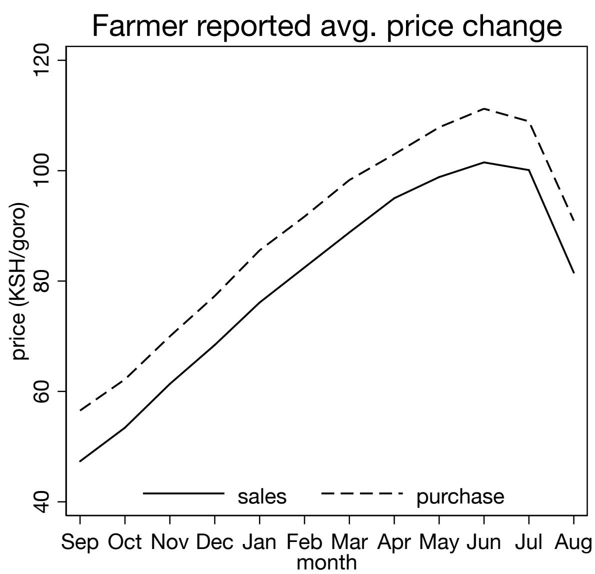


Figure 3: **Maize prices in local markets**. **Left panel**: farmer-reported average monthly maize prices for purchase and sales over 2007-2012, averaged over all farmers in our sample. Prices are in Kenyan shillings per goro goro (2.2kg). **Right panel**: farmers expectations of sales prices over the Sept2012-Aug2013 period, as reported in August2012 (solid red line), and actual observed sales prices in local markets over the same period (dotted line).



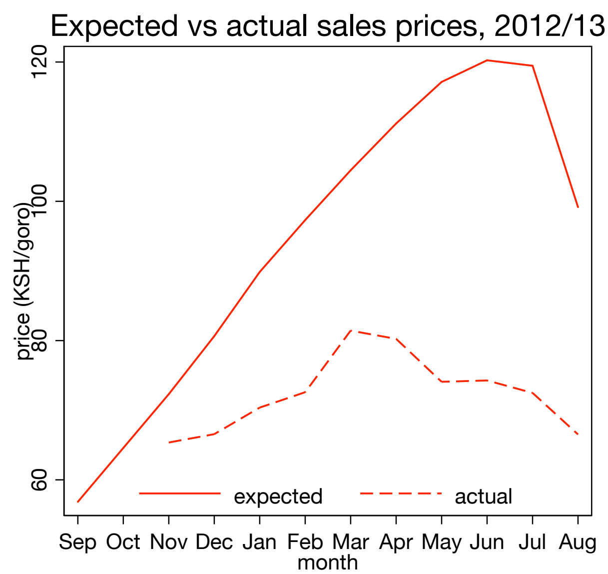


Figure 4: **Pooled treatment effects, assuming no spillovers**. The top row of plots shows how average inventories, net revenues, and log per capita consumption evolve over the study period in the treatment groups (T1 + T2) versus the control group, as estimated with fan regressions. The bottom row shows the difference between the treatment and control, with the bootstrapped 95% confidence interval shown in grey (1000 replications drawing groups with replacement).

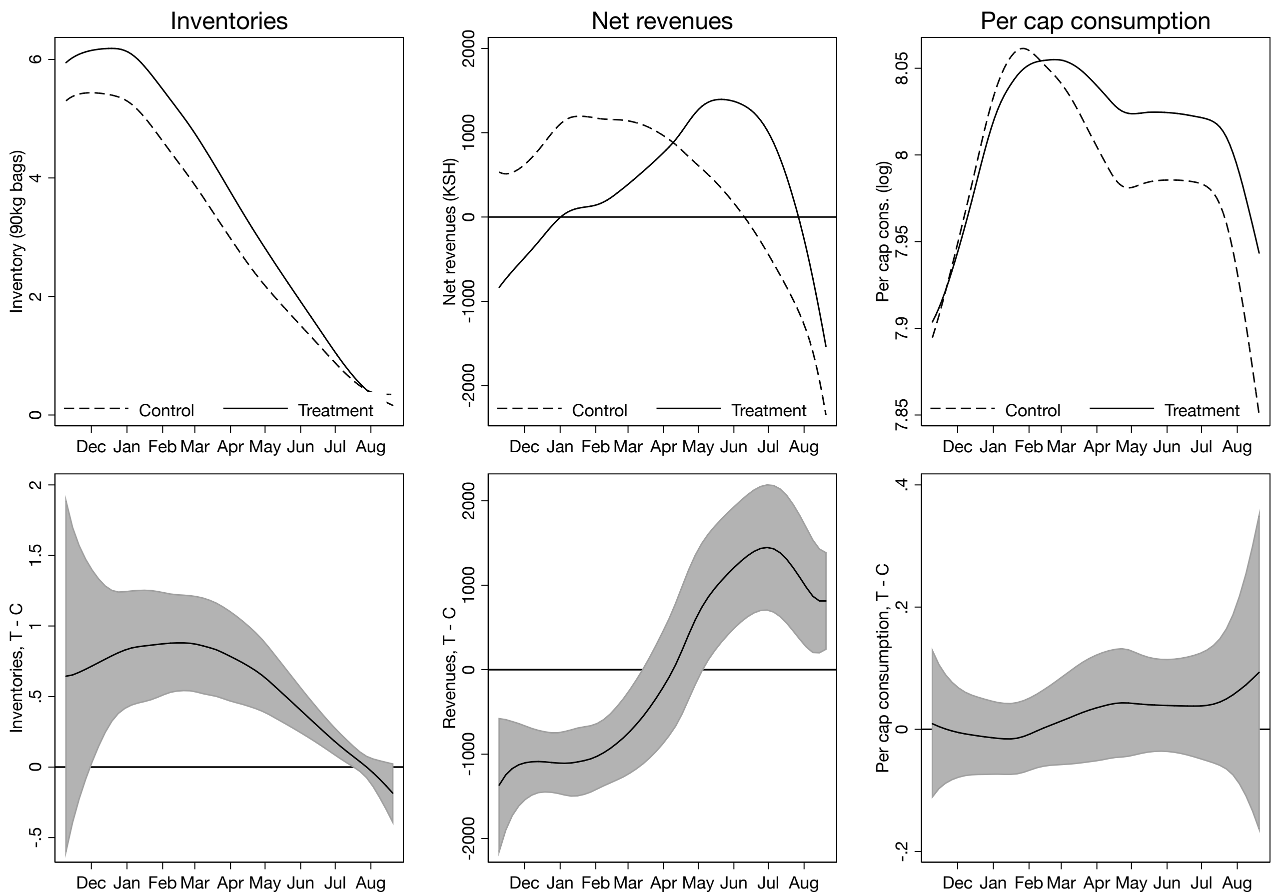


Figure 5: **Treatment effects by loan timing, assuming no spillovers**. Plots shows how average inventories, net revenues, and log per capita consumption evolve over the study period for farmers assigned to T1 (blue line), T2 (red line), and C (black dashed line), as estimated with fan regressions.

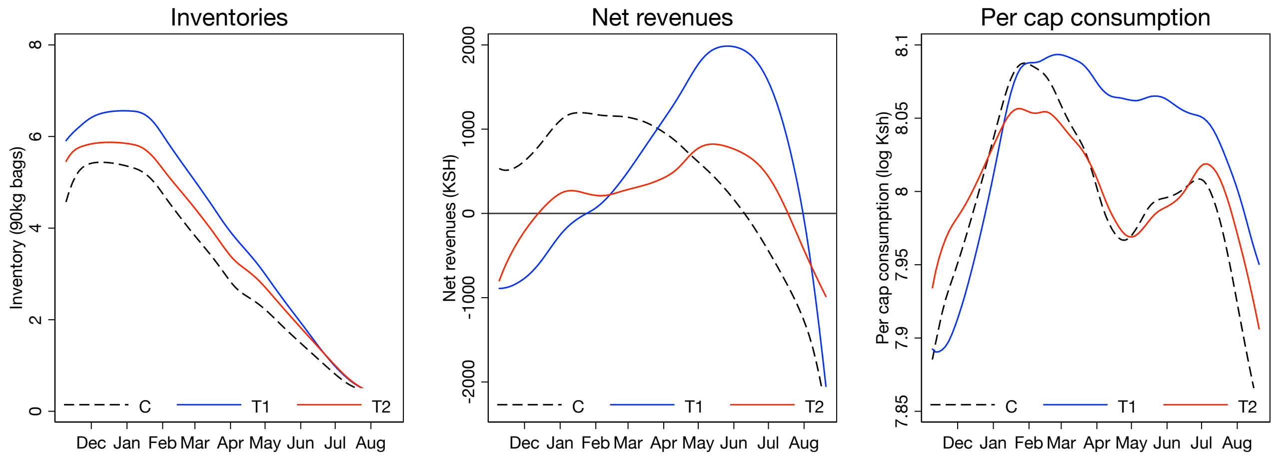


Figure 6: **Revenue treatment effects by loan timing, assuming no spillovers**. Plots show the difference in net revenues over time for T1 versus C (left), T2 versus C (center), and T1 versus T2 (right), with bootstrapped 95% confidence intervals shown in grey.

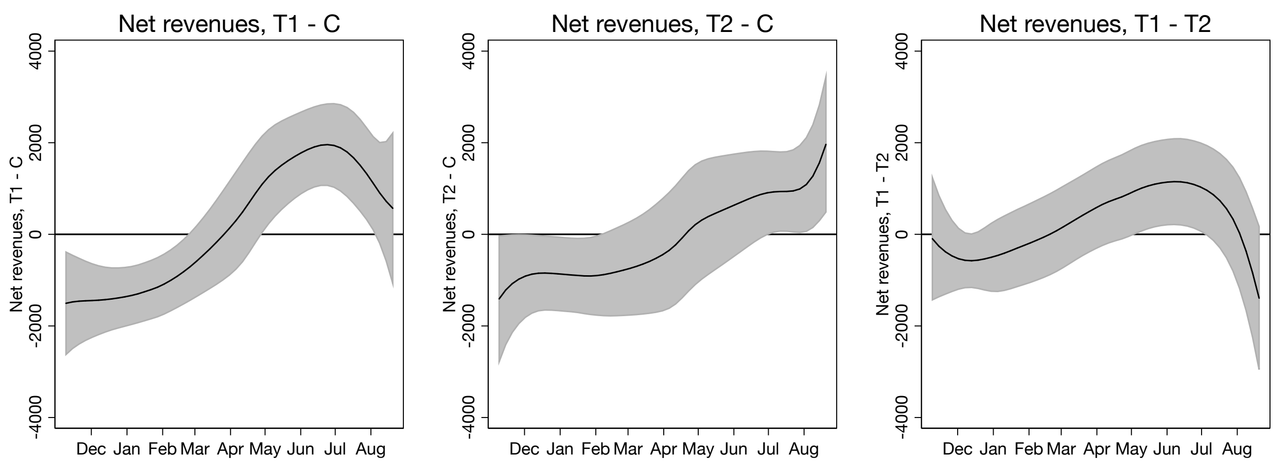


Figure 7: **Market prices for maize as a function of local treatment intensity**. The left panel shows the average sales price in markets in high-intensity areas (solid line) versus in low-intensity areas (dashed line) over the study period. The right panel shows the average difference in log price between high- and low-intensity areas over time, with the bootstrapped 95% confidence interval shown in grey.

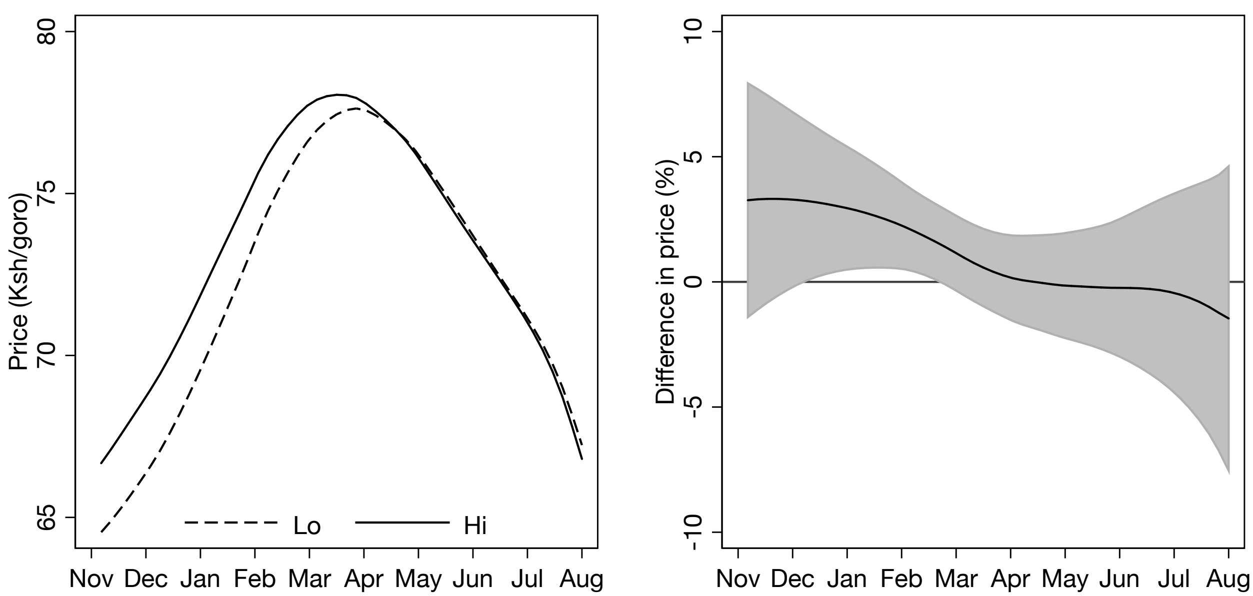


Figure 8: **Treatment effects by treatment intensity**. Average inventories, net revenues, and log per capita consumption over the study period in the pooled treatment groups (T1 + T2) versus the control group, split apart by high intensity areas (orange lines) and low-intensity areas (black lines).

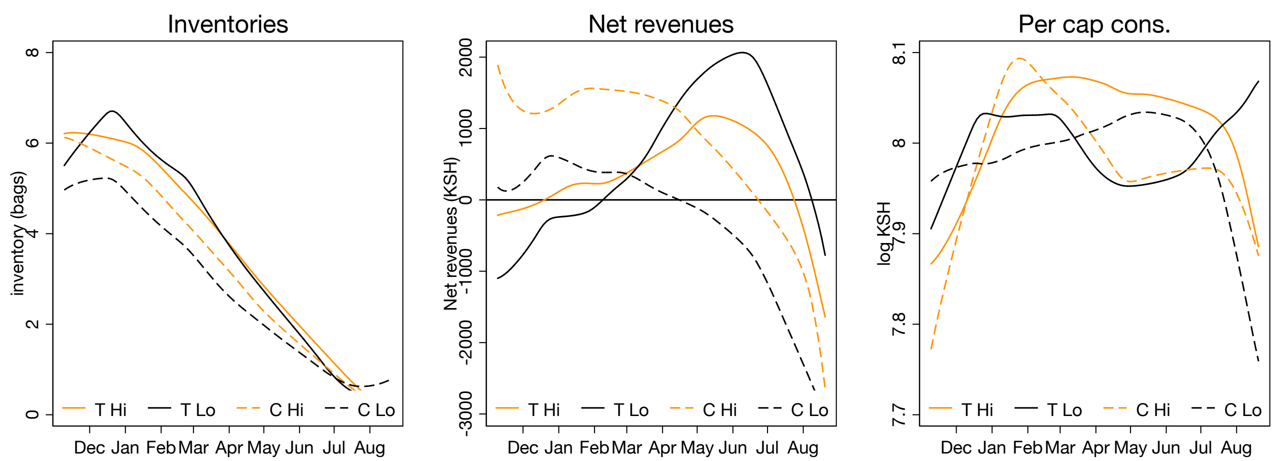


Table 1: **Summary statistics and balance among baseline covariates.** The first three columns give the means in each treatment arm. The 4th column gives the total number of observations across the three groups. The last four columns give differences in means normalized by the Control sd, with the corresponding p-value on the test of equality.

	C	T1	T2	Obs	C - T1		C - T2	
					sd	*p-val*	*sd*	*p-val*
Male	0.33	0.30	0.29	1,589	0.08	0.20	0.08	0.15
Number of adults	3.20	2.98	3.03	1,510	0.10	0.09	0.08	0.16
Kids in school	3.07	2.91	3.08	1,589	0.08	0.17	-0.01	0.93
Finished primary	0.77	0.73	0.70	1,490	0.09	0.15	0.16	0.01
Finished secondary	0.27	0.25	0.26	1,490	0.05	0.44	0.03	0.63
Total cropland (acres)	2.40	2.57	2.31	1,512	-0.05	0.39	0.03	0.65
Total school fees (1000 Ksh)	29.81	27.47	27.01	1,589	0.06	0.33	0.07	0.22
Total cash savings (trim)	5,389.84	5,019.01	4,447.26	1,572	0.03	0.66	0.07	0.24
Has bank savings acct	0.43	0.41	0.43	1,589	0.03	0.65	-0.00	0.95
Taken bank loan	0.08	0.08	0.08	1,589	0.01	0.84	0.02	0.70
Taken informal loan	0.25	0.25	0.24	1,589	-0.00	1.00	0.02	0.72
Off-farm wages (Ksh)	3,797.48	3,678.41	4,152.24	1,589	0.01	0.88	-0.03	0.64
Business profit (Ksh)	1,801.69	2,433.02	2,173.79	1,589	-0.10	0.31	-0.06	0.41
Net seller 2011	0.30	0.31	0.34	1,428	-0.01	0.92	-0.09	0.18
Autarkic 2011	0.06	0.06	0.08	1,589	0.00	0.96	-0.07	0.26
% maize lost 2011	0.01	0.01	0.02	1,428	-0.02	0.73	-0.04	0.53
2012 LR harvest (bags)	11.03	11.27	11.09	1,484	-0.03	0.64	-0.01	0.91
Calculated interest correctly	0.73	0.73	0.70	1,580	0.01	0.90	0.06	0.32
Digit span recall	4.58	4.58	4.56	1,504	-0.00	0.97	0.02	0.78

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

Table 2: **Treatment effects at the individual level.** Regressions include round fixed effects and strata fixed effects, with errors clustered at the group level.

	Inventories		Prices		Revenues		Consumption	
	(1) Pooled	(2) By round	(3) Purchase price	(4) Sales prices	(5) Pooled	(6) By round	(7) Pooled	(8) By round
Treatment	0.61***		-22.28	-3.62	282.99		0.02	
	(0.11)		(19.55)	(19.39)	(218.32)		(0.03)	
Treatment - Round 1		0.89***				-1091.34***		-0.01
		(0.24)				(295.25)		(0.04)
Treatment - Round 2		0.77***				534.48		0.05
		(0.15)				(429.56)		(0.04)
Treatment - Round 3		0.18*				1340.64***		0.03
		(0.11)				(388.18)		(0.04)
Constant	209.99**	205.62**	-11568.08	-47616.89***	-630244.67***	-601691.91***	-8.97	-7.95
	(87.35)	(87.49)	(15573.70)	(17699.52)	(229995.15)	(225529.44)	(21.83)	(21.85)
Observations	3816	3816	1914	1425	3776	3776	3596	3596
Mean of Dep Variable	2.67	2.67	2982.02	2827.58	334.41	334.41	8.00	8.00
R squared	0.49	0.49	0.30	0.47	0.13	0.13	0.21	0.21
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 3: **Effects of sub-treatments.** Regressions include round fixed effects and strata fixed effects, with errors clustered at the group level.

	Inventories		Prices		Revenues		Consumption	
	(1) Pooled	(2) By round	(3) Purchase price	(4) Sales prices	(5) Pooled	(6) By round	(7) Pooled	(8) By round
T1	0.77***		-47.81**	-4.84	541.95**		0.04	
	(0.13)		(23.20)	(21.43)	(248.78)		(0.03)	
T2	0.46***		2.47	-2.32	36.03		0.01	
	(0.13)		(22.47)	(23.05)	(248.15)		(0.03)	
T1 - Round 1		1.25***				-1218.96***		-0.00
		(0.27)				(353.43)		(0.05)
T1 - Round 2		0.91***				924.50*		0.08*
		(0.19)				(512.50)		(0.05)
T1 - Round 3		0.18				1840.70***		0.04
		(0.13)				(483.92)		(0.04)
T2 - Round 1		0.54**				-951.27***		-0.01
		(0.27)				(347.35)		(0.05)
T2 - Round 2		0.65***				156.58		0.01
		(0.16)				(503.66)		(0.05)
T2 - Round 3		0.18				851.70**		0.02
		(0.12)				(410.53)		(0.04)
Observations	3816	3816	1914	1425	3776	3776	3596	3596
Mean of Dep Variable	3.03	3.03	2936.14	2887.46	501.64	501.64	8.02	8.02
SD of Dep Variable	3.73	3.73	425.20	437.86	6217.09	6217.09	0.66	0.66
R squared	0.49	0.50	0.30	0.47	0.13	0.14	0.21	0.21
T1 = T2 (pval)	0.02		0.04		0.04		0.19	

Table 4: **Market prices for maize as a function of local treatment intensity.**

	(1) Admin	(2) Admin	(3) Nearest	(4) Nearest
Hi Intensity	2.64*	2.51*	2.81*	2.70*
	(1.25)	(1.32)	(1.41)	(1.46)
Time	0.73***	0.75***	0.78***	0.81***
	(0.22)	(0.21)	(0.24)	(0.23)
Hi Intensity * Time	-0.37	-0.37	-0.39	-0.42
	(0.27)	(0.27)	(0.27)	(0.27)
Constant	68.93***	69.62***	68.54***	69.25***
	(1.10)	(1.12)	(1.34)	(1.33)
Observations	491	491	491	491
R squared	0.07	0.09	0.08	0.10
Controls	No	Yes	No	Yes

Data are for 52 market points across 17 sublocations, and are for November 2012 through August 2013. "Hi intensity" is a dummy for a sublocation randomly assigned a high number of treatment groups and "Time" is a time trend (month number). Standard errors are clustered at the sublocation level. Columns 1 and 2 match markets to sublocations using administrative data, columns 3 and 4 using location data on farmers and markets.

Table 5: **Individual level effects, accounting for treatment intensity.** Regressions include round fixed effects with errors clustered at the sublocation level. P-values on the test that the sum of the treated and treated*hi equal zero are provided in the bottom rows of the table.

	Inventories		Revenues		Consumption	
	(1) Pooled	(2) Split	(3) Pooled	(4) Split	(5) Pooled	(6) Split
Pooled	0.86***		1118.63**		-0.01	
	(0.26)		(418.41)		(0.05)	
Hi intensity	0.20	0.16	528.00	219.67	0.01	-0.01
	(0.37)	(0.31)	(573.50)	(521.01)	(0.04)	(0.05)
Pooled*Hi	-0.40		-1139.31**		0.04	
	(0.28)		(513.32)		(0.06)	
T1		1.17***		925.61***		-0.00
		(0.23)		(284.88)		(0.06)
T1*Hi		-0.68**		-589.44		0.06
		(0.24)		(461.87)		(0.07)
T2		0.47		768.93*		-0.03
		(0.27)		(426.16)		(0.06)
T2*Hi		-0.11		-1046.82*		0.05
		(0.31)		(515.01)		(0.06)
Observations	3816	4250	3776	4207	3596	3995
R squared	0.48	0.48	0.11	0.12	0.19	0.20
p-val P+PH=0	0.00		0.95		0.44	
p-val T1+T1H=0		0.00		0.38		0.16
p-val T2+T2H=0		0.02		0.36		0.52

A Appendix

Figure A.1: **Pilot data on maize inventories and marketing decisions over time**, using data from two earlier pilot studies conducted with One Acre Fund in 2010/11 with 225 farmers (top row) and 2011/12 with 700 different farmers (bottom row). *Left panels*: inventories (measured in 90kg bags) as a function of weeks past harvest. The dotted line is the sample median, the solid line the mean (with 95% CI in grey). *Right panels*: average net sales position across farmers over the same period, with quantities shown for 2010/11 (quantity sold minus purchased) and values shown for 2011/12 (value of all sales minus purchases).

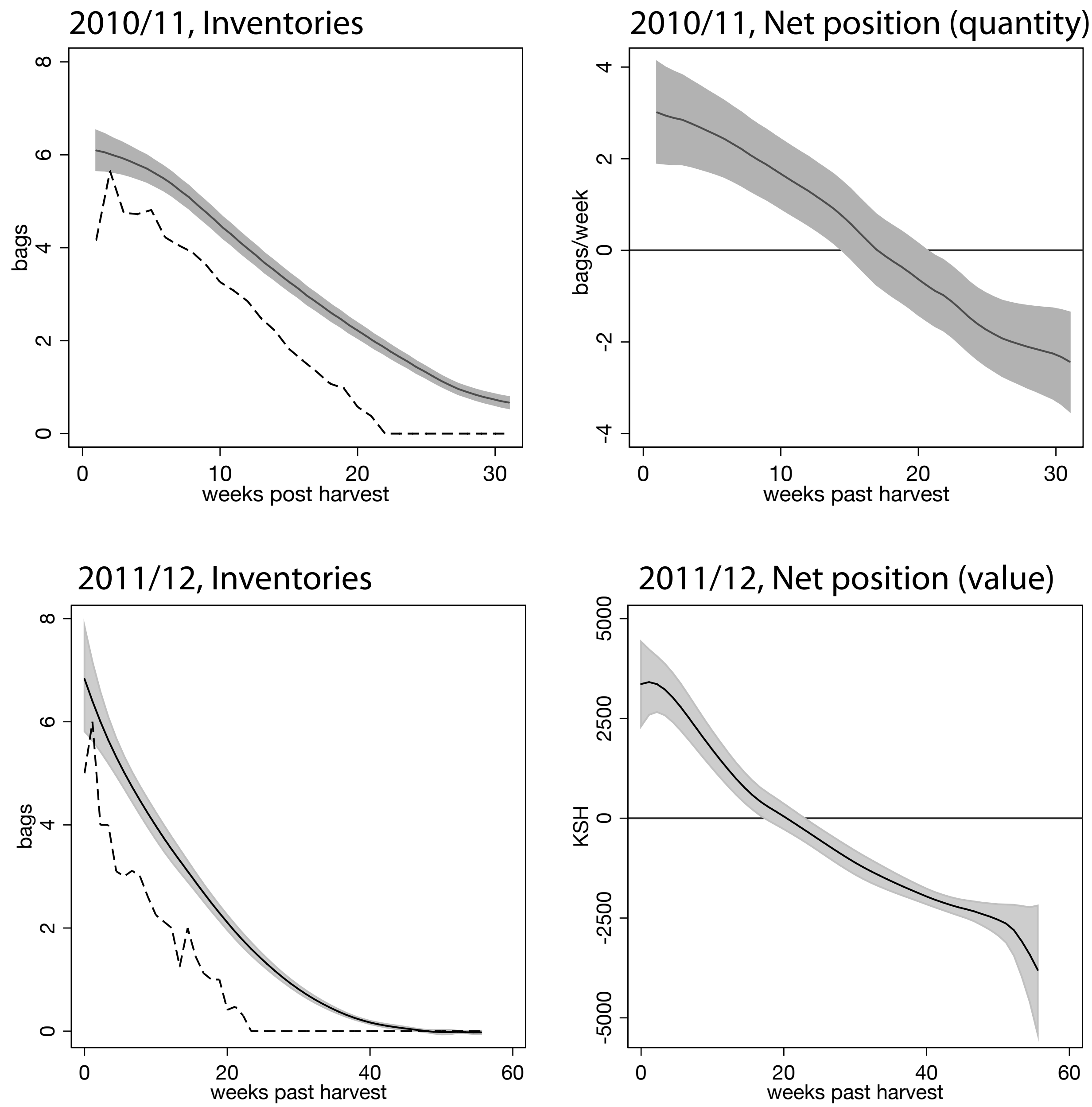


Figure A.2: **Study timeline**. The timing of the interventions and data collection are show at top, and the timing of the main agricultural season is shown at the bottom.

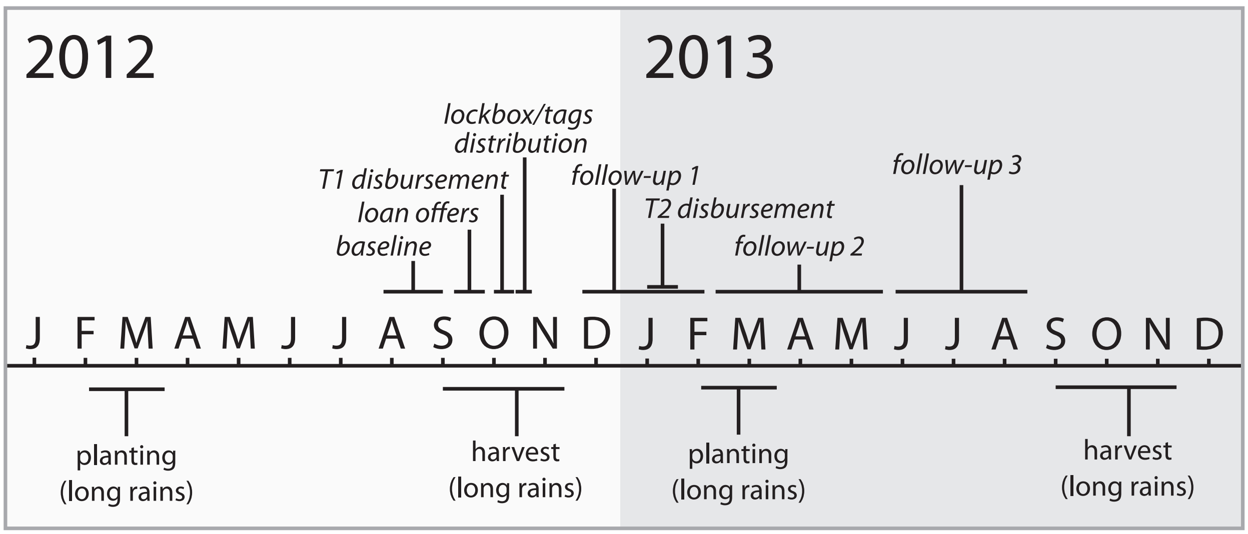


Figure A.3: **Robustness of price effect estimates**. *Left panel*: difference in prices between high and low-density markets over time for the full sample (black line) and for the sample with each sublocation dropped in turn (grey lines). *Center panel*: price effects under the “true” treatment assignment (black line) and 1000 placebo treatment assignments (grey lines). *Right panel*: randomization-inference based p-values on the test that the price difference is zero, as derived from the center panel.

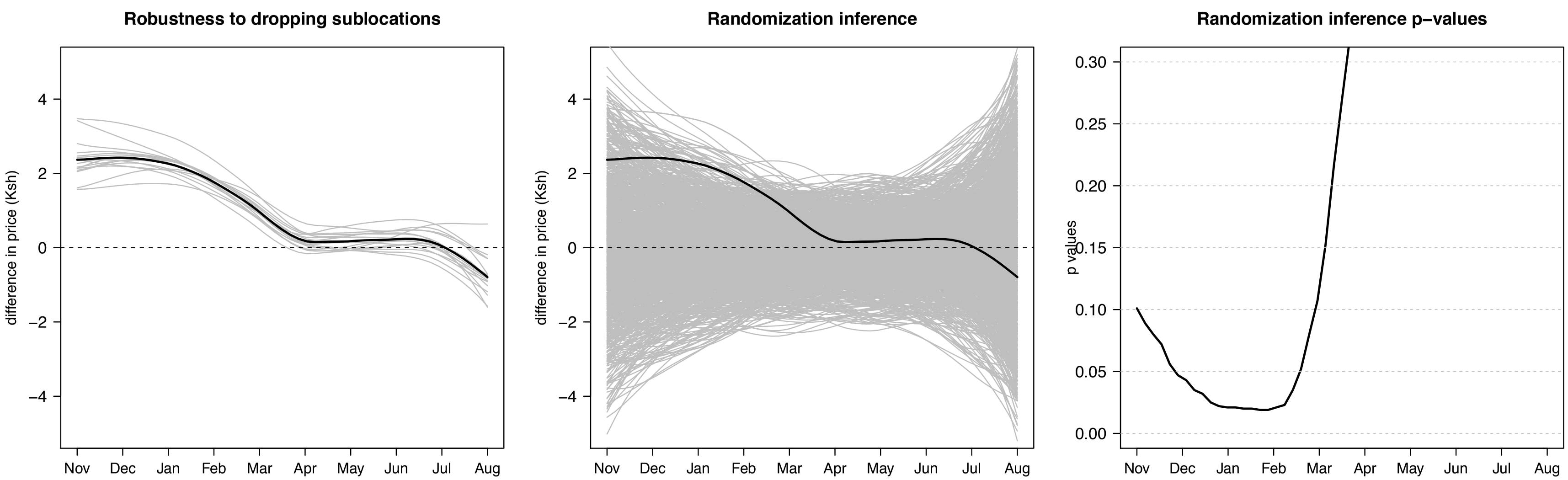


Table A.1: **Balance among baseline covariates, high versus low treatment intensity areas.** The first two columns give the means in the low or high treatment intensity areas, the 3rd column the total number of observations across the two groups, and the last two columns the differences in means normalized by the standard deviation in the low intensity areas, with the corresponding p-value on the test of equality.

	Lo	Hi	Obs	Lo - Hi	
				sd	*p-val*
Male	0.32	0.31	1,589	0.02	0.72
Number of adults	3.11	3.07	1,510	0.02	0.74
Kids in school	3.15	2.98	1,589	0.09	0.11
Finished primary	0.71	0.75	1,490	-0.08	0.13
Finished secondary	0.27	0.25	1,490	0.04	0.51
Total cropland (acres)	2.60	2.35	1,512	0.08	0.15
Number of rooms in hhold	3.31	3.08	1,511	0.08	0.10
Total school fees (1000 Ksh)	29.23	27.88	1,589	0.04	0.51
Average monthly cons (Ksh)	15,586.03	14,943.57	1,437	0.05	0.38
Avg monthly cons./cap (log Ksh)	7.98	7.97	1,434	0.02	0.77
Total cash savings (KSH)	5,776.38	6,516.09	1,572	-0.04	0.56
Total cash savings (trim)	5,112.65	4,947.51	1,572	0.01	0.82
Has bank savings acct	0.42	0.42	1,589	-0.01	0.91
Taken bank loan	0.07	0.09	1,589	-0.06	0.30
Taken informal loan	0.25	0.24	1,589	0.02	0.72
Liquid wealth	87,076.12	98,542.58	1,491	-0.12	0.06
Off-farm wages (Ksh)	3,965.65	3,829.80	1,589	0.01	0.84
Business profit (Ksh)	1,859.63	2,201.34	1,589	-0.04	0.53
Avg %Δ price Sep-Jun	121.58	138.18	1,504	-0.21	0.00
Expect %Δ price Sep12-Jun13	105.89	128.19	1,510	-0.37	0.00
2011 LR harvest (bags)	10.52	8.70	1,511	0.08	0.03
Net revenue 2011	-2,175.44	-4,200.36	1,428	0.03	0.45
Net seller 2011	0.34	0.30	1,428	0.08	0.16
Autarkic 2011	0.06	0.07	1,589	-0.04	0.53
% maize lost 2011	0.01	0.01	1,428	0.00	0.95
2012 LR harvest (bags)	11.57	10.94	1,484	0.07	0.19
Calculated interest correctly	0.68	0.74	1,580	-0.12	0.03
Digit span recall	4.49	4.60	1,504	-0.10	0.08
Maize giver	0.25	0.27	1,589	-0.05	0.37
Delta	0.14	0.13	1,512	0.07	0.28

See Table 1 and the text for additional details on the variables.

Table A.2: **Effects of tags.** Regressions include round fixed effects and strata fixed effects, with errors clustered at the group level.

	(1) Inventories	(2) Inventories	(3) Revenues	(4) Revenues	(5) Consumption	(6) Consumption
Pooled	0.62***		277.75		0.02	
	(0.11)		(216.70)		(0.03)	
T1		0.78***		533.40**		0.04
		(0.14)		(248.43)		(0.03)
T2		0.47***		33.05		0.00
		(0.12)		(245.65)		(0.03)
tags	0.23	0.23	217.85	218.45	-0.04	-0.04
	(0.16)	(0.16)	(305.87)	(305.95)	(0.05)	(0.05)
Observations	4250	4250	4207	4207	3995	3995
R squared	0.50	0.50	0.13	0.14	0.21	0.21
pooled-tags p-val	0.02		0.84		0.25	
T1-tags p-val		0.00		0.33		0.14
T2-tags p-val		0.17		0.56		0.45

Table A.3: **Effects on inventories, net quality sold, and net revenues, by treatment and treatment intensity**. Errors are clustered at the sublocation level. The omitted group is individuals in the control group in round 1.

	(1) Inventories	(2) Net quantities	(3) Net revenues
T - R1	1.39***	-0.21	-730.69**
	(0.37)	(0.13)	(314.22)
T - R1 * Hi	-0.78	-0.33	-502.00
	(0.48)	(0.21)	(484.28)
T - R2	1.09***	0.41**	1243.03**
	(0.34)	(0.17)	(575.99)
T - R2 * Hi	-0.49	-0.31	-929.25
	(0.36)	(0.25)	(823.98)
T - R3	0.12	1.00***	2809.83***
	(0.28)	(0.28)	(841.15)
T - R3 * Hi	0.05	-0.73**	-2045.81*
	(0.30)	(0.32)	(975.51)
R1 * Hi	0.43	0.35	657.77
	(0.62)	(0.21)	(483.43)
R2 * Hi	0.17	0.09	423.51
	(0.40)	(0.21)	(664.21)
R3 * Hi	-0.02	0.22	656.53
	(0.31)	(0.35)	(1106.79)
R2	-1.34**	-0.52	-1473.21
	(0.59)	(0.40)	(1096.07)
R3	-1.99*	-1.44*	-4079.27*
	(1.04)	(0.74)	(2027.71)
Observations	3816	3801	3776
R squared	0.48	0.12	0.12
p-val P1+P1H=0	0.07	0.01	0.00
p-val P2+P2H=0	0.00	0.58	0.60
p-val P3+P3H=0	0.09	0.11	0.15