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

Marshall Burke*

November 14, 2013

QUITE PRELIMINARY. PLEASE DO NOT CITE WITHOUT PERMISSION

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, and Jeremy Magruder for useful discussions, and thank seminar participants at Berkeley and Stanford 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, and I gratefully acknowledge funding from the Agricultural Technology Adoption Initiative. 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 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 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.[3]

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. As a likely result of these price effects, I find that treated farmers in high-density areas stored significantly more than their control counterparts, but their maize profits were indistinguishable from control. 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.

[3]The pre-analysis plan is registered here: https://www.socialscienceregistry.org/trials/67, and is available upon request.

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 appears does not appear beneficial on average. First, many small businesses or potential micro-entrepreneurs simply might not actually face profitable investment opportunities (Banerjee et al., 2013; Fafchamps et al., 2013; Karlan, Knight, and Udry, 2012; Banerjee, 2013).[4] 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 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.[5]

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 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

[4]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.

[5]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. . . . I believe 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).

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.[6] 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 (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

[6]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.

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.[7]

These increases also appear to be a lower bound on typical increase observed in the smaller markets in our study area, which (relative to these much larger markets) are characterized with much smaller "catchments" and less outside trade. We asked farmers at baseline to estimate

[7]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%.

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.[8] Second, pest-related losses appear surprisingly low in our

[8]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

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 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.

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.

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.[9]

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.[10] So a farmer who committed 5 bags when offered the October loan would receive $5*1500 = 7500$Ksh in cash in October ($\sim$\$90 at current exchange rates), and would be required to

[9]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).

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

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.[11] 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 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

[11] 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.

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.[12] 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 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

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

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 stupidity 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[13]:

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

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). We begin by clustering errors at the sublocation level when estimating (3). Future versions of the

[13]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.

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.[14]

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 for much of the year, on average about 20% more than the control group mean (Column 1 in Table

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

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 20-30% 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.[15] 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.[16]

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.

[15] Assuming 25% OAF density, population-level saturation = 0.25*0.8*0.63*0.65 = 8.2% (assuming 25pct of population is OAF, 80pct of these are enrolled in study, 63pct of them are in T1 + T2, and 65% who are offered the loan sign up). Because OAF client farmers are typically higher yielding than other smallholders in the area due to their higher average input use, they could represent more than 10% of the local supply – but we do not have the data to verify this.

[16] 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.

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 sublocation 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 5% higher in the immediate post-harvest months. This is consistent with the individual level results presented above: point estimates suggest that both T1 and T2 farmers became net purchasers rather than net sellers in this period (see middle panel, Figure 5), meaning supply would have shifted in (and demand out) in both areas, but more in high-intensity areas. 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 and outward shifts in demand 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.

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.[17] 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.

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

[17] 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.

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.[18] 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

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

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 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).

Finally, 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.

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

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

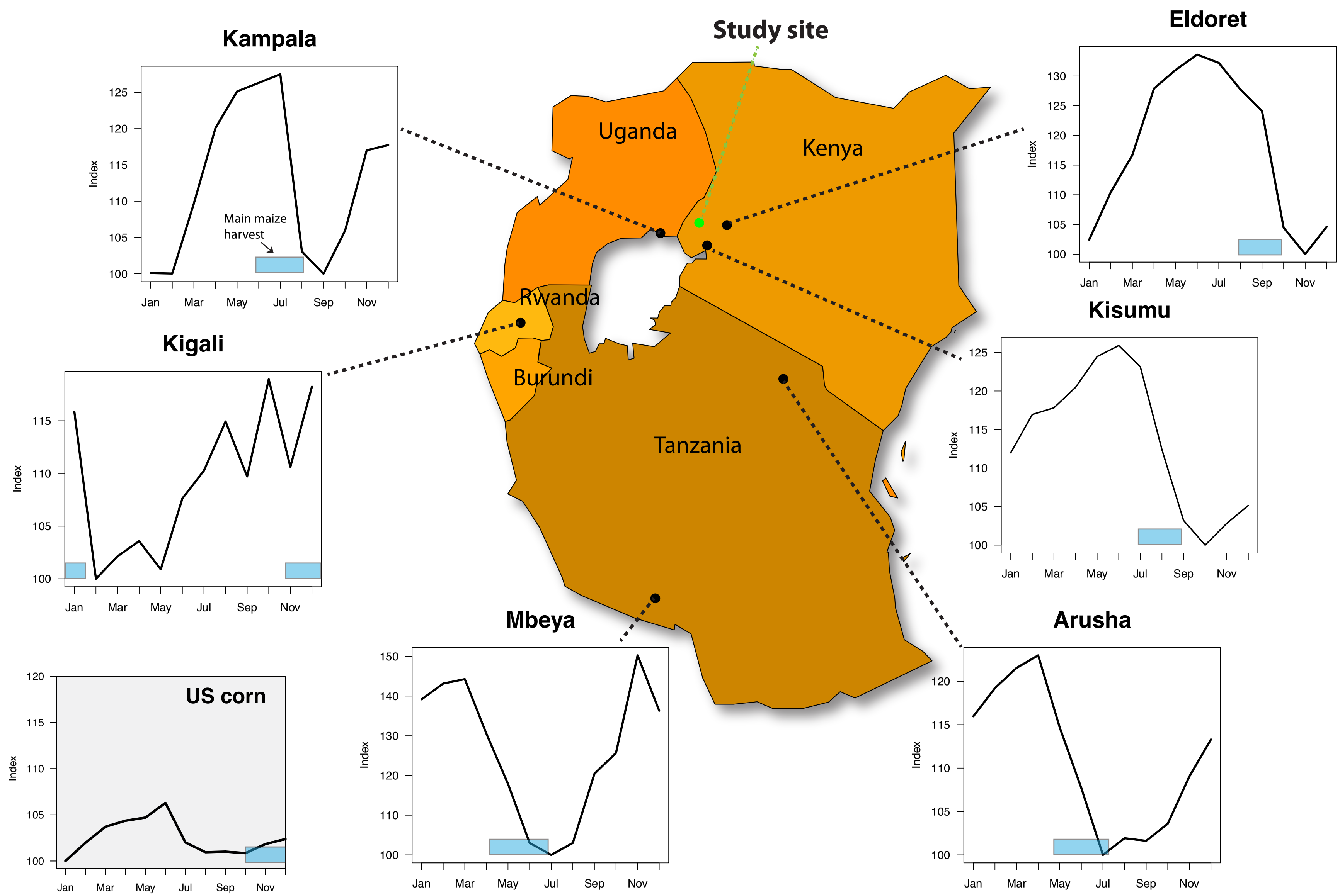


Figure 2: **Study design**. Randomization occurs at three levels. First, treatment intensity was randomized across 17 sublocations (top row, each box represents a sublocation). Second, treatment was randomized at the group level within sublocations (second row, each box representing a group in a given 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.

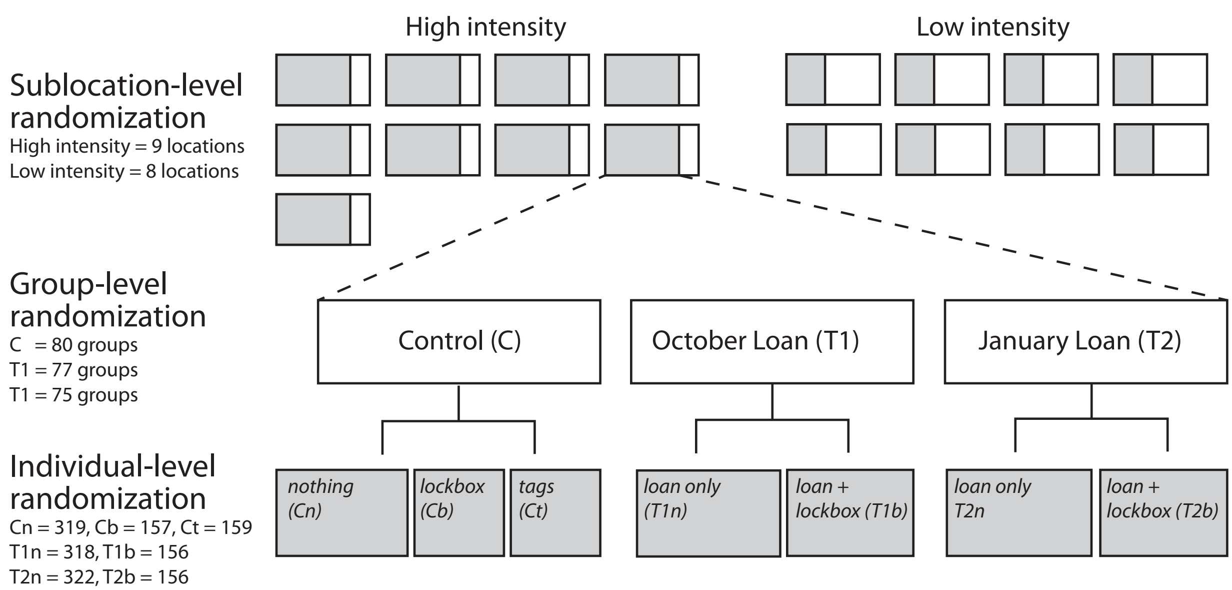


Figure 3: **Maize prices in local markets**. **Left panel**: farmer-reported average monthly maize prices for purchase and sales over 2007-2012, averaged over all farmers in our sample. Prices are in Kenyan shillings per goro goro (2.2kg). **Right panel**: farmers expectations of sales prices over the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

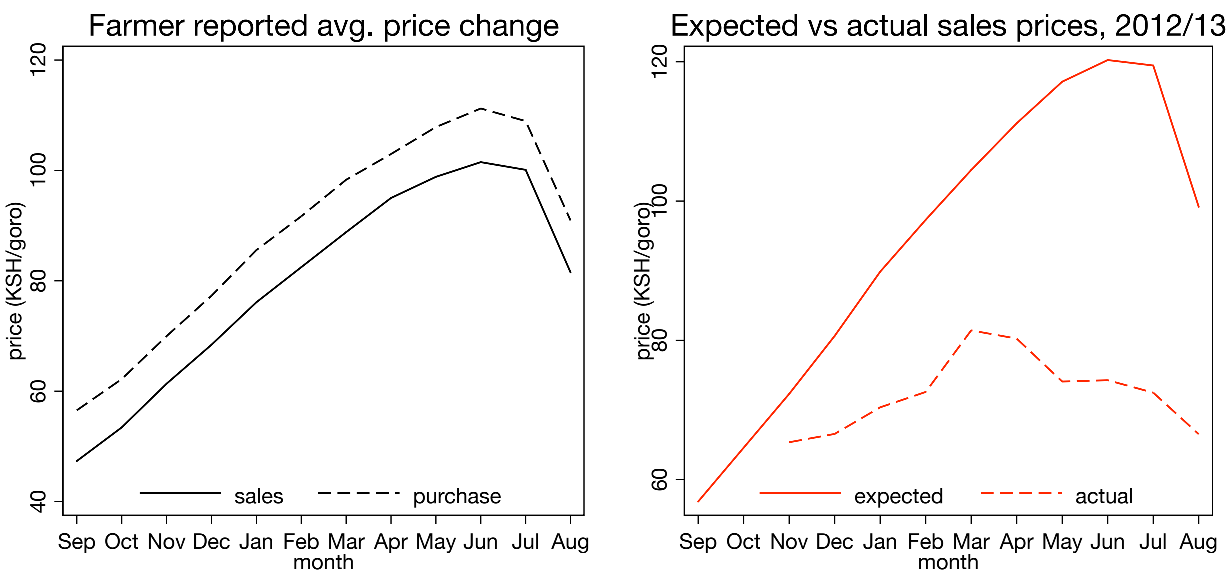


Figure 4: **Pooled treatment effects, assuming no spillovers**. The top row of plots shows how average inventories, net revenues, and log per capita consumption evolve over the study period in the treatment groups (T1 + T2) versus the control group, as estimated with fan regressions. The bottom row shows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

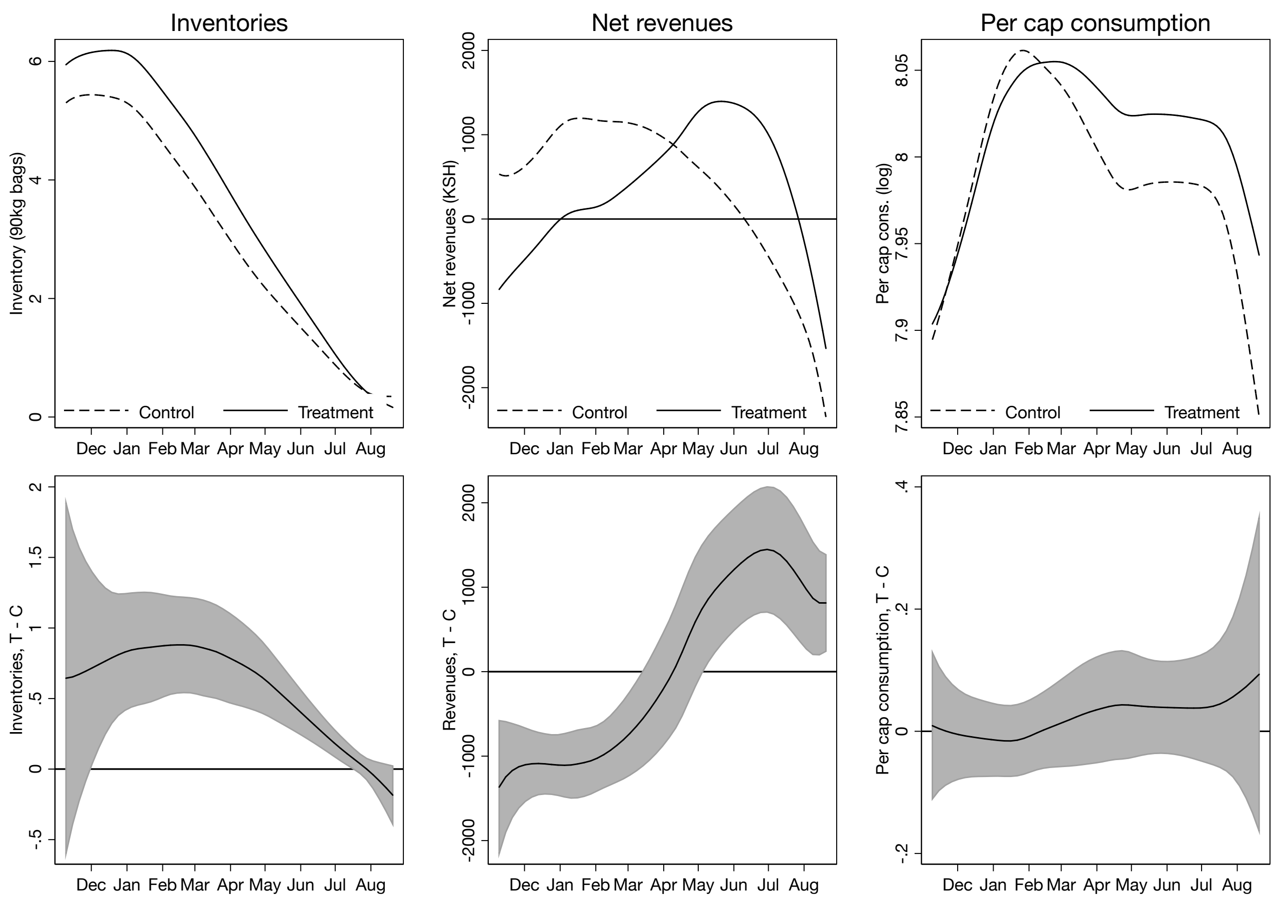


Figure 5: **Treatment effects by loan timing, assuming no spillovers**. Plots shows how average inventories, net revenues, and log per capita consumption evolve over the study period for farmers assigned to T1 (blue line), T2 (red line), and C (black dashed line), as estimated with fan regressions.

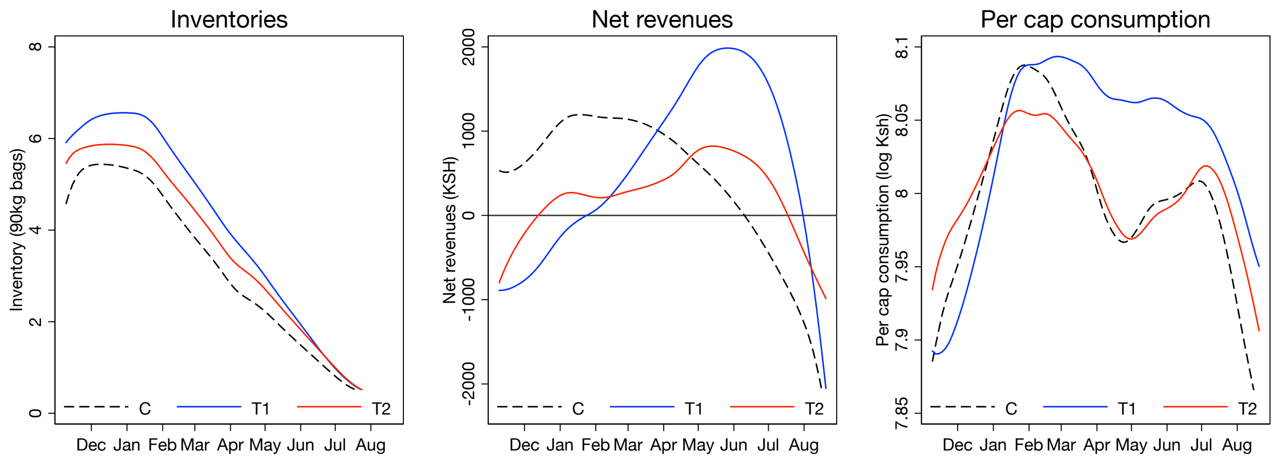


Figure 6: **Revenue treatment effects by loan timing, assuming no spillovers**. Plots show the difference in net revenues over time for T1 versus C (left), T2 versus C (center), and T1 versus T2 (right), with bootstrapped 95% confidence intervals shown in grey.

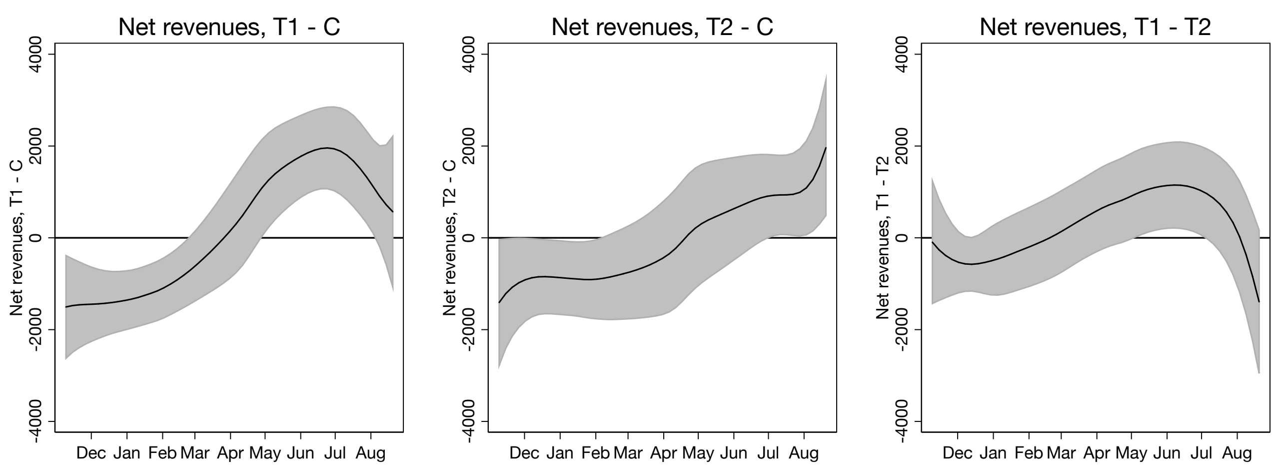


Figure 7: **Market prices for maize as a function of local treatment intensity**. The left panel shows the average sales price in markets in high-intensity areas (solid line) versus in low-intensity areas (dashed line) over the study period. The right panel shows the average difference in price between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

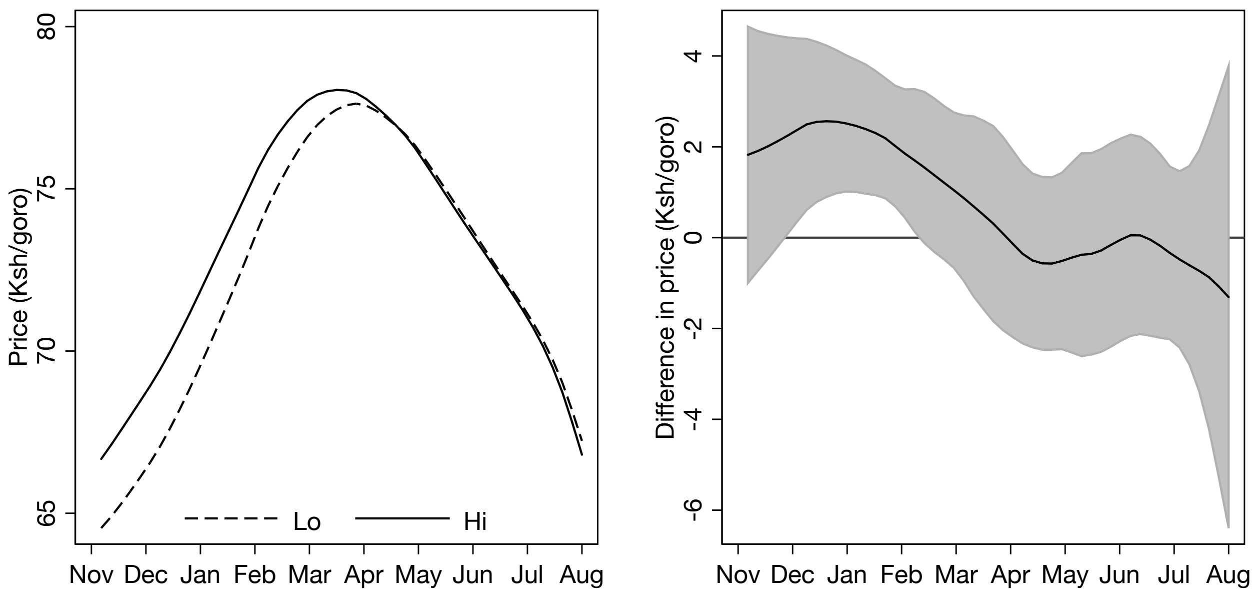


Figure 8: **Treatment effects by treatment intensity**. Average inventories, net revenues, and log per capita consumption over the study period in the pooled treatment groups (T1 + T2) versus the control group, split apart by high intensity areas (orange lines) and low-intensity areas (black lines).

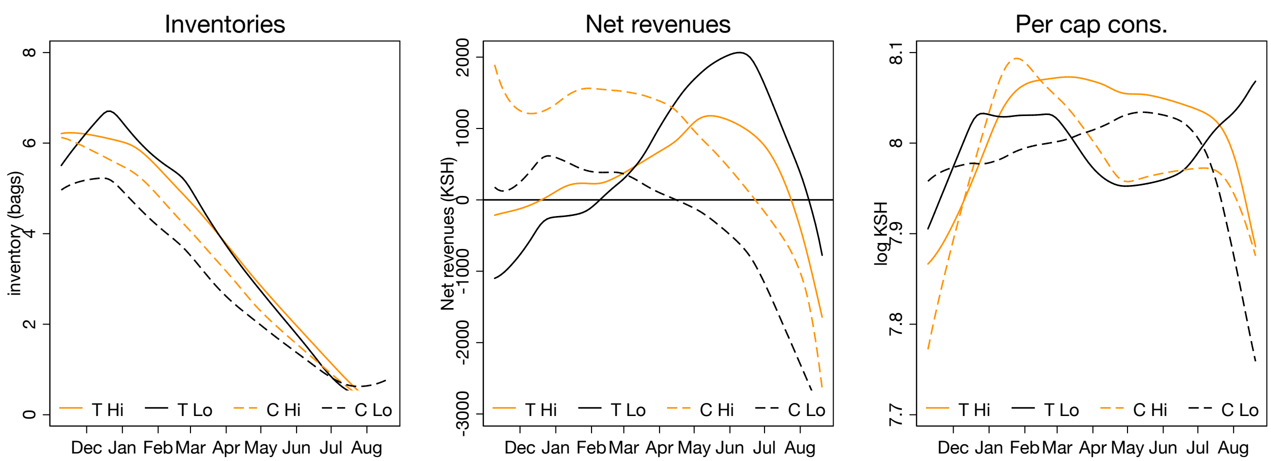


Table 1: **Summary statistics and balance among baseline covariates.** The first three columns give the means in each treatment arm. The 4th column gives the total number of observations across the three groups. The last four columns give differences in means normalized by the Control sd, with the corresponding 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
Number of rooms in hhold	3.25	3.10	3.04	1,511	0.04	0.40	0.06	0.20
Total school fees (1000 Ksh)	29.81	27.47	27.01	1,589	0.06	0.33	0.07	0.22
Average monthly cons (Ksh)	15,371.38	15,065.23	14,876.93	1,437	0.03	0.70	0.04	0.53
Avg monthly cons./cap (log Ksh)	7.96	8.00	7.95	1,434	-0.05	0.41	0.02	0.81
Total cash savings (KSH)	8,021.50	5,436.93	4,880.81	1,572	0.09	0.09	0.10	0.04
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
Liquid wealth	97,280.92	93,353.54	94,400.81	1,491	0.04	0.56	0.03	0.66
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
Avg %Δ price Sep-Jun	133.18	131.93	135.05	1,504	0.02	0.80	-0.02	0.71
Expect %Δ price Sep12-Jun13	117.26	132.06	117.38	1,510	-0.28	0.04	-0.00	0.97
2011 LR harvest (bags)	9.03	9.44	9.29	1,511	-0.03	0.66	-0.02	0.78
Net revenue 2011	-4,088.62	-5,878.45	-770.66	1,428	0.07	0.28	-0.13	0.31
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
Maize giver	0.26	0.27	0.25	1,589	-0.01	0.81	0.02	0.78
Delta	0.13	0.13	0.14	1,512	-0.01	0.90	-0.08	0.24

"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)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Pooled	By round	Purchase price	Sales prices	Pooled	By round	Pooled	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).

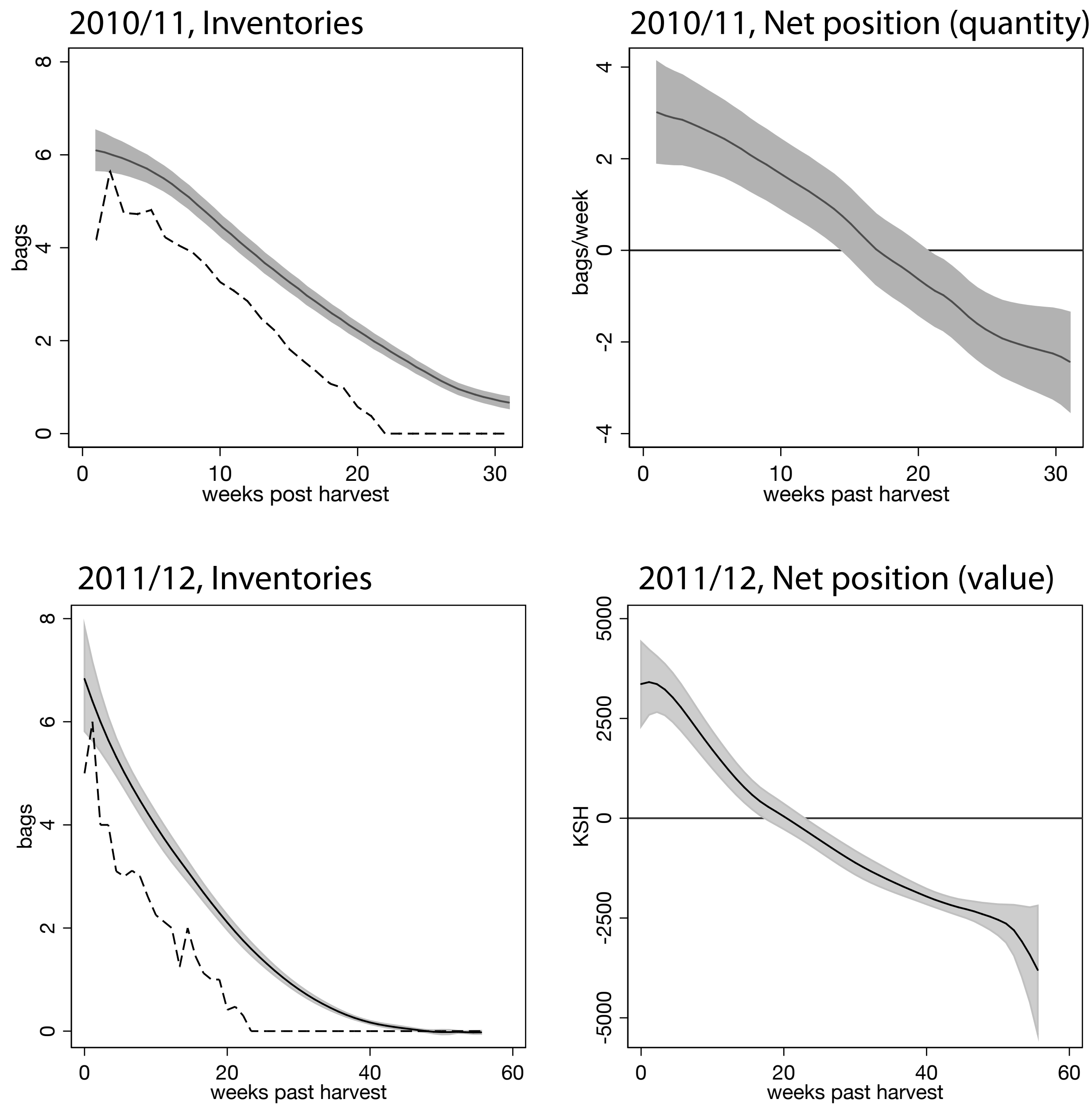


Figure A.2: **Study timeline**. The timing of the interventions and data collection are show at top, and the timing of the main agricultural season is shown at the bottom.

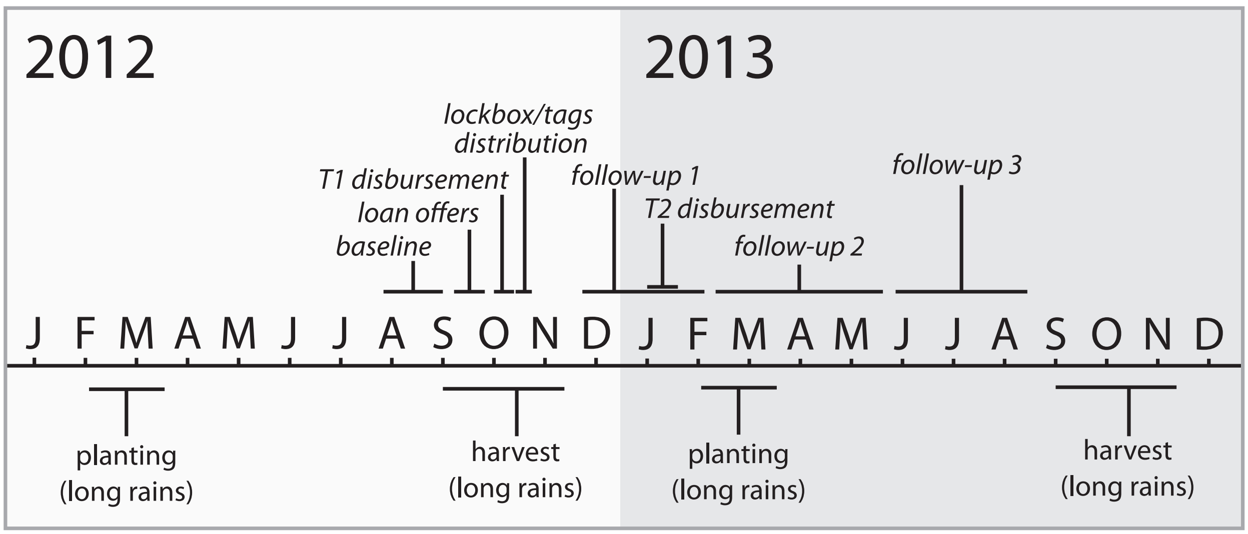


Figure A.3: **Robustness of price effect estimates**. *Left panel*: difference in prices between high and low-density markets over time for the full sample (black line) and for the sample with each sublocation dropped in turn (grey lines). *Center panel*: price effects under the "true" treatment assignment (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.

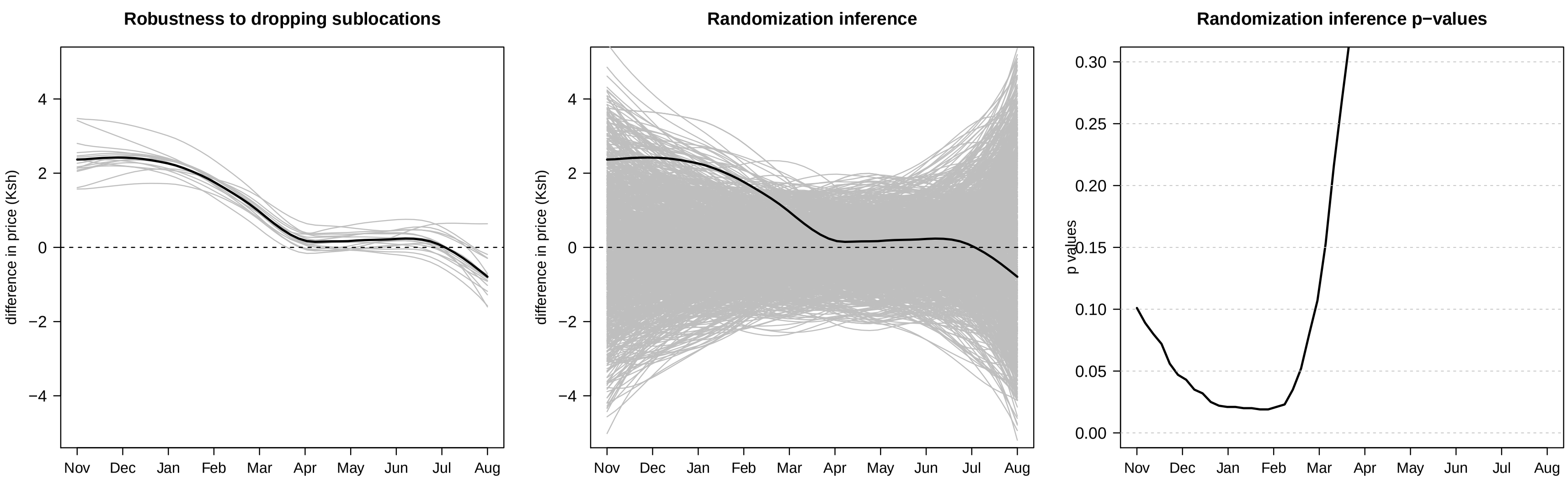


Table A.1: **Balance among baseline covariates, high versus low treatment intensity areas.** The first two columns give the means in the low or high treatment intensity areas, the 3rd column the total number of observations across the two groups, and the last two columns the differences in means normalized 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

When Commitment Fails
- Evidence from a Regular Saver Product in the Philippines*

Anett Hofmann†

Version 11 February 2014

(a current draft can be found here)

Abstract

Recent literature promotes commitment products as a new remedy for overcoming self-control problems and savings constraints. Committing to a welfare-improving contract requires knowledge about one's preferences, including biases and inconsistencies. If agents are imperfectly informed about their preferences, they may choose ill-suited commitment contracts. I designed a regular-instalment commitment savings product, intended to improve on pure withdrawal-restriction products by mimicking the fixed-instalment nature of loan repayment contracts. I conduct a randomised experiment in the Philippines, where individuals from a general low-income population were randomly offered to take up the product. Individuals chose the stakes of the contract (in the form of a default penalty) themselves. The result is that a majority appears to choose a harmful contract: While the intent-to-treat effect on bank savings for individuals assigned to the treatment group is four times that of a withdrawal-restriction product (offered as a control treatment), 55 percent of clients default on their savings contract. The explanation most strongly supported by the data is that the chosen stakes were too low (the commitment was too weak) to overcome clients' self-control problems. Moreover, both take-up and default are *negatively* predicted by measures of sophisticated hyperbolic discounting, suggesting that those who are fully aware of their bias realise the commitment is too weak for them, and avoid the product. The study suggests that research on new commitment products should carefully consider the risk of adverse welfare effects, particularly for naïve and partially sophisticated hyperbolic discounters.

*I would like to extend my gratitude to Oriana Bandiera, Maitreesh Ghatak, and Gharad Bryan, for their invaluable support and advice throughout this research project. I am deeply indebted to Ann Mayuga, Faith McCollister, Megan McGuire, Yoeri Suykerbuyk and Eva Ghirmai of IPA Philippines, without whom this project would not have been possible. I also thank the entire Survey and Marketer team of the IPA Regular Savers project. I am grateful for advice from Dean Karlan, Stefano DellaVigna, Matthew Levy, George Loewenstein, Lance Lochner, Francesco Nava, Johannes Spinnewijn, as well as from Betty Wilkinson and Michiko Katagami at the Asian Development Bank, and John Owens at the Rural Bankers' Association of the Philippines. Thanks to Jonathan de Quidt, Erina Ytsma, Claudia Steinwender and Michele Piffer for helpful comments and discussions. I also thank 1st Valley Bank of Cagayan de Oro, Philippines, for a productive collaboration. I gratefully acknowledge the financial support of the Yale Savings and Payments Research Fund at Innovations for Poverty Action, sponsored by a grant from the Bill & Melinda Gates Foundation. This research further received generous support from the Royal Economic Society. All errors and omissions are my own.

†Department of Economics and STICERD, London School of Economics. Email: a.hofmann1@lse.ac.uk.

1 Introduction

Commitment is popular. Contrary to predictions of the standard neoclassical model, the last decade has seen a surge of evidence documenting a demand for (self-)commitment contracts - roughly understood as a voluntary restriction of one's future choice set, in order to overcome intrapersonal conflicts.[1] Applications are as broad as the scope of human ambition, and range from gym memberships, diet clubs and pension savings to self-imposed binding deadlines for academic papers.[2] More informal arrangements include taking only a fixed amount of cash (and no credit cards) when going shopping, not keeping alcohol or chocolate in the house, and putting one's alarm clock at the other side of the room.[3] In the context of developing countries, documented demand for commitment devices goes back to the literature on rotating savings and credit organisations (ROSCAs),[4] the wandering deposit collectors of South Asia and Africa,[5] and more recent studies on newly introduced commitment savings products.[6]

Why do people self-commit? Commitment entails the voluntary imposition of constraints on future choices, thereby putting a cost on flexibility, which is weakly welfare-reducing from a neoclassical perspective. Three types of models are frequently cited to rationalise the observed demand for commitment: Models of quasi-hyperbolic discounting (Strotz (1955), Laibson (1997) and O'Donoghue and Rabin (1999)), models of temptation and self-control (Gul and Pesendorfer (2001), Gul and Pesendorfer (2004), Banerjee and Mullainathan (2010)), and dual-self models suggesting the existence of a long-run planning self and a short-run doing self (Fudenberg and Levine (2006)). All of these models generate preferences that are inconsistent over time, and generally suggest that agents are more impatient over current trade-offs (now vs. tomorrow) than over future trade-offs (one year vs. one year plus one day). As a result, they procrastinate activities that involve immediate costs and later rewards (saving for a new TV, going to the gym), and do too much of activities that involve immediate gratification but later costs (using high-interest credit cards, buying temptation goods). If individuals with such preferences realise their own time-inconsistency, they will have a positive willingness to pay for commitment devices which eliminate tempting options from their future choice sets (or make them more expensive), thus allowing them to follow through with their plans (to save, to eat healthily, to exercise). In theory, this will increase their welfare from an ex-ante (or long-run) perspective.

Is commitment a good idea? Especially in the development literature, the answer seems to be yes. Recent years have seen a multitude of papers promoting commitment savings in particular as a remedy for behavioural savings constraints, and thus as a possible way out of (credit-constraint based) poverty traps. Commitment savings have been hailed as increasing savings levels (Ashraf et al. (2006b)), agricultural input use (Brune et al. (2011)), pension contributions (Benartzi and Thaler (2004)), microenterprise investment (Dupas and Robinson (2013)), and chances of successful smoking cessation (Giné et al. (2010)).

[1]This paper focuses purely on self-commitment. It does not address commitment contracts adopted with strategic motives with respect to others. Furthermore, the paper abstracts from commitments entered into for convenience or other immediate benefits. As an example, the purchase of Christmas gifts in October qualifies as self-commitment if the agent fears not having enough money left in December, but not if the agent's motivation is purely to avoid the Christmas rush.

[2]See DellaVigna and Malmendier (2006) for gym memberships as a commitment device, Benartzi and Thaler (2004) for 401(k) pension savings, and Ariely and Wertenbroch (2002) for academic assignments.

[3]For an overview of commitment devices, see Bryan et al. (2010). For a humorous illustration, see popular articles and Internet videos on the 'money-shredding alarm clock '.

[4]See Besley et al. (1993), Anderson and Baland (2002), Ambec and Treich (2007) or Gugerty (2007).

[5]See e.g. Besley (1995) on West Africa's susu collectors.

[6]See Ashraf et al. (2006b), Brune et al. (2011) and Dupas and Robinson (2013) for the use of withdrawal-restriction savings accounts. Also see Duflo et al. (2011) for commitment to fertilizer use via advance purchase.

But are people good at choosing the ’right’ commitment contract? And if not, can commitment be harmful? By construction, correctly choosing a welfare-improving contract requires some knowledge about one’s preferences, including possible biases and inconsistencies: To determine whether a contract will enable an agent to follow through with a plan, the agent needs to anticipate how his future selves will behave under the contract. Consequently, if the agent is overconfident, or imperfectly informed about his own future preferences, the contract may result in undesirable behaviour, and the agent may be hurt, rather than helped. Given that the very nature of most commitment contracts is to impose penalties (usually of a monetary or social nature) for undesirable behaviour, adopting a commitment device that is ill-suited to one’s preferences may ’backfire’ and become a threat to welfare.[7]

This paper argues that commitment can be harmful if agents select into the wrong commitment contract - and that they frequently do. I conduct a randomised experiment where individuals could sign up for a new commitment savings account with fixed regular instalments, and where they are given the chance to choose the ‘stakes’ of the contract (in form of a default penalty) themselves. I find that the *average* effect on bank savings is large and significant: The Intent-to-Treat (ITT) effect on bank savings is roughly four times that of a conventional withdrawal-restriction product that was offered as a control treatment. However, a striking feature of the results is that the *median* client appears to choose a ‘harmful’ contract: 55 percent of clients default on their savings contract, and incur the associated penalty. The magnitude and timing of defaults is difficult to reconcile with rational expectations, and suggestive of individuals making ‘mistakes’ in contract choice. The explanation most strongly supported by the data is that the chosen stakes were too low (the commitment was too weak) to overcome clients’ self-control problems. In addition, both take-up and default are *negatively* predicted by measures of sophisticated hyperbolic discounting, consistent with the notion that those who are fully aware of their bias realise the commitment is too weak for them, and stay away. The results from a subsequent repeat marketing stage with the offer of ‘pre-ordering’ the product for a second round support the impression that a significant share of clients took up the commitment contract by mistake. Using a model of commitment under partial sophistication, I formally show that commitment contracts can reduce welfare if the commitment is not strong enough to discipline the agent, resulting in costly default. I further show that such insufficient commitment contracts may be selected by time-inconsistent agents with ‘partially sophisticated’ preferences - i.e. agents who are neither completely unaware nor fully aware of their time-inconsistency, but anywhere in between those two extremes. Alternative explanations for default that find some support in the data are income overoptimism, household conflict, and poor financial literacy. A pure stochastic shock explanation appears unlikely.

I partnered with 1st Valley Bank, a rural bank based in Mindanao (Philippines). The sample population of 913 individuals was obtained by conducting a door-to-door baseline survey in low-income areas in proximity to two selected bank branches. The baseline survey elicited time preferences, with a random half of individuals receiving real monetary rewards. Further elicited measures included perceived time-inconsistency, risk aversion, financial claims from others, cognitive ability, financial literacy, intrahousehold bargaining power, household demographics, and measures of saving, borrowing, and household expenditures. After the baseline survey, all individuals were provided with a marketing treatment, which included a personalised savings plan for an upcoming expenditure and a free

[7] Consider any type of commitment contract with front-loaded fees, such as retirement savings products with acquisition or management costs. Fees are generally subtracted from the contributions during the first few years of the contract, generating a ’J curve’ in the asset value. Cancelling or defaulting on the contract during early years generates high negative returns. A similar argument can be made for front-loaded gym membership costs.

non-commitment savings account with 100 pesos (U.S. \$2.50) opening balance.[8] Personal savings plans featured a self-chosen goal date, goal amount, and a weekly or bi-weekly instalment plan. The idea was to encourage individuals to save for their lump-sum expenses (such as school fees, business capital, or house repairs), rather than following the common practice of borrowing at high informal moneylender rates. At the end of the marketing visit, a randomly chosen 50 percent (the 'Regular Saver' group) were offered a new commitment product called 'Achiever's Savings Account' (ASA). ASA committed clients to make fixed regular deposits and pay a penalty upon default, which effectively made all features of the personal savings plan binding. The default penalty was chosen by the client upon contract signing, and framed as a charity donation.[9] There was no compensation for the restrictions, no added help (such as deposit collectors or text message reminders), and a standard market interest rate.[10]

As a control treatment, an additional 25 percent of the sample (the 'Withdrawal Restriction' group) were offered the commitment savings account studied in Ashraf et al. (2006b), Giné et al. (2010), Brune et al. (2011), and Karlan and Zinman (2013): The 'Gihandom' savings account (Visayan for 'dream') allowed individuals to commit to either the goal date or the goal amount from their savings plan, by restricting withdrawals before the goal had been reached. This account did not include any obligation to make further deposits after the opening balance. The remaining 25 percent of the sample received no further intervention after the marketing treatment, and constitute the control group. For those in the control group (and those who rejected the commitment accounts), none of the savings plan features were binding. Since individuals' expenditures were due at different times, the outcome variable of interest are individuals' savings at the time of their goal date. The study concluded with a comprehensive endline survey, a 'customer satisfaction survey' for ASA clients, and a repeat marketing stage where ASA clients could opt to 'pre-order' the product for a second round.[11]

I find that demand for commitment is high, even in a general low-income population with little previous bank exposure: Take-up rates were 27 percent for ASA and 42 percent for Gihandom, in spite of the fact that all individuals were given a free standard savings account (with 100 pesos) immediately prior to the commitment offer.[12] Offering ASA was more effective at increasing savings: By the time individuals reached their goal date (an average of 130 days later), bank savings in the Regular Saver group had increased by 585 pesos (U.S. \$14) relative to the control group, whereas bank savings in the Withdrawal Restriction group had increased by 148 pesos (U.S. \$3.50, as measured by the Intent-to-Treat effect) relative to the control.[13] The control group saved an average of 27 pesos. The scale of effects suggests that a commitment product with fixed regular instalments is highly effective at increasing savings on *average*. However, this average hides a lot of heterogeneity in the case of both products: 55 percent of ASA clients defaulted on their savings contract, incurring penalties (charity donations) between 150 and 300 pesos - the equivalent of a day's wage (the stated treatment effect already accounts for these charges). The penalty for the withdrawal-restriction product Gihandom was less salient, but existent: 79 percent of Gihandom clients made no further deposits after the opening balance. For those

[8] At the time of marketing (October 2012), the exchange rate was 42 Philippine pesos per U.S. dollar.

[9] The concept is roughly comparable to the Stickk initiative (www.stickk.com), where people are asked to set their own stakes, but applied to the requirements of a developing country context.

[10] The bank's standard interest rate as of September 2012 was 1.5 percent per annum, and decreased to 1 percent in November 2012. This interest rate was the same across all offered accounts. The inflation rate for 2012 was 3.1 percent.

[11] Pre-orders were not legally binding, but involved a cost through substantial paperwork.

[12] The difference in take-up rates may be partly driven by liquidity concerns: ASA required an opening balance equal to the first weekly deposit (minimum 150 or 250 pesos), whereas Gihandom could be opened with 100 pesos.

[13] The Intent-to-Treat (ITT) effect measures the effect of being *offered* the product. An increase of 585 pesos (148 pesos) corresponds to 27 percent (7 percent) of median weekly household income in our sample.

who had chosen to make their goal amount binding (45 percent), this meant their initial savings were tied up indefinitely, or until the bank would exhaust their account with dormancy fees.[14]

Using conventional measures of actual time-inconsistency and a novel measure of *perceived* time-inconsistency (sophistication), the data suggest that present-biased preferences by themselves do not predict take-up of a commitment product, but they do predict default. In contrast, sophistication drives both take-up and default: As an agent's degree of sophistication rises, he becomes less likely to adopt commitment, and less likely to default, conditional on take-up. This is consistent with the interpretation that partial sophistication about time-inconsistency leads agents to adopt weak commitment contracts, and subsequently default. Highly sophisticated agents are more cautious about adopting commitment, but have higher chances of success when they do choose to commit. The notion of 'weak commitment' is supported by the observation that 80 percent of ASA clients chose the minimum admissible default penalty (150 or 250 pesos, depending on the savings goal). Finally, the data is strongly bi-modal, in the sense that almost all clients either (i) stop depositing immediately after the opening balance or (ii) complete their savings plan in full. I interpret this as evidence against a shock explanation, where individuals rationally default following large random shocks to their income or expenditures.

This paper builds and expands on the literature in three ways. First, to the author's knowledge, it is the first paper to explicitly discuss adverse effects of a commitment product in a savings context, and one of the first papers to do so for commitment devices in general. This makes it closest in spirit to DellaVigna and Malmendier (2006), who show that U.S. consumers choose gym contracts which are cost-inefficient given their attendance frequency. It also relates to the theoretical work of DellaVigna and Malmendier (2004). In the realm of commitment savings, the literature has largely focussed on positive *average* effects, highlighting the promising role that commitment savings could play in overcoming behavioural savings constraints. However, welfare inference critically depends on the distribution of effects in the population. I establish that these effects may be very heterogeneous, including the possibility of a majority being hurt by the product. The results of this paper complement previous findings: Ashraf et al. (2006b) find that a withdrawal-restriction product increased savings by 81 percent on average after 12 months, but 50 percent of the 202 clients made no further deposits after the opening balance of 100 pesos. Out of 62 clients who selected an amount goal, only six reached this goal within a year, suggesting that the remainder may have their initial savings tied up indefinitely.[15] Similarly, Dupas and Robinson (2013) document that offering Kenyan women savings accounts with withdrawal restrictions led to a 45 percent increase in daily business investment on *average*, but 43 percent of women made no further deposits after opening the account. Finally, Giné et al. (2010) offered smokers in the Philippines a commitment contract for smoking cessation, in which smokers would deposit savings into a withdrawal-restriction account, and forfeit their savings to charity if they failed a nicotine test after 6 months. The authors point out that offering the commitment contract increased the likelihood of smoking cessation by 3 percentage points. Looking at heterogeneity, 66 percent of smokers who took up the product failed the nicotine test, forfeiting an average of 277 pesos in savings. Interestingly, those who defaulted on their contract had chosen much lower stakes relative to those who succeeded (successful quitters saved 1,080 pesos on average). While the direction of causality is unclear, this is consistent with the idea that individuals tend to choose commitment products which are too weak to overcome their

[14]Dormancy fees are very common with Philippine banks, and commonly start after two years of inactivity.

[15]Neither Ashraf et al.'s SEED product, nor the Gihandom product used in this study are fool-proof, in the sense that clients could have borrowed the goal amount for five minutes from a friend, deposited it at the bank, and received their savings back. Neither study finds any evidence that this happened.

self-control problems. In summary, a closer look at the heterogeneity behind average treatment effects in the literature reveals that adverse effects of commitment products may be widespread.

Second, the paper provides the first analysis of a commitment savings product with fixed regular instalments in a randomised setting. The product design mimics the fixed instalment structure found in loan repayment contracts, thus providing an incentive to make regular future deposits and smooth consumption. It is motivated by empirical evidence suggesting that microloans and informal loans are often taken out for consumption purposes, or for recurring business expenditures - rather than as a one-off investment.[16] With loans that are not directly required for income generation, the question arises why individuals are willing to pay substantial loan interest charges rather than choosing to save. Especially for those who borrow in frequent cycles, the long-term difference between expensive loan cycles and equivalent savings cycles reduces to (i) one initial loan disbursement and (ii) a binding fixed-instalment structure that is rarely available in savings products.[17] The idea that time-inconsistent agents benefit from commitment to regular fixed instalments has been suggested by Fischer and Ghatak (2010), Bauer et al. (2012), and Hofmann (2013). If it is true that a significant share of the demand for microcredit and informal borrowing is just a demand for commitment to fixed instalments, then we should expect to see that the introduction of a fixed-instalment microsavings product will result in (i) substantial increases in saving and (ii) a reduction in the demand for loans. I find strong support for an increase in savings, and a large but statistically insignificant reduction of loan demand. Furthermore, the paper provides the first direct comparison of a regular-instalment savings product with a pure withdrawal-restriction product.[18] I estimate an average effect on bank savings that is about four times the effect of the withdrawal restriction product. This is consistent with the theoretical work of Amador et al. (2006), who show that when individuals value both commitment and flexibility, the optimal contract involves a minimum (per-period) savings requirement.

Third, the paper proposes a novel measure of sophistication for time-inconsistent agents. Previous literature has often assumed a one-to-one mapping from the take-up of a commitment product to the presence of (fully) sophisticated time-inconsistency (with the notable exception of Tarozzi and Mahajan (2011), who follow a structural approach). Such a one-to-one mapping does not allow for the possibility that individuals may take up commitment products by mistake. I propose a survey-based measure of sophistication, which relies on the interaction between observed time-inconsistency (as measured by conventional time discounting questions), and measures of self-perceived temptation.

The paper proceeds as follows. Section 2 describes the experimental design employed. Section 3 explains the survey instrument, with particular view to the measurement of time-inconsistency and sophisticated hyperbolic discounting. Section 4 outlines the empirical strategy. Section 5 presents empirical results. Section 6 outlines a model of commitment under partial sophistication. Section 7 discusses other explanations. Section 8 concludes.

[16] See e.g., Mullainathan et al. (2007).

[17] Informal arrangements like ROSCAs may constitute an exception, but they are inflexible to an individual member's needs. Also, ROSCAs were not widely available in the study region. Deposit collectors (see Ashraf et al. (2006a)), if available, provide another alternative - but a deposit collection service does not commit the individual to deposit any particular amount, and the individual may be tempted to deposit the minimum necessary to avoid social sanctions.

[18] The withdrawal-restriction product tested in this study (Gihandom) directly corresponds to the SEED product in Ashraf et al. (2006b): Terms and Conditions are identical, and the study locations are 70km (2h by local bus) apart. The magnitude of estimated effects is comparable, considering that Ashraf et al. (2006b) estimate an ITT of 411 pesos after 12 months in a sample of previous savings account holders, whereas I estimate an ITT of 148 pesos after 4.5 months (on average) in a general low-income population.

2 Experimental Design

I designed and implemented the commitment savings product ASA in cooperation with 1st Valley Bank, based in Mindanao, Philippines. 1st Valley Bank is a small rural bank that offers microcredit, agricultural insurance, salary loans, and pension products to its clients. The bank agreed to offer both the regular-instalment product ASA and the withdrawal-restriction product Gihandom in two of their bank branches: Gingoog and Mambajao. Gingoog is a city of 112,000 people in northern Mindanao, and Mambajao is a municipality of 36,000 people on Camiguin Island. For these two bank branches, both ASA and Gihandom constituted new product additions at the time of the study.[19]

The sample was obtained through door-to-door visits in all low and middle income areas in proximity to the bank branches. In each household, the survey team identified the person in charge of managing the household budget (in 94 percent of the cases, this was the mother of the family), and invited them to take part in a household survey. The baseline survey was completed with all individuals who (i) had some form of identification,[20] (ii) claimed to have a large upcoming expenditure (such as school fees, house repairs, or business expansions) and (iii) agreed to receive a visit from a financial advisor (to talk about how to manage household expenses). After the baseline survey, individuals were randomly assigned to three groups: 50 percent of individuals were assigned to a 'Regular Saver' (R) group, and 25 percent each were assigned to a 'Withdrawal Restriction' (W) and a control (C) group.

Approximately one week after having completed the household survey, individuals received a visit from a bank marketer. Of 913 surveyed individuals, 852 could be re-located.[21] Marketers engaged individuals in a conversation about how to manage large lump-sum expenses, and talked about the benefits of saving. Focussing on one particular expenditure, individuals were encouraged to make a formal 'Personal Savings Plan', which contained a purpose, a goal amount, a goal date, and a fixed equal instalment plan with due dates (see Figure 10). Median savings goals were 2400 pesos across all groups (roughly comparable to a median household's weekly income of 2125 pesos), with a median weekly instalment of 150 pesos. Common savings goals were school tuition fees, house repairs, and Christmas gifts (see Table IX for an overview of savings plan characteristics). The duration of savings plans was limited to 3–6 months, so that the outcome could be observed by the study. The median duration was 137 days. In addition, everyone was offered a standard non-commitment savings account (henceforth called 'ordinary savings account') as a 'welcome gift' from the bank, along with an encouragement to use this account to save for the expenditure. This ordinary savings account contained a free 100 pesos opening balance, which also constitutes the minimum maintening balance.[22]

At the end of the visit, individuals in group R were asked whether they wanted to commit to the fixed-instalment structure outlined in their Personal Savings Plan by taking up the ASA product, and the product features were explained. In contrast, individuals in group W were offered the option to restrict withdrawals of their savings until they reached either the goal amount or the goal date specified in their Personal Savings Plan, implemented through the use of the Gihandom product. It is to be

[19]Gihandom had been previously offered at other 1st Valley Bank branches, but not at the two branches in question.

[20]A valid form of identification was required by the bank to open a savings account. Accepted forms of identification included a birth certificate, tax certificate, voter's ID, barangay clearance, and several other substitutes.

[21]A test for equality of means in the probability of being reached by the marketer across treatment groups yields an F-statistic with a p-value of 0.16. Individuals in group R were as likely to be reached as individuals in group C, but slightly less likely than individuals in group W.

[22]Individuals were able to close this account and retrieve funds by visiting the bank, but incurred a 50 peso closing fee. During the period of observation (September 2012 until 15 April 2013), no client closed their account.

expected that the marketing treatment itself influenced individuals' savings behaviour, as evidenced in the literature on mental accounting.[23] However, up to the point of offering the commitment products, the marketing script was identical across all groups, which prevents a bias of the estimated treatment effects. The effect of marketing as such cannot be identified, as there was no marketing-free group. However, the fact that the control group saved an average of 27 pesos until their respective goal dates indicates that, given a non-negativity constraint on bank savings, the effect of marketing was small.

Out of 852 individuals located for the marketing visit in September and October 2012, 788 accepted the free ordinary savings account, and 748 agreed to make a savings plan. In group R, 114 clients (out of 423 offered) accepted the ASA product.[24] In group W, 92 (out of 219 offered) accepted the Gihandom product. Table X summarizes the take-up results.

The regular-instalment product ASA committed clients to a specific instalment plan with weekly or bi-weekly due dates. An account was considered in default from the day the client fell *three* instalments behind. In case of a default, the account was closed, an 'Early Termination Fee' was charged, and any remaining savings were returned to the client. A termination fee that is directly linked to the instalment structure distinguishes the ASA product from withdrawal-restriction or standard accounts, and represents its key commitment feature. No fee was charged after sucessful completion. The amount of the fee was chosen by the client upon signing the ASA contract: Each client signed a 'Voluntary Donation Form', which specified a termination fee that would be donated to charity in case of a default. Clients were given a choice of three national (but not locally-based) Philippine charities.[25] While the instalment structure may appear rigid at first sight, a variety of flexibility features were included to allow clients to adapt to changing circumstances: First, clients could fall up to two instalments behind at any given time, making it theoretically possible to miss every other instalment, and pay a double instalment in alternate weeks. To encourage timely depositing, a small 10 peso ($0.25) admin fee had to be paid upon making up a missed past instalment, but this fee did not accumulate over time. Deposits towards future weeks could be made at any time, as long as they were in increments of the weekly instalment. This was a practical requirement, as the client's progress was monitored by making ticks on a collection card for each successful week (see Figure 10). The possibility of making future deposits early effectively provided a form of insurance against uneven income streams. Withdrawals during the savings period were only possible by allowing default to occur.

Enforceability of the termination fee was facilitated through the account opening balance: To complete the opening of an ASA, clients had to deposit an account opening balance equal to their first weekly instalment, but at least 150 pesos for savings goals below 2500 pesos, and at least 250 pesos for savings goals of 2500 pesos and above. The same threshold applied for the termination fee: Clients could choose a termination fee as high as they liked, but no lower than a minimum of 150 or 250 pesos, respectively. Consequently, the minimum termination fee could always be enforced. Higher termination fees could be enforced only if the client continued to save, or if their opening balance exceeded the minimum. Note that by nature of the contract, all ASAs were either successfully completed or in default by the goal date,[26] and any remaining savings were transferred to clients' ordinary savings accounts.

[23]Most prominently, see Thaler (1985) and Shefrin and Thaler (1988).

[24]One member of the control group was mistakenly offered the ASA product and accepted, which means a total of 115 ASA accounts were opened. This constitutes a mild contamination of the control group, and means that the estimated ASA treatment effect is a lower bound on the true effect.

[25]Altruism and attitudes towards charities were measured in the baseline, and are available as a control variable.

[26]After the goal date, there was a one-week grace period to make any outstanding deposits, but no client made use of this.

The withdrawal-restriction account Gihandom was simpler in structure: Clients chose to restrict withdrawals before either their goal date or their goal amount (specified at contract signing) was reached. Out of 92 Gihandom clients, 39 chose the amount goal, and 53 chose the date goal. The goal amount can be interpreted as the stronger restriction, since additional deposits need to be made in order to receive savings back. Formally, there was no limit on how long individuals could take to reach the goal amount. However, as is common for Philippine banks, significant dormancy fees were applied after two years of inactivity. While the marketers encouraged individuals to deposit the first weekly instalment from their savings plan as an opening balance, the formal minimum opening balance for Gihandom was 100 pesos. The difference of 50–150 pesos (depending on the savings goal) in the mandatory opening balances between ASA and Gihandom is a possible explanatory factor in the difference between take-up rates. Finally, two features were common to both ASA and Gihandom: First, opening balances for both products were deliberately collected one week *after* contract signing. The practical motivation behind this was to give individuals time to prepare for the expense. The theoretical motivation was to free the decisionmaker from temptation in the contract-signing period – a sophisticated hyperbolic discounter should choose a welfare-maximising contract when asked in period 0, but not necessarily when asked in period 1.[27] Second, both products shared the same emergency provisions: In case of a medical emergency or death in the family, a relocation to an area not served by the bank branch, or a natural disaster (as declared by the government),[28] clients could close their account and access their savings without any penalties. Within the six months of observation, no client exercised this option.

In order to identify the treatment effect of a commitment to fixed regular instalments, individuals were left to themselves during the savings period, without any help from deposit collectors or reminders. After all goal dates had been reached, a comprehensive endline survey was administered. The endline survey focused on all types of savings (including at home and in other banks), outstanding loans, expenditures, changes in income, and various types of shocks experienced since the baseline survey. In addition, existing ASA clients were offered the option to sign up for a 'Pre-Order' of the product: Clients were informed that the bank may decide to offer ASA for a second round, conditional on sufficient demand. While the Pre-Order did not involve a financial commitment, it involved the completion of a new savings plan, a new 'Voluntary Donation Form', and a decision on a new termination fee (to deter cheap talk).

3 The Survey Instrument

The household survey administered at the beginning of the study had two objectives: First, to measure factors commonly suspected to influence the demand for (commitment) savings products. Second, the survey data on savings, loans, income, and expenditures provides the baseline for the estimation of treatment effects (see Section 5.1).

[27] This approach is similar to that in Benartzi and Thaler (2004), who let employees commit to allocate *future* salary increases to their pension plan. It could be argued that the late collection of opening balances effectively just delayed when individuals entered the commitment contract. In a purely financial sense, this is true. However, signing the contract was associated with substantial paperwork, as well as a non-financial commitment to the marketers, who personally collected the opening balance after one week. Out of 159 individuals who initially signed the ASA contract, 45 failed to deposit an opening balance. The corresponding number for Gihandom is 24 out of 116 initially signed contracts.

[28] Provided appropriate documentation, i.e. a hospital bill, death certificate, or proof of relocation.

I measured time-inconsistent preferences using the common method of multiple price lists (MPLs): Individuals were asked to choose between a fixed monetary reward in one period and various larger rewards in a later period. A randomly chosen half of the sample received real rewards, whereas for the other half, the questions were hypothetical. After a set of questions using a near time frame (now versus one month), the same set of questions was repeated for a future time frame (one month versus two months). The outcome of interest was the size of the later reward necessary to make the individual switch from preferring the (smaller) earlier reward to the (larger) later reward. For illustration, consider the following sample questions:

1. Would you prefer to receive P200 guaranteed today, or P250 guaranteed in 1 month?

2. Would you prefer to receive P200 guaranteed in 1 month, or P250 guaranteed in 2 months?

The earlier reward was kept constant at 200 pesos, while the later reward gradually increased from 180 to 300 pesos. Individuals whose preferences satisfy standard exponential discounting will be time-consistent – i.e., the amount necessary to make them switch from the earlier reward to the later reward will be the same whether they are asked to choose between now and one month, or between one month and two months. I identify as hyperbolic discounter those who are impatient in the present, but patient in the future, i.e., the reward needed to make them wait for one month is larger in the present than in the future (thus the term 'present biased'). In the opposite direction, individuals who exhibit more patience now than in the future are classed as 'future biased'. An individual who always prefers the earlier reward in all questions (for both near and future time frames) is classified as 'impatient'. The two sets of questions were separated by at least 15 minutes of other survey questions, in order to prevent individuals from anchoring their responses to earlier answers. An ad-hoc randomisation based on individuals' birthdays determined who played the game with real rewards (see Appendix B for details). For those with real rewards, one of their choices was paid out, selected at random by drawing a ping pong ball with a question number from a black bag. To prevent uncertainty about whether future payments would be guaranteed (causing an upward bias of the present-bias measure), both cash and official post-dated bank cheques were presented during the game.

I find 16.6 percent of individuals to be present-biased, and 18.9 percent of individuals to be future-biased. No systematic difference is apparent between those offered real and those offered hypothethical rewards.[29] These estimates are slightly below comparable estimates in the literature, but show a similar tendency for future bias to be as common as present bias (Ashraf et al. (2006b) find 27.5 percent present-biased and 19.8 percent future-biased, Giné et al. (2012) find 28.5 and 25.7 percent, respectively, Dupas and Robinson (2013) find 22.5 and 22 percent, and both Brune et al. (2011) and Sinn (2012) find 10 percent present-biased and 30 percent future-biased). Explanations that have been proposed for future bias include utility from anticipation (Loewenstein (1987), Ameriks et al. (2007)), varying degrees of future uncertainty (Takeuchi (2011), Sayman and Öncüler (2009)), and survey noise.

In addition to a standard measure of preference reversals, it is vital to the analysis to obtain a measure of sophistication. In particular, this measure should not in itself be derived from a demand for commitment. The approach pursued in this paper relies on a simple idea: Multiple price lists provide a measure of actual time-inconsistency, independent of an individual's awareness of said inconsistency.

[29] At 17.9 percent, present bias was more frequent among those with hypothetical rewards, than among those with real rewards (15.2 percent), but the difference is not significant. This suggests that a bias from uncertainty is unlikely. A detailed comparison of real and hypothetical incentives is beyond the scope of this paper, and will be provided in a separate working paper on the elicitation of time preferences.

If *observed* inconsistency could be interacted with a measure of *perceived* inconsistency, a measure of sophisticated hyperbolic discounting would result.

Such an awareness measure exists: The self-control measure proposed by Ameriks et al. (2007), henceforth referred to as ACLT. Using survey questions, the authors elicit how individuals would *optimally* like to allocate a fixed resource over time. They then ask which allocation individuals would be *tempted* to consume (if not exercising self-control), and finally, which allocation they *expect* they would consume in the end. While originally intended to identify the parameters of the Gul and Pesendorfer (2001) model, the questions require the individual to critically assess future temptations and their (hypothetical) response to them. The resulting measure reflects an individual's *perceived* (rather than actual) self-control problems. This makes the ACLT questions, interacted with a measure of observed time-inconsistency (e.g., through MPLs), a promising candidate to measure sophistication.

The setup is as follows: Respondents were presented with a hypothetical scenario of winning 10 certificates for "dream restaurant nights". In this scenario, each certificate entitled the holder and a companion to an evening at any local restaurant of their choice, including the best table, an unlimited budget for food and drink, and all gratuities. The certificates could be used starting immediately, and would be valid for two years. Any certificates not used after two years would expire. I presented the ACLT scenario along with an example list of local middle-class restaurants which were chosen to be above what respondents could usually afford, and which were regarded as highly desirable. This was intended to prevent simple substitution of certificates into everyday consumption (given the low income levels in the sample, respondents were used to eat either at home, or in simple street eateries, *carinderias*). In addition, the restaurant framing has the added benefit of being directly linked to consumption, thus avoiding the common concern with cash rewards that money is fungible and does not have to be associated with an immediate consumption shock (cf. Frederick et al. (2002)). In line with the ACLT design, I then asked the following questions:

1. Think about what would be the ideal allocation of these certificates for the first and the second year. From your current perspective, how many of the ten certificates would you ideally like to use in year 1 as opposed to year 2?

2. Some people might be tempted to depart from this ideal allocation. For example, there might be temptation to use up the certificates sooner, and not keep enough for the second year. Or you might be tempted to keep too many for the second year. If you just gave in to your temptation, how many would you use in the first year?

3. Think about both the ideal and the temptation. Based on your most accurate forecast of how you would actually behave, how many of the nights would you end up using in year 1 as opposed to year 2?

The answers to these questions provide two important measures: Perceived self-control (from $(3) - (1)$, $expected - ideal$) and perceived temptation (from $(2) - (1)$, $tempted - ideal$). However, these measures were designed for the Gul-Pesendorfer model, which does not directly translate into the $\beta\delta$-model of hyperbolic discounting which underlies this analysis. The models are not nested, and there is no direct equivalent to self-control and temptation in the model of hyperbolic discounting. From the perspective of the $\beta\delta$-model (where self-control does not exist), we would expect the two measures to be the same - namely the difference between the optimal ex-ante allocation, and (the individual's perception of) the allocation that results in a subgame perfect equilibrium between the different selves. Following this logic, both measures are equally suitable to assess an individual's awareness of their time-inconsistency.

For the purposes of my empirical analysis, I choose to focus on $tempted - ideal$ as an awareness measure for time-inconsistency. I then interact awareness of time-inconsistency with observed time-inconsistency (in MPLs), and obtain a measure of sophisticated hyperbolic discounting: $tempted - ideal * presentbias$.[30] The reason for focussing on the temptation measure is as follows: Suppose costly self-control does exist. An individual who exercises full self-control, and thus realises the (ex-ante) ideal allocation, might still have a demand for commitment. While a commitment device would not change the de-facto allocation he consumes, it can increase his utility by removing temptation, and thus the need to exercise costly self-control. Therefore, a low or zero measure of $expected - ideal$ (i.e., good self-control) might still be associated with a demand for commitment, while time-consistent preferences would not. As a result, for the purposes of analyzing demand for a commitment savings product, the temptation measure provides a better indication of whether individuals feel they could benefit from commitment. In this sense, perceived temptation is closely related to the concept of sophistication.

I observe that 81.6 percent of individuals report strictly positive values of temptation, with a median temptation of two certificates. Given the much lower frequency of observed present bias, the question arises how to interpret temptation without present bias. This paper remains agnostic about the precise theoretical connection between models of self-control and those of hyperbolic discounting, and instead offers a simple intuition: If an individual reports to be tempted, but behaves in a time-consistent fashion, this may be due to the exercise of self-control. This hypothesis is supported by the data: Conditional on a given level of temptation, non-present biased individuals report significantly better self-control than present-biased individuals.

In addition to the measures for present bias and sophistication ($tempted - ideal * presentbias$), the survey obtained measures of financial claims from others, risk aversion, cognitive ability, financial literacy, bargaining power within one's household, distance to the bank branch (via GPS coordinates), and frequency of income or expenditure shocks, as well as an indicator for having an existing bank account. These measures are discussed in Appendix B.

Table I presents summary statistics for the main observed covariates from the survey. Tests for equality of means across treatment groups were conducted to verify that the randomisation was balanced. Randomisation into treatment groups occurred shortly after the baseline survey, which means that covariates were available at the time of randomisation. A star next to a variable in Table I indicates that the randomisation was stratified on this variable. In three cases, means were statistically different across treatment groups: Income, impatience and risk aversion. Income and impatience have no predictive power in any of the later regressions. In particular, wealthier individuals are no more likely to take up a commitment product than poorer individuals. Risk aversion does have predictive power for the take-up of Gihandom (W-group). Robustness checks are reported in Table XI and Table XII.

4 Empirical Strategy

The primary objective of the study was to analyse the demand for and the effects of a regular-instalment commitment savings product, and to compare its performance to traditional withdrawal-restriction

[30] I censor values of temptation and self-control at zero. I interprete observed negative values as measuring something other than temptation and self-control – e.g., not having time to go to restaurants as often as individuals would ideally like, or inability to understand the survey question. Negative values occurred in 4 (42) out of 910 cases for temptation (self-control).

TABLE I: SUMMARY STATISTICS BY TREATMENT ASSIGNMENT

	R-Group	W-Group	Control	F-stat P-value
Age*	43.8337	43.4493	44.25	0.8039
	(0.6029)	(0.8214)	(0.8412)	
Female*	0.9409	0.9430	0.9430	0.9912
	(0.0110)	(0.0154)	(0.0154)	
Education (yrs)	10.55604	10.39207	10.56388	0.8398
	(0.1662)	(0.2417)	(0.2513)	
HH Income	2890.89	2485.78	3194.43	0.0481
	(124.26)	(165.13)	(272.45)	
#HH members	5.07221	5.179825	5.429825	0.1081
	(0.0909)	(0.1399)	(0.1398)	
Real Rewards*	0.5033	0.5219	0.5263	0.8371
	(0.0234)	(0.0332)	(0.0331)	
Financial	0.3934	0.3877	0.3860	0.9867
Claims*	(0.0229)	(0.0324)	(0.0323)	
Existing	0.4683	0.4649	0.4254	0.5176
Savings Account	(0.0234)	(0.0331)	(0.0328)	
Impatience	0.3217	0.4035	0.3333	0.0959
	(0.0219)	(0.0326)	(0.0313)	
Present Bias*	0.1723	0.1614	0.1560	0.8388
	(0.0180)	(0.0246)	(0.0246)	
Perceived	2.3838	2.1850	2.4714	0.2249
Temptation	(0.0889)	(0.1122)	(0.1210)	
Risk Aversion	4.2254	4.6360	4.1316	0.0104
	(0.0932)	(0.1219)	(0.1289)	
Cognitive	2.9365	2.8860	2.9342	0.8870
Ability	(0.0592)	(0.0889)	(0.0955)	
Financial	1.8556	1.8377	1.8509	0.9767
Literacy	(0.0466)	(0.0682)	(0.0694)	
N	457	228	228	913

Note: A starred variable indicates that the randomisation was stratified on this variable.

commitment products. Given the heterogeneity of results, a deduced objective is to document possible risks of commitment contracts, in particular with respect to partially sophisticated hyperbolic discounting. The main outcomes of econometric interest are a range of treatment effects (on bank savings, other savings, loan demand, and expenditures), as well as predictors of take-up, contract outome (successful or default), and the pre-order decision (comparable to repeat take-up).

For the estimation of treatment effects, denote by R_i an indicator variable for assignment to the 'Regular Saver' treatment group – all individuals in this group were offered the ASA product. Denote by W_i an indicator variable for assignment to the 'Withdrawal Restriction' group – all individuals in this group were offered the Gihandom product. Treatment effects can be estimated using the equation

$$\Delta Y_i = \alpha_0 + \alpha_R R_i + \alpha_W W_i + \epsilon_i \tag{1}$$

where ΔY_i denotes the change in the outcome variable of interest. In Section 5.1, I focus on bank savings, but also provide estimates for total savings, loan demand, and expenditures (see Figure 6). Bank savings refers to the change in savings held at 1st Valley Bank. For ASA clients, this is the sum of their savings in ASA plus their savings in the non-commitment savings account provided to them. For Gihandom clients, it is the sum of their Gihandom savings plus their savings in the ordinary account. For everyone else (i.e., the control group and those who rejected the commitment product offered to them), bank savings refer to their ordinary savings account only (recall that individuals were encouraged to use the ordinary savings account to follow the personal savings plan provided to them). Summing all existing savings accounts per individual means that crowd-out between savings devices at the bank will not impact the analysis. However, individuals could have substituted away from home savings, or savings at other banks. To observe such effects, I also analyse a measure of other savings, obtained from survey data, which includes home savings, money lent out to others or safekept elsewhere, and money at other banks. The time frame for measuring savings runs from the date of the baseline survey visit to the individual's savings goal date – i.e., savings durations vary at the individual level. This is a consequence of focusing the marketing on particular expenditures: If savings were measured at the end of the study, even a successful saver would have a savings balance of zero if he has already paid for the expenditure.

An OLS estimation of equation 1 provides $\hat{\alpha}_R$ and $\hat{\alpha}_W$ – estimates of the Intent-to-Treat effects of the regular-instalment product ASA and the withdrawal-restriction product Gihandom. The ITT measures the mean causal effect of having been *offered* the product, which is likely to be an average of the effect of using the product, and of simply feeling encouraged to save because of the product offer. However, as has been outlined in Section 2, individuals in all groups received an identical marketing treatment. Only after a personal savings plan for an upcoming expenditure had been made, and an ordinary savings account had been opened, did the marketers offer individuals in groups R and W the possibility to bindingly commit to selected features in their savings plan. Under the assumption that the mere offer of commitment has no effect on savings (other than via encouraging people to use the product), the ITT will be a composite of the Treatment-on-the-Treated effect (TOT) on those who took up the product offered to them, and a zero effect (relative to the control) on those who did not take up the product.[31] In this case, the TOT can be estimated by dividing the ITT ($\hat{\alpha}_R$ and $\hat{\alpha}_W$) by the fraction of take-ups. Alternatively, equation 1 can be estimated using an instrumental variables approach, with

[31]See Imbens and Angrist (1994) and Duflo et al. (2007) for a discussion on ITTs and local average treatment effects.

takeup (ASA_i and $Gihandom_i$) as the regressors and assignment to treatment (R_i and W_i) as orthogonal instruments.

Predictors of the take-up, default and pre-order decision can be summarized in a binary choice equation. I use a probit model to estimate

$$Choice_i = \beta_0 + \beta X_i + \epsilon_i,$$

where $Choice_i$ can be an individual's decision to take up ASA (if in group R), to take up Gihandom (if in group W), to default on an ASA contract, or to pre-order ASA for a second round. The vector X_i contains demographics (age, gender, marital status, income, assets, household size, years of education), as well as all survey-based measures mentioned in Section 3. In addition, all binary choice regressions contain marketer fixed effects. This is to filter any noise from differences in marketer ability.

5 Results

5.1 Average Treatment Effects

Effect on Total Bank Savings

This section presents estimates of the effects of the two commitment treatments on individuals' total savings held at the partner bank. The outcome variable of interest is the change in a client's total savings balance at the partner bank, summed across ordinary savings accounts and any commitment savings products (ASA or Gihandom). The savings period is specific to each individual, starting with the date of the baseline survey, and ending with the goal date specified in an individual's personal savings plan.[32] The cost of this reliance on the goal date is that it diminishes the sample to those 748 individuals who a) could be located for the marketing visit and b) were willing to make a savings plan with the marketer. This form of attrition is orthogonal to assignment to treatment.

Column (1) in Table II estimates that assignment to the Regular Saver treatment group increased average bank balances by 585 pesos (U.S.$14) relative to the control group. This estimate already includes any charged termination fees due to default. In contrast, individuals assigned to the Withdrawal-Restriction group saved on average 148 pesos more than the control group. Noting that the average duration of savings periods was 130 days (about 4.5 months), this estimate is roughly comparable to the effect estimated in Ashraf et al. (2006b): In a sample of previous savings account holders, the authors find that their withdrawal-restriction product SEED increased average savings by 411 pesos after 12 months. Given that the product design of SEED and Gihandom was identical, the Gihandom estimates presented here also serve to replicate and confirm the results of Ashraf et al. (2006b). Furthermore, the estimates confirm a small but significant increase of 27 pesos in savings for the control group. Two interpretations are possible: First, the marketing treatment could have led to higher savings even in the absence of commitment products. Second, the savings increase could be a result of the monetary rewards received in the baseline survey. Randomisation into treatment groups was stratified on whether individuals had received real rewards, which ensures that the resulting income shock is exogenous to

[32] All accounts except for those of existing 1st Valley Bank clients were opened after the marketing stage, implying the observed change in savings is equal to the savings balance. For those 18 clients who had previously existing 1st Valley Bank savings accounts, the existing account was monitored instead of opening a new ordinary savings account. Existing bank clients were still offered commitment savings products, in accordance with their assignment to treatment.

TABLE II: SAVINGS OUTCOMES (OLS, PROBIT)

	(1) Change in Bank Savings	(2) Purchased Savings Goal	(3) Borrowed to Purchase Goal (given purchase)	(4) Change in Other Savings (survey-based)
Regular Saver Treatment (ASA)	585.4652***	0.1156**	0.0509	426.8112
	(129.2510)	(0.0486)	(0.0621)	(671.8442)
Withdrawal Restr. Treatment (Gihandom)	148.2429***	0.1322**	0.2109***	-328.1585
	(40.9269)	(0.0545)	(0.0808)	(705.4607)
Constant	27.1600***			63.4513
	(9.3987)			(531.0279)
Mean Dep. Variable		0.4992	0.1922	
R^2	0.02			0.00
Observations	748	615	307	603

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Entries in columns (1) and (4) represent OLS coefficients. Entries in columns (2) and (3) represent marginal coefficients of the corresponding probit regressions.

treatment. Note that savings increases are net of the 100 peso opening balance contained in the free ordinary savings account – this amount constituted the minimum account maintaining balance, and no client closed their ordinary savings account during the period of observation.

In addition to the ITT effects reported in Table II, an instrumental variables regression of the change in bank savings on an indicator for take-up of the commitment products (using assignment to groups R and W as orthogonal instruments) provides an estimate of the Treatment-on-the-Treated (TOT) effect (discussed in Section 4). The TOT regression suggests that taking up the regular-instalment product ASA increased savings by 1928 pesos, while taking up the withdrawal-restriction product Gihandom increased savings by 324 pesos. Both estimates are conditional on the assumption that being offered a commitment product has no effect on savings, other than through use of the product (equivalently, those who rejected the commitment products on average saved the same as clients in the control group). This assumption is supported by the fact that the marketing treatment was identical across all groups. The increased gap in the TOT effects of ASA and Gihandom relative to their ITT effects is a result of the higher take-up rate for the Gihandom product.

The remainder of Table II presents the results from a probit estimation of whether individuals purchased the savings goal (see Table IX) they had been saving for: At the end of the endline survey, individuals were asked whether they had purchased the good specified on their personal savings plan.[33] Note that this is a distinct question from whether individuals achieved a certain amount of money in a bank account – they could have saved for the good at home, or found a different way to pay for it. If respondents confirmed having purchased the desired good, they were further asked how they paid for it, and in particular whether they borrowed (from any source, including friends or family). Due to attrition in the endline survey, the sample for this estimation is limited to the 615 individuals who a) had made a savings plan during the marketing stage and were b) reached by the endline survey.[34] Exactly half of the individuals reported to have bought the good, or paid for the expenditure, that was named

[33]The survey team was informed about this savings purpose, in case individuals had forgotten.

[34]Both 'having a savings plan' and 'being reached by the endline survey' are orthogonal to assignment to treatment.

on their personal savings plan. These 307 individuals constitute the sample for the probit regression in column (3), which estimates the effect of treatment on the likelihood of borrowing for the purchase (conditional on purchase). Borrowing was not uncommon: Slightly below 20 percent of individuals chose loans or family borrowing as a means of affording the expenditure.

Table II confirms that both the Regular Saver treatment and the Withdrawal Restriction treatment increased an individual's chances of purchasing their savings goal. The coefficients for the two treatments are not significantly different from each other. However, column (3) shows that individuals in the Withdrawal Restriction group were significantly more likely to borrow in order to obtain the good: Converting the probit coefficients into marginal effects, assignment to group W increases the likelihood of borrowing by 19.6 percentage points (from 11.4 to 31 percent). In comparison, assignment to group R (being offered ASA) increased the likelihood of obtaining one's savings goal, but did not significantly affect the probability of borrowing for it. This may suggest that the ASA product indeed helped individuals to purchase a savings goal using their own money, and without the use of loans.

Figure 6 (Appendix A) shows the impact of the Regular Saver treatment and the Withdrawal-Restriction treatment on the cumulative distribution of changes in bank savings, total savings, outstanding loans, and expenditures.[35]

Testing for Crowd-Out of Savings

A caveat about the estimation presented above is that it is restricted to savings at the partner bank. During the baseline survey, 46 percent of the sample reported to have an existing savings or checking account. This number is partly driven by microentrepreneurs, who are required to hold an existing savings account when obtaining microloans (the pairwise correlation is 0.18). More than one quarter of bank account holders reported not to have used their account in the last 12 months, and dormant accounts were common. The regression in column (4) of Table II seeks to establish whether the savings increases observed at the partner bank constituted new savings, or whether a simple substitution from other sources of savings (at home, or at other institutions) took place.

The outcome variable in column (4) is the change in an individual's total savings balance outside of the partner bank, as measured by survey data: During the baseline survey, individuals were asked about their savings at home, money lent out or safekept by others, informal savings, and savings at other institutions. An incentive of 30 pesos was paid for showing an existing bank passbook. The same exercise was repeated during the endline survey six months later, except that individuals were now questioned about the savings they kept around the time of their goal date. Unfortunately, the survey data is very noisy, and coefficients are estimated with substantial imprecision.[36] The available evidence does not suggest that a substitution took place between savings increases at the partner bank, and savings at home or at other institutions. All coefficients are insignificant. Moreover, the coefficient for being assigned to the Regular Saver treatment is positive – if anything, individuals who were offered ASA may have been encouraged to save even more in other savings vehicles, in addition to deposits made to their ASA accounts. In contrast, the coefficient for the Withdrawal Restriction treatment is negative. While this could easily represent survey noise, it is consistent with the 'safekeeping' explana-

[35]The observation period ends with the goal date for bank savings and total savings, and with the date of the endline survey for outstanding loans and expenditures.

[36]To account for some outliers in the stated balances, the savings data has been truncated at 1 percent, reducing the sample from 615 to 603 observations.

tion discussed earlier: Individuals may decide to shift existing assets into an account where they know other members of their household will not be able to access them.

5.2 Heterogeneity: Descriptive Results

The ASA results were very bi-modal: At the time of their goal date (between December 2012 and April 2013), 51 ASA clients (45 percent) had reached their savings goal. They had completed all scheduled instalments with a median of 12 transactions,[37] and reached savings goals between 950 and 7150 pesos (U.S.$170). By design, accounts were closed after completion of the savings plan, and clients could withdraw their savings in order to pay for the planned lump-sum expenditure (any remaining savings were transferred onto clients' ordinary savings account). Many of these clients pro-actively enquired at the bank to roll over their account into a new ASA contract. While rolling over contracts was not an immediate possibility during the study period, the repeat marketing stage included the option to 'pre-order' the product for a second round, should the bank decide to offer the product again. The pre-order contract was not financially binding, but included substantial official paperwork. Two thirds of the successful clients took up this offer (see Table IV), devised a new savings plan, and chose a new termination fee. The bank has since decided to offer new ASA contracts to those enquiring about them at the branch.

The situation looked very different for the remaining 63 ASA clients (55 percent) who defaulted on their savings contract. After falling three deposits behind, their accounts were closed, and the initially agreed termination fee charged (and transferred to charity). What happened? Two possibilities emerge:[38] (i) Clients had chosen an ASA contract which was optimal for them in *expectation*, and then rationally defaulted upon observing a shock (in other words, a 'bad luck' scenario). Or (ii), clients chose the contract by mistake. If the 'bad luck' explanation is true, the timing of the defaults should depend on the shock distribution: If shocks are independently distributed across individuals and time, and hazard rates are small, the timing of defaults should be roughly uniform over time. In sharp contrast, Figure 1 illustrates that clients had a tendency to default either right from the start, or not at all: Out of 63 defaults, 35 clients stopped depositing immediately after the opening balance, 10 clients made one more deposit, another 10 made between three and five deposits, and only 8 clients made more than five deposits (see Figure 1). Approximating transactions with weeks (85 percent of clients chose weekly instalments), Figure 1 also illustrates the expected default timing given a hazard rate of 0.028 per week. This estimate is obtained from the endline survey: The sample population was questioned about the occurrence of 17 types of common emergencies (sickness, loss of job, bad business, flood damage) including a flexible 'other' category. 45% reported at least one shock within 6 months, with an average of 0.72 shocks, equivalent to 0.028 shocks per week. This hazard rate is neither consistent with the overall frequency of defaults (observed 55 percent versus predicted 29 percent based on a 12 week contract), nor with the steep timing of defaults. A much higher hazard rate of 0.56 shocks per week would be needed to explain that 56 percent of all defaults happened immediately after opening. A hazard rate this high is problematic: It predicts an overall default frequency of 99 percent within 6 weeks, which contradicts both the observed 45 percent 'success rate' on contracts lasting 12-24 weeks, as well as the thick tail of the default distribution (13 percent of defaults occur more than 6 weeks after opening). The observed default timing is difficult to reconcile with the exponential pattern that would be generated

[37] One transaction can cover several weeks' instalments.

[38] Strictly speaking, this assumes that an individual would not take up a contract if he knows that default is certain.

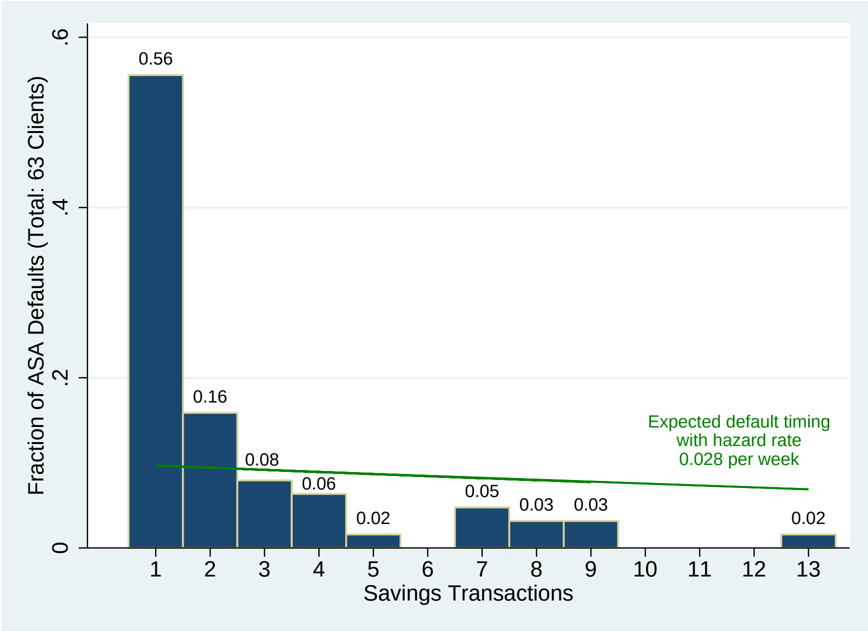


Figure 1: Savings Transactions: Defaulted ASA Clients

by any i.i.d. hazard rate. Unless there was an aggregate shock which affected all defaulting clients immediately after opening their accounts, a pure shock explanation seems unlikely. Evidence of aggregate shocks is discussed in Section 7.2. Heterogeneous hazard rates are discussed in Section 7.1.

The second possibility requires a deviation from rational expectations: Individuals could have chosen their contract by mistake. Mistakes (defined as choices that are not optimal under rational expectations) can happen if individuals have incorrect beliefs about their future preferences or their income distribution, including the probability of shocks to either of the two. Section 6 outlines why a time-inconsistent agent with incorrect beliefs about the degree of his time-inconsistency is likely to select into a commitment contract that is too "weak" to overcome his self-control issues, leading to default. Looking at the data, it is notable that 80 percent of individuals chose the minimum permissible termination fee for their savings goal (P150 for goals below P2500, and P250 for goals of P2500 and above), roughly equivalent to a day's wage. The observed combination of minimum penalties and high default rates raises the question whether individuals underestimated the amount of commitment it would take to make them save. This is consistent with the observed tendency to default soon after account opening, as individuals start behaving according to their true degree of time-inconsistency upon entering the depositing phase. Could rational expectations about stochastic future time-inconsistency explain the data? If individuals had correct beliefs on average about their future preferences, they should realize which contract (and in particular, which penalty) will be effective for them on average. Moreover, risk-averse preferences imply that individuals who internalize the risk of default should either sign up for stronger commitments (to be on the safe side), or stay away from commitment. This is inconsistent with the frequency of observed defaults, and the tendency to choose the minimum penalty.

Figure 7 (Appendix A) lists the chosen termination fees of the 114 ASA clients, and contrasts them with how much was charged ('successful' indicates that no fee was charged). Not all chosen fees were enforceable: Whenever clients chose a fee strictly above the minimum and later defaulted on their contract, the charged fee was the lesser of chosen fee and savings balance at the time of default. The minimum fee was always enforceable through the opening balance.

Unfortunately, it is safe to conclude that the ASA contract likely reduced the welfare of a significant share of its adopters. For the 35 clients who defaulted immediately after depositing the opening balance, losing the opening balance (through the termination fee) was the only economic consequence of the contract, thus leaving them worse off. For those who defaulted later during their savings plan (thus making a shock explanation more likely), an argument can be made that the contract helped them to achieve savings which they would not otherwise have been able to achieve, at a negative return of 150 or 250 pesos (which is still less than common interest payments to local moneylenders). A cautious estimate of the frequency of 'mistakes' is provided by the pre-order results: 55 percent of all ASA clients (71 percent of defaulting clients and 35 percent of successful clients) chose not to order the product again (see Table IV).

TABLE III: ACCOUNT USAGE

Average # of Deposits (includes opening balance)	Mean	Median	#accounts
ASA (all)	6.76	5	114
______(successful)	11.98	12	51
______(default)	2.52	1	63
Gihandom Account (all)	1.68	1	92
______(date-based)	1.68	1	53
______(amount-based)	1.69	1	39
Control Savings Account	0.43	0	788

TABLE IV: ASA PRE-ORDER

	Yes	No	
Successful	33	18	51
Default	18	43	63
	51	63	114

For the Gihandom accounts, both benefits and risks were less pronounced: Out of 92 accounts, only five reached the goal amount specified in their savings plan (three were date-based, two were amount-based). The 53 clients who had opted for a binding date goal received their savings back after the savings period completed. Their median savings were 100 pesos (average 286 pesos), which is equivalent to the minimum opening balance. Out of 39 clients who had opted for binding goal amounts, 35 were still open at the end of the six-month observation period (average savings 141 pesos).[39] 85 percent of all amount-based Gihandom accounts (and 79 percent of Gihandom accounts overall) had no further deposits after the opening balance. This creates a parallel between Gihandom and ASA defaults: Similarly to the ASA clients who made no further deposits, amount-based Gihandom accounts effectively lose their opening balance if they do not continue to deposit. A difference between the two commitment products is that the penalty for discontinuing to save on an amount-based Gihandom account increases with every deposit, while the default penalty for ASA is fixed.

Finally, 582 clients opened exclusively an ordinary savings account – either because they were assigned to the control group, or because they rejected the commitment product offered to them. Out of these clients, one reached the goal amount specified within their savings plan. Summary statistics for transactions in all accounts can be found in Table III.

[39]Two accounts were closed after reaching the goal amount, and another two were closed after the bank mistakenly treated them as date-based accounts.

5.3 Heterogeneity: Regressions

In an attempt to resolve the puzzles presented in the previous section, this section analyses empirical predictors of take-up for the two commitment products, as well as default and pre-order decisions.

5.3.1 Predicting Take-Up of the Commitment Savings Products

Columns (1) to (3) in Table V present the results of a probit regression of the ASA take-up decision on a number of potential determinants, limiting the sample to the Regular Saver group (R), where ASA was offered.[40] The first notable fact is that not a single demographic factor seems to correlate with the take-up decision. Age, gender, income, assets, marital status, education and household size all appear to be insubstantial for the decision to take up the regular-instalment product.

The main factors which do predict ASA take-up are the proposed measure of sophisticated hyperbolic discounting (see Section 3) and a measure of cognitive ability (see Figure 9 for a sample question from the cognitive ability test). The positive predictive power of cognitive ability is reassuring: The ASA product is more complex in its rules than traditional savings accounts (but no more complex than a loan contract). The significance of cognitive skills suggests that those clients who were more likely to understand the rules were also more likely to take up the product. This may also be interpreted as evidence against possible manipulation by the bank marketers.

Present bias on its own is not a predictor of take-up, consistent with the intuition that *perceived* time-inconsistency, rather than *actual* time-inconsistency, determines demand for a commitment product. Perhaps more surprisingly, the association of commitment take-up and sophisticated hyperbolic discounting is significant and *negative*. Recall from Section 3 that sophistication is measured as the interaction of present bias (from multiple price list questions) and self-reported temptation. In other words, those who exhibit hyperbolic preference reversals, but at the same time report *low* levels of temptation, are more likely to take up the product. In contrast, those who report being strongly tempted tend to stay away from the product. To interpret interaction coefficients, note that present bias is a binary variable, whereas temptation is in the interval $[0, 10]$. A possible explanation for this link comes from theory: Partially sophisticated agents (i.e. those with time-inconsistent preferences and positive but low self-perceived temptation) have a positive demand for commitment. They take up the product and choose a low default penalty, which they anticipate will be sufficient to make them save. In contrast, agents who perceive themselves as strongly tempted have two choices: Either they take up the product with a penalty that is sufficiently large to make them save, or they stay away from the product completely. The latter choice may be optimal if the required effective penalty is very high: Given a constant probability of 'rational default', in which a shock (say, the loss of one's business) makes it optimal for an individual to discontinue their contract, agents with a higher default penalty have more to lose. As a result, for a fully sophisticated agent with medium to high time-inconsistency, a low penalty may not be effective, and a high penalty may be too risky in the face of uncertainty.

Column (3) looks beyond 'deep' individual characteristics and investigates correlations with other choices. While neither distance to the bank branch nor estimated shock frequency significantly affect take-up probability, individuals with an existing bank account (at any local bank) were more likely

[40] The sample for the regression is restricted to those clients who could be located for the marketing visit. Inability to locate individuals for marketing is orthogonal to treatment group assignment.

TABLE V: PREDICTING DEMAND FOR COMMITMENT (PROBIT)

Commitment Take-Up	ASA (1)	ASA (2)	ASA (3)	Gihandom (1)	Gihandom (2)	Gihandom (3)
Age	-0.0015	-0.0004	-0.0002	0.0017	0.0009	0.0011
	(0.0019)	(0.0020)	(0.0020)	(0.0029)	(0.0028)	(0.0028)
Female	0.0328	0.0592	0.0504	0.2418	0.2347	0.2299
	(0.0914)	(0.0936)	(0.0875)	(0.1687)	(0.1501)	(0.1520)
Married	0.0076	0.0095	0.0191	-0.0932	-0.0848	-0.0942
	(0.0650)	(0.0640)	(0.0631)	(0.0939)	(0.0872)	(0.0890)
HH Income	0.0018	0.0013	-0.0030	0.0003	-0.0054	-0.0054
	(0.0081)	(0.0080)	(0.0079)	(0.0118)	(0.0108)	(0.0109)
Assets	0.0007	-0.0048	-0.0083	0.0220	0.0253	0.0255
	(0.0145)	(0.0143)	(0.0143)	(0.0224)	(0.0213)	(0.0216)
HH Members	0.0125	0.0105	0.0130	0.0229	0.0202	0.0216
	(0.0100)	(0.0097)	(0.0095)	(0.0157)	(0.0154)	(0.0158)
Education (yrs)	-0.0055	-0.0075	-0.0094	0.0276***	0.0316***	0.0316***
	(0.0066)	(0.0064)	(0.0064)	(0.0094)	(0.0093)	(0.0093)
Present Bias	0.0757	0.0636	0.0805	0.0809	0.1020	0.1057
	(0.0866)	(0.0870)	(0.0864)	(0.1260)	(0.1256)	(0.1301)
Soph. Present Bias	-0.0622**	-0.0579**	-0.0620**	-0.0363	-0.0524	-0.0537
(Pres.Bias*Temptation)	(0.0291)	(0.0290)	(0.0294)	(0.0491)	(0.0510)	(0.0534)
Perceived Temptation	-0.0114	-0.0067	-0.0050	-0.0058	0.0005	0.0010
	(0.0123)	(0.0124)	(0.0124)	(0.0212)	(0.0205)	(0.0207)
Impatience	-0.0047	0.0053	0.0008	-0.0124	-0.0224	-0.0228
	(0.0476)	(0.0467)	(0.0464)	(0.0717)	(0.0684)	(0.0687)
Financial Claims	-0.0022	-0.0076	-0.0029	0.1079	0.1166*	0.1166*
	(0.0426)	(0.0418)	(0.0414)	(0.0663)	(0.0638)	(0.0639)
Risk Aversion		-0.0049	-0.0057		0.0497***	0.0499***
		(0.0106)	(0.0105)		(0.0167)	(0.0168)
Cognitive Ability		0.0353*	0.0364*		0.0157	0.0170
		(0.0188)	(0.0187)		(0.0236)	(0.0238)
Financial Literacy		0.0425*	0.0331		-0.0115	-0.0119
		(0.0246)	(0.0250)		(0.0322)	(0.0322)
HH Bargaining Power		0.0063	0.0054		0.0444***	0.0458***
		(0.0113)	(0.0113)		(0.0164)	(0.0167)
Distance to Bank			-0.0271			0.0078
			(0.0207)			(0.0262)
Exist. Savings Account			0.1009**			-0.0023
			(0.0444)			(0.0668)
#Emergencies last yr			-0.0168			-0.0217
			(0.0277)			(0.0493)
Marketer FE	YES	YES	YES	YES	YES	YES
Mean Dep. Variable	0.2687	0.2687	0.2687	0.4115	0.4115	0.4115
Observations	402	402	402	209	209	209

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Entries in the table represent the marginal coefficients of the probit regressions.

to take up the product. Given a widespread scepticism towards banks in the study area, this may be interpreted as a sign of trust in and familiarity with the banking system.

Columns (4) to (6) present the same regressions applied to take-up for the withdrawal-restriction product Gihandom, limiting the sample to group W (where Gihandom was offered). Most strikingly, there is no overlap in the factors predicting ASA and Gihandom. If the products were perceived as close substitutes, and individuals in need of commitment merely took up whichever commitment product was offered to them, then the empirical analysis should find that the same factors which predict ASA take-up also predict take-up of Gihandom. A look at the data confirms that the sets of determinants for the two products are mutually exclusive, suggesting that individuals perceived ASA and Gihandom rather differently. Specifically, Gihandom take-up is predicted by high education (measured in years of schooling), high risk aversion (choosing a 'safe' lottery in Figure 8), high household bargaining power (measured using questions on who decides what in a household), and strong claims from others on own liquid assets. Considering a 94 percent female sample population, this combination of factors is reminiscent of the evidence presented in Anderson and Baland (2002): In their study, the authors argue that Kenyan women use commitment devices (ROSCAs) to protect their savings against claims from their husbands. They propose an inverted U-shaped relationship between women's power in their household, and participation in ROSCAs. While the Kenyan context studied by Anderson and Baland (2002) is different from the Philippine context studied here, the evidence in Table V is consistent with the explanation that women took up Gihandom to 'safeguard' their savings from intra-household conflicts. The withdrawal restriction featured in the Gihandom account is well-suited to preventing other household members from accessing savings, but allows the woman to retain flexibility regarding when to make deposits. The estimated linear relationship of commitment take-up with household bargaining power is unable to capture the proposed inverted U-shape. However, both household bargaining power and female education may be associated with an increased autonomy of the woman in planning to build up savings of her own. Finally, the strong predictive power of risk aversion is consistent with a precautionary savings motive: Those women who are particularly concerned about consumption variance and the possibility of shocks will be more interested in putting savings aside for future hard times.

No evidence currently suggests that demand for the withdrawal-restriction product Gihandom is associated with intra-personal conflicts and time-inconsistency. A reservation must be made with respect to statistical power: The sample of group W is half the size of group R, reducing the precision of estimates. Summing up, the evidence currently available suggests that demand for ASA is related to time-inconsistency and partial sophistication, while demand for Gihandom appears to be related to household bargaining and safekeeping motives.

5.3.2 Predicting Default and Repeat Take-Up

Table VI presents marginal coefficients from probit regressions with ASA default as well as the ASA pre-order decision as the dependent variable. A take-up regression (column (3) from Table V) has been added for comparison. In addition to the regressors from the take-up regressions, Table VI also includes the number of emergencies (illness or death of household members, unemployment, damage due to natural disasters, and a range of other income and expenditure shocks) which the household suffered since the baseline survey.

TABLE VI: ASA DEFAULTS & REPEAT TAKE-UPS (PROBIT)

Dependent Variable	ASA Take-Up	Default (R-Sample)	Default (takeup-Sample)	Pre-Order (takeup-Sample)
Age	-0.0002	-0.0022	-0.0049	-0.0070
	(0.0020)	(0.0016)	(0.0040)	(0.0045)
Female	0.0504	0.1100	0.3332*	0.0129
	(0.0875)	(0.0923)	(0.1758)	(0.2002)
Married	0.0191	0.0122	0.0541	-0.1862
	(0.0631)	(0.0540)	(0.1270)	(0.1601)
HH Income	-0.0030	0.0031	0.0205	0.0149
	(0.0079)	(0.0057)	(0.0161)	(0.0179)
Assets	-0.0083	-0.0090	-0.0383	-0.0026
	(0.0143)	(0.0115)	(0.0236)	(0.0283)
HH Members	0.0130	0.0119	0.0124	-0.0052
	(0.0095)	(0.0077)	(0.0138)	(0.0181)
Education (yrs)	-0.0094	-0.0029	0.0000	-0.0171
	(0.0064)	(0.0053)	(0.0120)	(0.0143)
Present Bias	0.0805	0.1084*	0.4424*	-0.5468**
	(0.0864)	(0.0651)	(0.2532)	(0.2398)
Soph. Present Bias	-0.0620**	-0.0429*	-0.1498	0.2882**
(Pres.Bias*Temptation)	(0.0294)	(0.0238)	(0.1307)	(0.1186)
Perceived Temptation	-0.0050	-0.0212**	-0.0698***	-0.0037
	(0.0124)	(0.0105)	(0.0237)	(0.0298)
Impatience	0.0008	0.0004	0.0262	0.0451
	(0.0464)	(0.0373)	(0.0893)	(0.1010)
Financial Claims	-0.0029	-0.0092	0.0134	0.0102
	(0.0414)	(0.0331)	(0.0852)	(0.0926)
Risk Aversion	-0.0057	-0.0180**	-0.0676***	0.0139
	(0.0105)	(0.0085)	(0.0193)	(0.0251)
Cognitive Ability	0.0364*	0.0362**	0.0705*	-0.0226
	(0.0187)	(0.0143)	(0.0381)	(0.0427)
Financial Literacy	0.0331	-0.0166	-0.1440***	0.0472
	(0.0250)	(0.0203)	(0.0403)	(0.0455)
HH Bargaining Power	0.0054	-0.0111	-0.0746***	0.0803***
	(0.0113)	(0.0090)	(0.0233)	(0.0271)
Distance to Bank	-0.0271	-0.0130	-0.0371	0.0632
	(0.0207)	(0.0163)	(0.0541)	(0.0619)
Exist. Savings Account	0.1009**	0.0329	-0.0658	0.1776*
	(0.0444)	(0.0363)	(0.0865)	(0.0935)
#Emergencies last yr	-0.0168	-0.0004	0.0559	-0.0778
	(0.0277)	(0.0211)	(0.0612)	(0.0689)
#Emergencies		-0.0000	0.1257*	-0.0047
since baseline		(0.0182)	(0.0673)	(0.0660)
Marketer FE	YES	YES	YES	YES
Mean Dep. Variable	0.2687	0.1468	0.5463	0.4630
Observations	402	402	108	108

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Entries in the table represent the marginal coefficients of the probit regressions.

Column (2) in Table VI can be understood as an analysis of which type of individuals took up the commitment product 'by mistake', proxied by take-up and subsequent default (this interpretation abstracts from the possibility of rational default). The results provide further support to the partial sophistication hypothesis: Among those randomly assigned to group R, present-biased individuals are significantly more likely to take up the ASA product and then default. This effect is particularly strong for agents who report low levels of temptation, representing naive and partially sophisticated hyperbolics. In contrast, more sophisticated hyperbolics are *less* likely to default: Note that temptation is in [0, 10] with a median of 2. Aggregating the coefficients for present bias (0.11*), sophistication (-0.04*) and temptation (-0.02**) yields a lower likelihood of default for all time-inconsistent agents with perceived temptation values higher than the median. This in line with the proposed explanation: Sophisticated hyperbolic discounters either don't select into the product (if the effective penalty would be prohibitively high), or they make sure to choose a contract which is incentive-compatible for their preferences (which could be through the size of the weekly instalment, or the size of the penalty). The data confirm that ASA clients with higher perceived temptation are indeed more likely to choose a penalty strictly above the minimum. However, due to lack of variation in penalties, this relationship is not significant.

Column (3) analyses default occurence in the take-up sample, and should be interpreted as correlational evidence only (since the regression conditions on an endogenous variable). The marginal coefficient on present bias has quadrupled, and kept its significance. Interestingly, the link between present bias (as proxied by observed time-inconsistency) and default seems to be much stronger than the link between present bias and take-up. This is consistent with the theoretical intuition that an agent's take-up decision should be driven by *perceived* time-inconsistency, as proxied by the sophistication measure. In contrast, once the agent has adopted the contract, actual time-inconsistency will determine the success of the contract (in addition to a sophistication effect). The temptation measure now has strong predictive power on its own, even when not interacted with present bias. Expanding on the discussion from Section 3, individuals who report being positively tempted but do not exhibit hyperbolic preference reversals in MPLs could be either one of two things: a) they are time-inconsistent, but incorrectly classified as time-consistent in MPL questions, or b) they are subjectively feeling tempted but behaving time-consistently, possibly due to the exercise of costly self-control. In both cases, higher awareness of temptation will prompt individuals to choose more manageable (incentive-compatible) contracts – either through higher penalties or through lower weekly deposits (conditional on income).

Moving on to the pre-order (repeat take-up) decision, the effects of present bias (-0.55**) and sophistication (0.29**) are now large and significant. The aggregate coefficient for a present-biased individual with the median value of perceived temptation is approximately zero. This has a convenient interpretation: Relatively naive hyperbolic discounters (those with below-median reported temptation) are unlikely to take up the ASA product again. From the previous analysis, there is a high chance that these individuals defaulted on their contract, and at the same time had not anticipated the default risk. These clients have 'burnt their fingers'. The result that such individuals do not take up the product again is encouraging, insofar as it suggests learning about their preferences. The reverse holds true for present-biased individuals with above-median reported temptation (sophisticated hyperbolics): The aggregate coefficient on their time preferences is positive, suggesting a higher likelihood to pre-order ASA for a second round. This is consistent with the conjecture that sophisticated hyperbolic discounters choose the 'right' contract, which is incentive-compatible with their true preferences, and optimal in expectation. This does not imply a one-to-one mapping from successful ASA completion to the decision to pre-order: A sophisticated client who chose a contract that was optimal in expectation, but then rationally defaulted following a shock, might well decide to take up the product again (supported by an imperfect mapping from account status to pre-order decision in the data, see Table IV).

A number of other factors can help in explaining the observed default rates. The most obvious candidate - the occurence of shocks during the savings period - finds some support in the take-up sample ('emergencies since baseline', column (3)). Shock occurence was estimated by asking for common income or consumption emergencies during the endline survey.[41] The positive correlation of defaults with shocks, in combination with the fact that 45 percent of clients completed their ASA contract successfully, suggests that a significant portion of the take-up sample likely did choose a contract which was optimal for them in expectation. Clients without shocks could complete their plan successfully, while those with shocks rationally defaulted. The theoretical prediction that shock realisation should be irrelevant to the pre-order decision (as it does not affect contract optimality in expectation) is supported by the data (see column (4) of Table VI). Other factors predicting default include financial literacy (-), household bargaining power (-), risk aversion (-) and cognitive ability (+). Financial literacy is perhaps the least surprising: Individuals with poor numeracy skills tend to do worse at managing their household finances, and may fail to allocate a portion of the household budget to regular ASA deposits. The positive significance of cognitive ability in the take-up sample is likely due to a multi-collinearity with financial literacy (the pairwise correlation is 0.33). It disappears when financial literacy is excluded from the regression. In contrast, the positive significance of cognitive ability within the R-sample is robust. This can be partially explained by the predictive power that cognitive ability has for take-up of the ASA product. A similar puzzle arises from the negative correlation of risk aversion with default (but not with take-up). An explanation requires a closer look at how the measure was obtained: The risk aversion measure is a score in [1, 6], indicating which lottery individuals chose from a set of lottery options with increasing expected value and increasing variance (see Figure 8). If preferences are characterised by reference dependence (with the no-risk lottery A as a reference point) and loss aversion, then the choice of a safe lottery would measure loss aversion rather than risk aversion. A high degree of loss aversion can be associated with a lower likelihood to default on the ASA product (to avoid loss of the penalty). Finally, it is interesting to note that household bargaining power is unrelated to take-up of the regular-instalment product ASA, but strongly related to defaulting on it (within the take-up sample). A possible explanation is that intra-household conflicts played no role in individual's motives to take up the product – but that, much like a shock, individuals soon learnt that it caused household conflicts to try and put aside a portion of the household budget every week, beyond the reach of other household members. This can be interpreted as a learning process in adopting a new savings technology (loan repayment is similar in structure, but may be easier to agree on in a household because of the higher penalties involved). Consequently, clients with low bargaining power may have yielded to these disagreements, and defaulted on their contracts. The large positive association of household bargaining power with the pre-order decision provides further support for a learning explanation: Once individuals had learnt about the difficulties of regularly diverting a share of the household budget, only those with sufficient autonomy in their household chose to take up the product again.

[41] There is a risk that clients who defaulted had a stronger incentive to report shocks, in order to preserve their self-image or reputation. However, the endline survey was framed as coming from a research organisation, with no direct link to the bank. The survey was identical across the sample, and made no reference to ASA or Gihandom. Note that attrition in the endline survey was compensated by imputing the median shock value for those who did not participate.

5.3.3 Heterogeneous Treatment Effects

Table VII examines treatment effect heterogeneity across a number of dimensions of interest. The regression set-up is identical to that in column (1) of Table II: The change in savings held at the partner bank is regressed on indicators for assignment to the treatment groups. In addition, the treatment indicator for the Regular Saver group is interacted with variables which have been shown to predict take-up or default, or which are of interest in themselves. Interaction variables include present-biased preferences, the self-reported sophistication measure, having an existing savings account, household bargaining power, and household income.

Heterogeneity in treatment effects is most pronounced for existing savings account holders. Existing savings account holders increased their savings balances by 622 pesos more than those without an existing account after being offered the Regular Saver product. Put differently, the intent-to-treat effect of the Regular Saver product was 909 pesos for existing savings account holders, and only 287 pesos for those without existing accounts. This seems particularly surprising in light of the fact that, in *absence* of the Regular Saver treatment, existing account holders saved only 75 pesos more than those without existing accounts. The evidence suggests that existing account holders were not necessarily active savers before the intervention, but felt strongly motivated by the Regular Saver treatment. A possible explanation relates to mistrust and negative preconceptions towards banks, which were common in the population.[42] Existing account holders were more likely to be familiar with basic bank transactions, and more trusting of the banking system as a whole.

It is worth noting that treatment effects appear to be relatively uniform across measures of present bias and sophistication. Theory predicts that a present-biased agent with a low degree of sophistication is likely to select into a commitment contract that is too weak to be effective given his preferences, resulting in default soon after take-up (see Section 6). After taking into account the default penalty, savings with the commitment product should be weakly smaller than savings without the commitment product. The positive association between (naive) present bias and default is supported empirically by the regressions in Table VI. The negative effect of present-biased preferences on savings should be mitigated or even reversed with increasing levels of sophistiction: The agent is more likely to choose an incentive-compatible contract, increasing the chances of successfully reaching his savings goal. The sign of the aggregate coefficient on sophisticated present-biased preferences relative to time-*consistent* behaviour is theoretically ambiguous, as illustrated by O'Donoghue and Rabin (1999). Column (1) of Table VII shows that all estimates of treatment effects with respect to measures of present bias and sophistication are small and insignificant. A likely reason are the composition effects inherent in ITT estimates: Individuals with sophisticated time-inconsistent preferences were much less likely to select into the product (see Table V). Thus, a lower percentage of sophisticated agents were 'treated'. Theoretical arguments in Section 6 confirm that a sophisticated agent may choose to stay away from commitment, if the effective default penalty is prohibitely high in the presence of shocks. If his preferences are such that he cannot achieve his savings goal in autarky, he will choose not to save.

Columns (3) and (4) show that the estimated treatment effect is relatively uniform across household income level, as well as an above-median indicator for cognitive ability. Column (6) suggests that successfully maintaining the Regular Saver product ASA was facilitated by having a certain degree

[42] It was a common belief that banks were "not for poor people". In addition, some individuals believed that savings deposited at a bank would likely be lost if the bank became insolvent. Deposit insurance does exist in the Philippines, but may be associated with years of waiting time. See e.g., Dupas et al. (2012) on trust-related challenges in banking the poor.

TABLE VII: HETEROGENEOUS TREATMENT EFFECTS: CHANGE IN BANK SAVINGS

	(1)	(2)	(3)	(4)	(5)
Regular Saver (R)	707.1748***	481.4925***	287.4050***	455.4759***	351.9444***
	(257.1454)	(134.9550)	(63.6388)	(133.2911)	(114.1216)
Withdrawal Rest. (W)	129.4941***	146.3426***	147.2989***	154.5030***	148.8766***
	(40.6258)	(39.6980)	(40.7323)	(42.6279)	(41.4302)
R * Present bias	-58.8396				
	(543.3738)				
Present bias	57.6709				
	(83.9149)				
R*Soph. Present Bias	20.7603				
(R*Pres.Bias*Temptation)	(81.9725)				
Soph. Present Bias	-6.0420				
	(18.1875)				
R * Temptation	-48.8363				
	(62.2914)				
Temptation	-4.9396				
	(9.2861)				
R*High Cognitive Ability		261.2952			
		(291.0735)			
High Cognitive Ability		-30.7091			
		(36.4527)			
R * Existing SA			621.5350**		
			(269.7281)		
Existing SA			74.9955*		
			(41.3577)		
R * Income				44.6249	
				(56.0615)	
Income				7.6109	
				(5.4829)	
R * HH power					88.1604*
					(50.5904)
HH power					6.9664
					(10.3868)
Constant	34.9483	41.1479**	-7.0002	2.7796	8.8140
	(26.7103)	(17.9218)	(18.9202)	(17.7092)	(29.9048)
R^2	0.02	0.02	0.04	0.02	0.03
Observations	720	748	748	745	748

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

of household bargaining power: Individuals who report to be the primary decisionmaker in many aspects of household budgeting respond to the Regular Saver treatment with larger savings increases than those with low bargaining power. Using a score [0, 5], each one-point increase in bargaining power corresponds to an increase of 88 pesos in savings after being offered the Regular Saver product. This effect is consistent with the incidence of household conflicts caused by the weekly ASA instalments (see Section 5.3.2). Note that individuals with high bargaining power did not save more absent treatment – it is the interaction of sufficient bargaining power and the Regular Saver treatment which helped individuals to save. This explanation differs markedly from a 'safekeeping' motive: If individuals took up ASA to mitigate household bargaining issues, we would expect the interaction coefficient to be negative (as those with low power would benefit more from treatment).

6 Theory: Commitment under Partial Sophistication

The following section develops a formal understanding of the interaction between commitment and partial sophistication. Focusing specifically on a regular-instalment savings product, it sheds light on (i) why sophisticated hyperbolic discounters can benefit from commitment to fixed instalments, (ii) why commitment reduces welfare if it is too weak to be effective, (iii) why partially sophisticated hyperbolic discounters are likely to select into such weak commitment contracts, and (iv) why those with high perceived degrees of time-inconsistency may avoid commitment. The model extends the autarky savings framework in Basu (2012) to allow for partial sophistication, stochastic income (creating a need for flexibility), and a regular-instalment commitment savings product. The commitment design differs from previous models of commitment products by a default penalty that is conditional on per-period contributions, and the simultaneous absence of any withdrawal restrictions. A version of the regular-instalment design with full sophistication and deterministic income is discussed in Hofmann (2013).

Consider an agent who chooses whether to save for a nondivisible good which costs the lump-sum $2 < p < 3$ and yields a benefit $b > 3$. The agent lives for 3 periods and receives a per period income of 1 (barring shocks), which he can either consume or save. He cannot borrow. His instantaneous utility is twice differentiable and strictly concave, with $u'(c) > 0$, $u''(c) < 0$, and $u'(0) = \infty$. Throughout, assume the interest rate is $R = 1$ and $\delta = 1$ for simplicity. Define s_t as the amount of savings that he sends from period t to $t+1$, so that $c_t = y_t + s_{t-1} - s_t \geq 0$. Lifetime utility as evaluated in each period is given by the discounted sum of the instantaneous utilities:

$$U_t = u(c_t) + \beta \sum_{k=t+1}^{3} u(c_k)$$

For $\beta < 1$, the agent is *present-biased*: He exhibits a lower rate of discount over current trade-offs (t vs. $t+1$) than over future trade-offs ($t+s$ vs. $t+s+1$, $s > 0$). Following O'Donoghue and Rabin (1999), the agent's degree of sophistication about his present bias is captured in the parameter $\tilde{\beta} \in [\beta, 1]$, which he believes he will use in all future periods. In particular, the agent believes in period t that her utility function in period $t+s$ will be

$$U_{t+s} = u(c_{t+s}) + \tilde{\beta} \sum_{k=t+s+1}^{3} u(c_k).$$

For a fully sophisticated agent, $\tilde{\beta} = \beta$. A fully naive agent believes he will behave time-consistently in the future, captured in $\tilde{\beta} = 1$.

A need for flexibility is introduced through stochastic income shocks: With a per-period probability of λ, the agent loses his income in that period. This shock has a variety of interpretations: It can be interpreted directly as a loss of income, e.g., from redundancy, bad business, or illness of an income-earning household member. With a minor modification, it can be interpreted as a reduced-form taste shock: Suppose the sudden illness of a family member changes preferences such that utility stays unchanged if a hospital visit (at cost 1) is consumed and paid for, and drops to $u(c) = -\infty$ without a hospital visit. The implication of a shock is that the agent's lifetime income drops to (at most) 2, which means the nondivisible good can no longer be purchased. When a shock hits, any plans to save for the nondivisible are abandoned, and existing savings are optimally spread over the remaining periods for consumption. This results in a third interpretation: More generally, the shock λ corresponds to the probability that, for any time-consistent reason, the agent no longer finds it optimal to save for the nondivisible.[43]

While there is much ambiguity over the definition of welfare for time-inconsistent agents, the paper will follow the convention proposed by O'Donoghue and Rabin (1999): An agent's welfare is evaluated from an ex-ante perspective, and corresponds to the lifetime utility of the period 0 agent:

$$W = E[u(c_1) + u(c_2) + u(c_3)].$$

The advantage of this convention is that no particular period is favoured (since no consumption takes place in period 0).

6.1 First Best

I assume throughout that b, p are such that it is optimal for a time-consistent agent to save for the nondivisible. For $\lambda = 0$, consumption smoothing implies that the agent optimally distributes the required savings burden of $p - 1$ evenly over periods 1 and 2, and uses his period 3 income plus accumulated savings to purchase the good. The implied savings profile is $s_1 = \frac{p-1}{2} \equiv \bar{s}$, $s_2 = p - 1 = 2\bar{s}$. For $\lambda > 0$, there is a precautionary savings motive, even if the agent does not intend to save for the nondivisible. Denote such precautionary savings s_t^{No}. It can be shown that the optimal savings path is slightly increasing, i.e., $s_1 < \bar{s}$.[44] Since the present analysis focuses on regular-instalment products, I assume that desirability of the nondivisible still holds for fixed equal instalments $\bar{s}$:

$$\begin{aligned}
&(1-\lambda)^2[2u(\frac{3-p}{2}) + (1-\lambda)u(b) + \lambda u(p-1)] \\
&+ (1-\lambda)\lambda[u(\frac{3-p}{2}) + u(\frac{p-1}{2} - s_2^{No}) + E(u(y_3 + s_2^{No}))] \\
&+ \lambda[u(0) + E(u(y_2 - s_2^{No}) + u(y_3 + s_2^{No}))] \\
&\geq E[u(y_1 - s_1^{No}) + u(y_2 + s_1^{No} - s_2^{No}) + u(y_3 + s_2^{No})]
\end{aligned} \tag{2}$$

[43] Another time-consistent explanation why an agent may no longer wish to purchase the nondivisible are state-dependent preferences. In contrast to income shocks, this would not necessarily result in a precautionary savings motive.

[44] The probability of remaining shock-free (and thus obtaining the nondivisible) increases over time, from $(1-\lambda)^3$ ex-ante to $(1-\lambda)$ once period 2 has been reached without a shock. This makes it optimal to slightly skew the savings burden $p-1$ towards period 2. To see this formally, note that expected utility decreases in s_1 when evaluated at $s_1 = \bar{s}$: $dU/ds_1 = (1-\lambda)^2[-u'(1-s_1) + u'(2+s_1-p)] + (1-\lambda)\lambda[-u'(1-s_1) + u'(s_1 - s_2^{No})] < 0$ for $s_1 = \frac{p-1}{2} > 0.5$. By the envelope condition, $dU/ds_1 = \frac{\delta U}{\delta s_1} + \frac{\delta U}{\delta s_2} \cdot \frac{\delta s_2}{\delta s_1} = \frac{\delta U}{\delta s_1}$.

where s_t^{No} is chosen to optimally spread available current assets over the remaining future periods, conditional on not buying the nondivisible (i.e. $s_t^{No} = f(y_t, s_{t-1}, \lambda)$).

6.2 Autarky

The following analysis assumes that no shock has hit up to period t. If a shock does hit (i.e., if $y_t = 0$), the agent immediately gives up any plans to save for the nondivisible, and instead spreads available savings s_{t-1} optimally over the remaining periods. Denote such savings as s_t^{No}. For $\beta = \tilde{\beta} = 1$, the agent will always buy the nondivisible given the above condition (and absent shocks). The savings path will be perfectly smooth ($s_1 = \bar{s}, s_2 = 2\bar{s}$) if $\lambda = 0$, and slightly increasing ($s_1 < \bar{s}, s_2 = 2\bar{s}$) if $\lambda > 0$. If $\beta \leq \tilde{\beta} \leq 1$ (with at least one inequality strict), the three period selves engage in strategic interaction. Savings behaviour can be analysed by backward induction, taking into account the agent's belief about his future preferences.

Period 3

The agent will buy the nondivisible whenever he can afford it, i.e., whenever there is no shock, and $s_2 \geq p - 1$. Additional savings $s_2 > p - 1$ are simply consumed, as are savings that are not sufficient to buy the good. Since there are no future choices, the sophistication level does not influence behaviour at this stage. The consumption profile is

$$c_3 = \begin{cases} y_3 + s_2 - p + b & if\ y_3 + s_2 \geq p \\ y_3 + s_2 & if\ y_3 + s_2 < p \end{cases}$$

Period 2

The period 2 self knows the good will be bought if and only if he sends $s_2 \geq p - 1$, and absent shocks. He decides whether to send $s_2 = p - 1$, in which case the good is bought, or less. Due to consumption smoothing motives, it is never optimal to send $s_2 > p - 1 > 1$, which exceeds the magnitude of the shock. If the agent prefers not to save for the good, he will want to smooth s_1 over periods 2 and 3: $s_2^{No}(s_1) = argmax(u(y_2 + s_1 - s_2) + \beta E[u(y_3 + s_2)]$ subject to $0 \leq s_2 < p - 1$. This equation also describes his savings behaviour in case of a shock, where $y_2 = 0$, imposing an additional restriction $0 \leq s_2 < s_1$. He is willing to save $s_2 = p - 1$ if s_1 is such that

$$u(1 + s_1 - (p-1)) + \beta[(1-\lambda)u(b) + \lambda u(p-1)] \geq u(1 + s_1 - s_2^{No}) + \beta E[u(y_3 + s_2^{No})]. \tag{3}$$

Proposition 1. *The above equation holds if s_1 is bigger than some threshold value, $s_1 \geq s_{min}$.*

(Proofs of all propositions are in Appendix C.)

Proposition 2. *s_{min} is strictly decreasing in β. The effect of λ on s_{min} is ambiguous.*

As in period 3, the level of sophistication does not affect the analysis: The period 2 self knows his true current β, but may mistakenly think that his period 3 self will apply $\tilde{\beta} \geq \beta$ to future decisions. As there are no future decisions once period 3 has been reached, this is of no consequence. Also note that the period 2 self conditions his behaviour on the s_1 received from period 1, regardless of the beliefs held by the period 1 self.

Period 1

Analogue to the minimum s_1 threshold for period 2, it is useful to identify the maximum s_1 that period 1 is willing to save, conditional on purchase of the nondivisible. If this maximum is bigger than the minimum required, the agent is theoretically able to purchase the good (whether saving is successful in equilibrium depends on the coordination between the selves, which is discussed in the equilibrium subsection below). In period 1, sophistication first comes into effect: Period 1 anticipates period 2's decisions, but is overconfident that his future self will be more patient than he is, i.e., he believes his future self uses $\tilde{\beta} \geq \beta$. This belief affects not only his perception of s_{min}, but also directly enters his own optimality considerations via $\tilde{s}_2^{No} \equiv s_2^{No}(\tilde{\beta})$, his perception of s_2^{No}. It is easy to show that $\tilde{s}_2^{No}$ increases in $\tilde{\beta}$, and that $\tilde{\beta} \geq \beta$ implies $\tilde{s}_2^{No} \geq s_2^{No}$. The special case of full sophistication is obtained by setting $\tilde{s}_2^{No} = s_2^{No}$.

Conditional on the nondivisible *not* being purchased (i.e., period 2 is believed to save $\tilde{s}_2^{No} < p-1$), period 1 saves only for precautionary purposes: $s_1^{No} = argmax(u(y_1 - s_1) + \beta E[u(y_2 + s_1 - \tilde{s}_2^{No}) + u(y_3 + \tilde{s}_2^{No})])$ for $s_1^{No} \geq 0$ and $y_t = \{0, 1\}$. The occurence of a shock implies $y_1 = 0$ and thus $s_1^{No} = 0$. Taking into account that the nondivisible can only be bought if no shock hits in any period, the maximum that period 1 would be willing to pay for its expected purchase (i.e., for $s_2 = p-1$) can be found by comparing

$$\begin{aligned} u(1-s_1) &+ \beta(1-\lambda)^2(u(2+s_1-p)+u(b)) \\ &+ \beta(1-\lambda)\lambda(u(2+s_1-p)+u(p-1)) \\ &+ \beta\lambda(u(s_1-\tilde{s}_2^{No})+E[u(y_3+\tilde{s}_2^{No})]) \\ &\geq u(1-s_1^{No})+\beta E[u(y_2+s_1^{No}-\tilde{s}_2^{No})+u(y_3+\tilde{s}_2^{No})]. \end{aligned} \tag{4}$$

Define s_{max} as the maximum value of s_1 such that this inequality holds (if there is no such value, let $s_{max} = 0$).

Proposition 3. *s_{max} is strictly increasing in β.*

Proposition 4. *s_{max} weakly decreases in the amount of naiveté, $\tilde{\beta} - \beta$.*

By assuming desirability of the nondivisible for a time-consistent agent (inequation 2), it follows that $s_{max}(\beta = \tilde{\beta} = 1) \geq \frac{p-1}{2}$. We further know that $s_{max}(0) = 0$. In addition to the maximum which period 1 is willing to save in order to purchase the good, consider the *optimal* way in which period 1 would like to allocate the savings burden of $p-1$ across periods 1 and 2.

Proposition 5. *The optimal allocation of savings from period 1's perspective, denoted $s_1 = s_{opt}$, is characterized by*

$$u'(1-s_{opt}) = \beta[(1-\lambda)u'(2+s_{opt}-p) + \lambda u'(s_{opt}-\tilde{s}_2^{No})(1+\frac{\delta\tilde{s}_2^{No}}{\delta s_1}\cdot\frac{1-\tilde{\beta}}{\tilde{\beta}})]. \tag{5}$$

The term involving $\delta\tilde{s}_2^{No}/\delta s_1$ is a result of the time-inconsistency (for a time-consistent agent, the envelope condition applies): $\tilde{s}_2^{No}$ is chosen optimally given period 2's preferences (more specifically, period 1's belief thereof), which makes it suboptimal from period 1's perspective for $\tilde{\beta} < 1$. As a result, s_1 has a first-order positive effect on $\tilde{s}_2^{No}$.

Proposition 6. *s_{opt} is strictly increasing in β, and always smaller than s_{max}.*

Unfortunately, the effect of sophistication on s_{opt} is ambiguous. Holding β constant and increasing $\tilde{\beta}$ (sophistication falls), period 1 is more confident about period 2 following his interests in the future - in particular with respect to precautionary savings for period 3. As $\tilde{\beta}$ increases, it becomes more attractive to send savings to period 2, as the period 2 self is believed to spread them more equally across periods 2 and 3. On the other hand, period 1 no longer has to overcompensate for period 2's bias, sending excessive savings just to ensure some of them are passed on to period 3. It depends on the specific values of λ, $u''(c)$, $\tilde{\beta}$ and β which effect is stronger.

Autarky Equilibrium with Full Sophistication

Given a decreasing $s_{min}(\beta)$ and an increasing $s_{max}(\beta)$- function, there is a threshold level $\hat{\beta}$ such that $s_{min}(\beta) \leq s_{max}(\beta)$ for any $\beta \geq \hat{\beta}$. The fact that $\hat{\beta}$ is in the relevant interval (0, 1] follows from $s_{min}(0) > s_{max}(0)$ and $s_{min}(1) \leq s_{max}(1)$: The former follows from $s_{min}(0) > 1$, $s_{max}(0) = 0$. The latter is a consequence of desirability (equation 2), by which a time-consistent agent always purchases the good. Since the different period selves are perfectly able to anticipate each other's behaviour, the nondivisible will be purchased (absent shocks) for all $\beta \in [\hat{\beta}, 1]$. Absent shocks, equilibrium savings are

$$s_1 = \begin{cases} max(s_{min}, s_{opt}) & if\ \beta \in [\hat{\beta}, 1] \\ s_1^{No} & if\ \beta \in [0, \hat{\beta}) \end{cases}, \ s_2 = \begin{cases} p-1 & if\ \beta \in [\hat{\beta}, 1] \\ s_2^{No} & if\ \beta \in [0, \hat{\beta}) \end{cases}$$

If a shock occurs in any period, the individual gives up any plans to save for the nondivisible, and merely smoothes available assets $y_t + s_{t-1}$ over future periods, saving $s_t^{No} \geq 0$ for all t after the shock.

Importantly, it is ambiguous whether autarky savings will be above or below $\bar{s} \equiv \frac{p-1}{2}$. This will complicate the later analysis of the regular saver product (e.g., compare Figures 3 and 4). Considering that a time-consistent agent saves $\bar{s}$ (for $\lambda = 0$) or slightly below $\bar{s}$ (for $\lambda > 0$), this question corresponds to O'Donoghue and Rabin's (1999) *pre-emptive overcontrol*: A sophisticated hyperbolic discounter may both save more or less than a time-consistent agent, depending on the numerical values used for $(b - p)$ and $u''(c)$. In the following, scenarios with $s_{min}(\hat{\beta}) = s_{max}(\hat{\beta}) < \bar{s}$ will be referred to as "low autarky savings", and scenarios with $s_{min}(\hat{\beta}) = s_{max}(\hat{\beta}) > \bar{s}$ will be referred to as "high autarky savings".

Autarky Equilibrium with Partial Sophistication

Allowing for partial sophistication, the period 1 agent overestimates the patience of his future self, $\tilde{\beta} > \beta$. As discussed previously, this affects the s_{max}- and s_{opt}- function used by period 1 via precautionary savings. However, the main effect of partial sophistication is to cause period 1 to underestimate the amount of savings s_1 required to convince period 2 to save $s_2 = p - 1$ (and thus facilitate the purchase of the nondivisible). For ease of graphical illustration, assume $\tilde{\beta} = \beta + \gamma$. Denote the resulting perceived s_{min}-function as $\tilde{s}_{min}(\tilde{\beta}) = \tilde{s}_{min}(\beta + \gamma) = \tilde{s}_{min}(\beta) < s_{min}(\beta)$: For a constant sophistication level γ, perceived minimum savings $\tilde{s}_{min}$ can be expressed as a function of β.[45] This allows me to define thresholds in terms of β only:

Define $\check{\beta}$ such that $\tilde{s}_{min}(\beta) \leq s_{max}(\beta)$ for any $\beta \geq \check{\beta}$. For $\beta \in [\check{\beta}, 1]$, the period 1 agent will believe that he is able to save for the nondivisible. Note $\tilde{s}_{min}(\beta) < s_{min}(\beta)$ implies that $\check{\beta} < \hat{\beta}$. Furthermore,

[45]While a constant sophistication level γ is convenient for the purposes of graphical illustration, the model's results do not depend on assumptions about the functional relationship between $\tilde{\beta}$ and β, other than $\tilde{\beta} \geq \beta$.

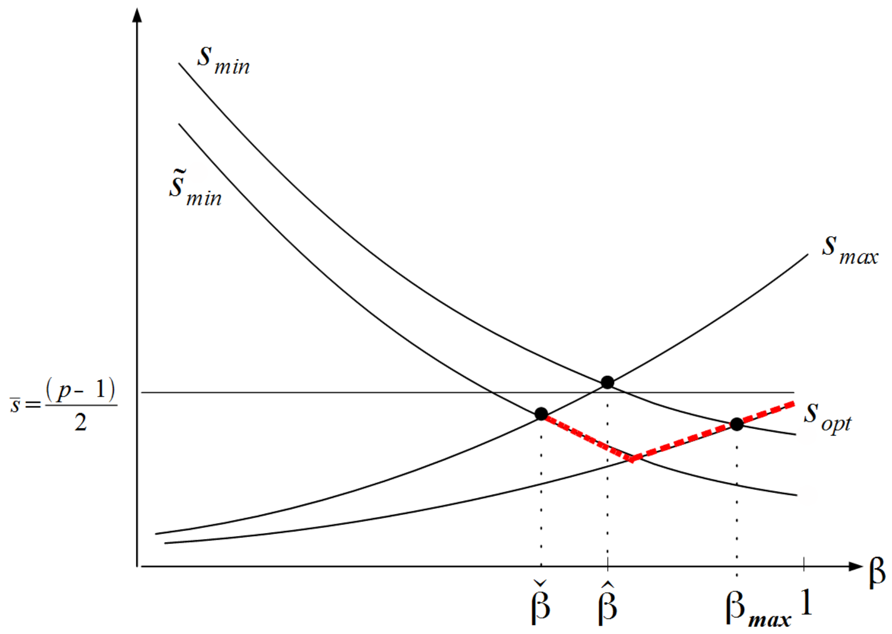


Figure 2: Autarky Equilibrium with Partial Sophistication

define β_{max} such that $s_{min}(\beta) \leq s_{opt}(\beta)$ for any $\beta \geq \beta_{max}$. For $\beta \in [\beta_{max}, 1]$, the optimal savings choice from period 1's perspective is more than required given period 2's true preferences. It follows that $\check{\beta} \leq \hat{\beta} \leq \beta_{max}$. Absent shocks in period 1 and 2, the autarky savings outcome is

$$s_1 = \begin{cases} max\{\tilde{s}_{min}, s_{opt}\} & if\ \beta \in [\check{\beta}, 1] \\ s_1^{No}(\beta) & if\ \beta \in [0, \check{\beta}) \end{cases}, \ s_2 = \begin{cases} p-1 & if\ \beta \in [\beta_{max}, 1] \\ s_2^{No}(\beta) & if\ \beta \in [0, \beta_{max}) \end{cases}$$

The savings path is illustrated by the red dashed line in Figure 2. The most remarkable feature of this savings function is that period 2 "eats" period 1's savings for $\beta \in [\check{\beta}, \beta_{max})$.[46] For $\beta \in [\check{\beta}, \hat{\beta})$, the agent believes he can save for the nondivisible, but is not genuinely able to do so given his true preferences. This is the region where $\tilde{s}_{min}(\beta) \leq s_{max}(\beta) < s_{min}(\beta)$. Period 1 sends $s_1 = \tilde{s}_{min}(\beta)$, anticipating that this will be enough to incentivise period 2 to save $p-1$. Period 2 responds by consuming the savings, transferring only $s_2^{No} < p-1$ to period 3.

Even more paradoxically, for $\beta \in [\hat{\beta}, \beta_{max})$, the agent fails to obtain the nondivisible because of a coordination failure between his different selves: $\tilde{s}_{min}(\beta) < s_{min}(\beta) \leq s_{max}(\beta)$, so the nondivisible could be bought if period 1 saved $s_1 \geq s_{min}(\beta)$. Instead, incorrect beliefs about this future preferences lead him to save $s_1 = max\{\tilde{s}_{min}, s_{opt}\} < s_{min}(\beta)$, and again period 2 consumes period 1's savings. It is only for $\beta \in [\beta_{max}, 1]$ that the savings sent by period 1 are sufficient for purchasing the nondivisible: As β rises, period 1 becomes sufficiently patient to save more than $\tilde{s}_{min}$ voluntarily, eventually reaching the point where $s_{opt}(\beta)$ becomes larger than the required true $s_{min}(\beta)$. Conditional on the absence of shocks, the nondivisible is purchased for the region $\beta \in [\beta_{max}, 1]$.

[46]This is comparable to the theoretical result in Duflo et al. (2011) for farmers' decision to save for fertilizer.

6.3 Equilibrium with a Regular Saver Commitment Product

The following section investigates the effect of offering agents a commitment to fixed regular contributions - as commonly found in loan contracts, pension savings, and other forms of regular saving. As pointed out by Fischer and Ghatak (2010) for the case of microloans, small frequent instalments may mediate time-inconsistency problems of hyperbolic discounters. In a savings setting, commitment to fixed instalments may help agents to reach savings goals, and smooth savings contributions.[47]

The Regular Saver product is defined as follows: Consider an agent who can commit in period 0 to deposit a fixed amount $\bar{s} = \frac{p-1}{2}$ in a bank account in both period 1 and 2. He also chooses a default penalty D, subject only to a limited liability constraint which prevents negative consumption. Once the agent fails to deposit $\bar{s}$ in a period, he is charged the default penalty D, but immediately receives back any accumulated savings. In addition, he is free to save at home independently of his bank contributions. His total cumulated savings (in the bank plus at home) which are transferred from period t to $t+1$ can then be captured as s_t. The penalty D is imposed in period 1 if $s_1 < \bar{s}$, and in period 2 if $s_1 \geq \bar{s}$, $s_2 < 2\bar{s}$. The contract is successfully completed with $s_1 \geq \bar{s}$, $s_2 \geq p-1$. The assumption that the contract is signed in period 0 simplifies things greatly, as the agent is not subject to temptation in this period.[48] As before, the savings outcome can be derived using backwards induction, with a contract-signing period 0 discussed at the end.

Period 3

Period 3 behaviour is identical to that in autarky. The agent will buy the nondivisible whenever he can afford it, i.e., whenever $s_2 \geq p-1$ holds, and absent shocks.

Period 2

Suppose the contract is still active in period 2. In other words, period 1 has not been hit by a shock, and has transferred $s_1 \geq \bar{s}$. Suppose a shock hits in period 2: At an asset level of $s_1 < 1$ and contractual savings of $s_2 = 2\bar{s} = p-1$, default is unavoidable. The resulting consumption level is $c_2 = s_1 - D - s_2^{No} \geq 0$, implying that a penalty of $D \leq s_1$ can be enforced. Absent shocks, period 2 is faced with the decision of whether to send $s_2 = 2\bar{s} = p-1$ (it is never optimal to send $s_2 > p-1$). He is willing to do so if he receives an s_1 that satisfies

$$u(1+s_1-(p-1)) + \beta[(1-\lambda)u(b) + \lambda u(p-1)] \geq u(1+s_1-D-s_2^{No}) + \beta E[u(y_3+s_2^{No})] \quad (6)$$

Since the inequality differs from the autarky case only in the penalty D, the same proof can be used to show that the nondivisible is bought for any $s_1 \geq s^B_{min}$. The threshold $s^B_{min}(\beta)$ will be strictly lower than $s_{min}(\beta)$ in the autarky case: The right-hand side of the inequality decreases when D is introduced, while the left-hand side stays unchanged. The effect of the penalty disappears for $s_1 < \bar{s}$: Period 1 has already defaulted on the contract and paid the penalty, so the contract is no longer active in period 2. As a result, $s^B_{min}(\beta) = s_{min}(\beta)$ for $s_1 < \bar{s}$. The two sections of the $s^B_{min}(\beta)$-function combine with a horizontal

[47] Section 6.1 argues that the first-best savings schedule is $s_1 = \bar{s}$ for $\lambda = 0$, and slightly increasing for $\lambda > 0$, i.e., $s_1 < \bar{s}$. For small λ, this effect is likely to be small. Commitment products with increasing savings schedules are possible, but may pose serious challenges to institutional implementation: The first-best schedule will depend on individual values of λ, $u''(c)$, p and b. The present analysis focuses on fixed-instalment products due to their empirical popularity and ease of administration.

[48] This assumes that the bank can enforce the penalty, even in the case that the agent defaults before depositing any savings. See Section 2 and footnote 27 on how I dealt with this issue in the study.

line at $s^B_{min}(\beta) = \bar{s} = \frac{p-1}{2}$. In this region, the s_{min} required by period 2 is lower than $\bar{s}$ if he faces the penalty, and higher than $\bar{s}$ if he does not. To keep the contract active and ensure that period 2 faces the penalty, the period 1 agent needs to save $s_1 \geq \bar{s}$. Therefore, the minimum s_1 needed to incentivise period 2 to save is $\bar{s}$. Finally, in the region where $s^B_{min}(\beta) > \bar{s}$, the period 2 agent is not willing to save for the nondivisible unless period 1 makes additional savings at home.

Period 1

Consider the maximum s_1 that period 1 is willing to save, once subjected to a penalty for $s_1 < \bar{s}$. Limited liability implies that the penalty cannot be enforced if there is a shock: With no income or previous savings, $c_1 = s_1 = 0$. Absent shocks, period 1 prefers to save for the nondivisible if

$$\begin{aligned} u(1-s_1) &+ \beta(1-\lambda)^2(u(2+s_1-p)+u(b)) \\ &+ \beta(1-\lambda)\lambda(u(2+s_1-p)+u(p-1)) \\ &+ \beta\lambda(u(s_1-D-\tilde{s}^{No}_2)+E[u(y_3+\tilde{s}^{No}_2)]) \\ &\geq u(1-D-s^{No}_1)+\beta E[u(y_2+s^{No}_1-\tilde{s}^{No}_2)+u(y_3+\tilde{s}^{No}_2)]. \end{aligned} \tag{7}$$

As described in Section 6.2, partial sophistication implies that the agent uses $\tilde{s}^{No}_2 \equiv s^{No}_2(\tilde{\beta})$ to assess period 2's behaviour. Full sophistication is nested with $\tilde{s}^{No}_2 = s^{No}_2$. In contrast to the equation for s^B_{min}, both sides of the s^B_{max}- inequality are affected by the penalty. Even for a devoted saver, the penalty is unavoidable if a shock hits in period 2, causing the left-hand side to decrease in D (discounted by $\beta\lambda$). On the right-hand side, the penalty is the consequence of a deliberate decision to default in period 1.

Proposition 7. *For small λ, and in the region $s_1 \geq \bar{s}$, adopting a regular-instalment product increases the maximum the agent is willing to save, i.e., $s^B_{max} > s_{max}$. A sufficient constraint on the shock frequency is $\lambda < \frac{u'(1)}{u'(0.5)}$. In the region $s_1 < \bar{s}$, adopting the regular-instalment product unambiguously decreases s_{max}.*

Note that inequality 7 is specific to the region $s_1 \geq \bar{s}$: The penalty is not charged in period 1 if the agent saves for the nondivisible. Consider the case where necessary savings are $s_1 < \bar{s}$, i.e., period 1 could ensure the good is bought even if he does not contribute $\bar{s}$. In this case, he faces a penalty whether or not he saves for the good. The penalty D enters in period 1 on both sides of the inequality (later periods are unaffected by D, as the contract is no longer active). The resulting threshold $s^B_{max}(\beta)$ is strictly *lower* than the original threshold $s_{max}(\beta)$.

Figure 3 shows that the two sections of the $s^B_{max}(\beta)$-curve combine with a vertical line. To see why, extend the lower section of $s^B_{max}(\beta)$ to the $\bar{s}$-line. For any β in this range, s_{max} is below $\bar{s}$ if the agent is charged the penalty even if he saves for the nondivisible, and s_{max} is above $\bar{s}$ if he is not charged. Since the penalty does not apply for $s_1 \geq \bar{s}$, the maximum that he is willing to pay is given by the upper part of the $s^B_{max}(\beta)$-curve. Even a high $s^B_{max} \geq \bar{s}$ does not rule out that the period 1 agent may optimally save $s_1 < \bar{s}$ for the nondivisible, deliberately incurring the penalty as a "premium" for procrastinating the savings burden onto period 2. Consider $s^B_{opt}(\beta)$, the optimal way in which period 1 would like to split the savings burden $p-1$ across periods 1 and 2, when subjected to the penalty. In autarky, $s_{opt} = \bar{s} > 0.5$ holds only for a time-consistent agent, and given $\lambda = 0$. In the presence of time-inconsistency and a positive shock frequency $\lambda > 0$, s_{opt} is strictly below $\bar{s}$. In consequence, the introduction of a penalty reduces optimal savings further, as the agent needs to pay both s_1 and the penalty D, as a premium for

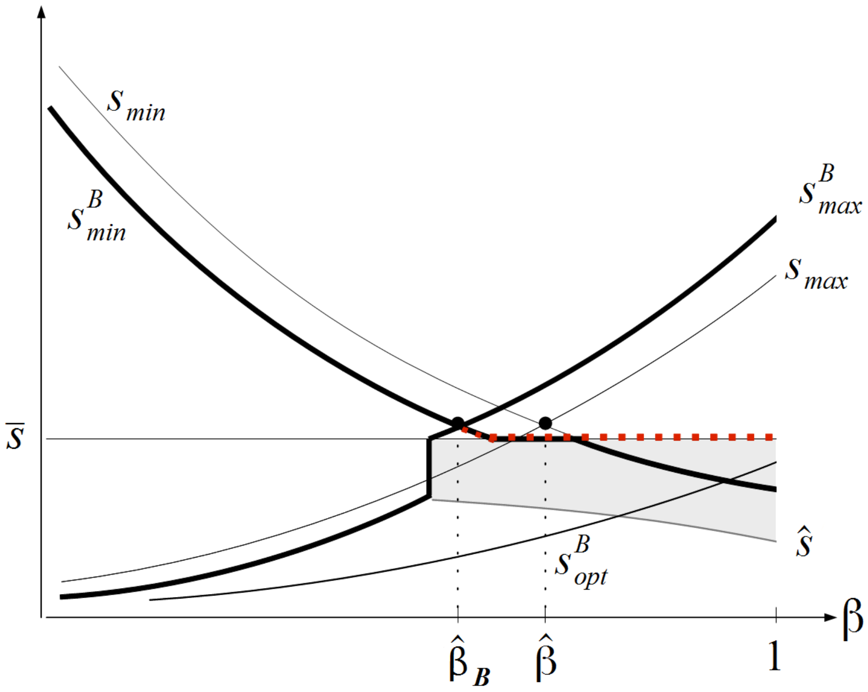


Figure 3: Regular Saver Equilibrium (high autarky savings)

procrastinating savings. Algebraically, $s^B_{opt} < s_{opt}$ follows from inequality 5, after allowing for the fact that period 1's consumption is now $c_1 = 1 - s_{opt} - D$.

Alternatively, period 1 may prefer to jump to $s_1 = \bar{s}$, rather than paying the penalty. The vertical part of s^B_{max} illustrates that it is never optimal to choose savings in the region $\hat{s}(\beta) < s_1 < \bar{s}$, where $\hat{s}(\beta)$ denotes the savings level which makes period 1 indifferent between saving $\hat{s}$ plus paying the penalty D, and saving $s_1 = \bar{s}$, thus avoiding the penalty.[49] Intuitively, if the necessary savings s_1 are such that $s_1 + D > \bar{s}$, then period 1 is trivially better off to save $\bar{s}$. Furthermore, the threshold $\hat{s}$ is strictly lower than $\bar{s} - D$ for $\beta > 0$: At equal instantaneous cost $s_1 + D = \bar{s}$, it is strictly preferable to save $\bar{s}$, for the sake of the additional consumption D in the next period. Finally, willingness to jump to $s_1 = \bar{s}$ requires that $s^B_{max} \geq \bar{s}$. As a result, $\hat{s}(\beta)$ is only defined for the range of β such that $s^B_{max}(\beta) \geq \bar{s}$ (see Figure 3).

Proposition 8. *The threshold $\hat{s}(\beta)$ weakly decreases in β. Equivalently, as β increases, a larger range $s_1 \in (\hat{s}(\beta), \bar{s})$ is strictly dominated by $\bar{s}$.*

Equilibrium and Contract Choice: Full Sophistication

With full sophistication, the nondivisible is purchased whenever $s^B_{max}(\beta) \geq s^B_{min}(\beta)$, which occurs for any $\beta \in [\hat{\beta}_B, 1]$. Equilibrium savings (absent shocks) are analogue to those for autarky, except for a

[49] Formally, $\hat{s}$ is the lowest value of s_1 which satisfies

$$\begin{aligned} &u(1-\hat{s}-D) + \beta(1-\lambda)(u(2+\hat{s}-p)) + \beta\lambda(u(\hat{s} - \bar{s}^{No}_2(\hat{s})) + E[u(y_3 + \bar{s}^{No}_2(\hat{s}))]) \\ &\leq u(1-\bar{s}) + \beta(1-\lambda)(u(2+\bar{s}-p)) + \beta\lambda(u(\bar{s} - D - \bar{s}^{No}_2(\bar{s})) + E[u(y_3 + \bar{s}^{No}_2(\bar{s}))]). \end{aligned}$$

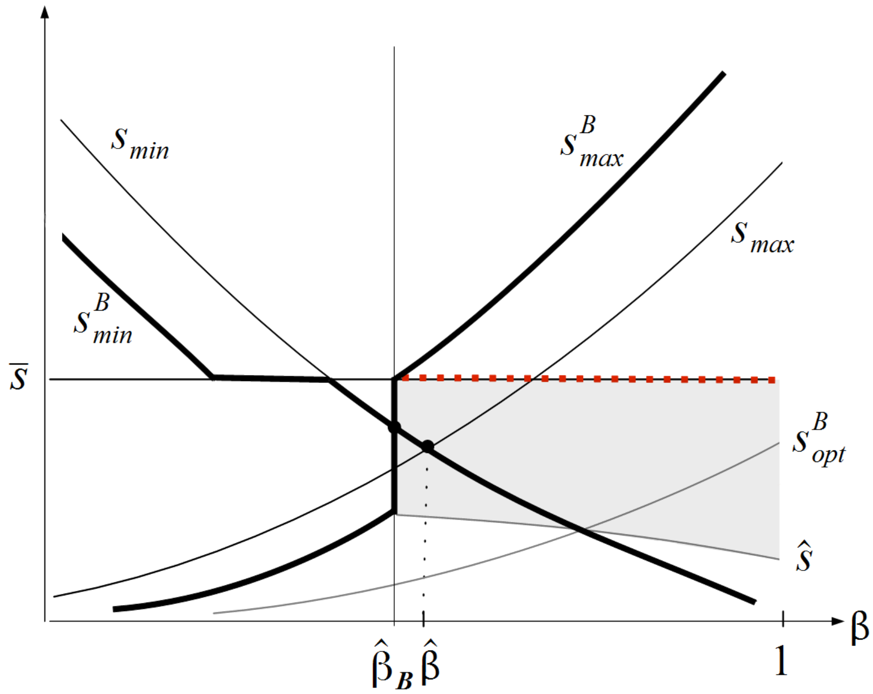


Figure 4: Regular Saver Equilibrium (low autarky savings)

lower savings threshold $\hat{\beta}_B < \hat{\beta}$, and a dominated region $s_1 \in (\hat{s}(\beta), \bar{s})$:

$$s_1 = \begin{cases} max(s^B_{min}, s^B_{opt}) & if\ \beta \in [\hat{\beta}_B, 1] \text{ and } max(s^B_{min}, s^B_{opt}) \notin [\hat{s}, \bar{s}) \\ \bar{s} & if\ \beta \in [\hat{\beta}_B, 1] \text{ and } max(s^B_{min}, s^B_{opt}) \in [\hat{s}, \bar{s}) \\ s_1^{No} & if\ \beta \in [0, \hat{\beta}_B) \end{cases}, \quad s_2 = \begin{cases} p-1 & if\ \beta \in [\hat{\beta}_B, 1] \\ s_2^{No} & if\ \beta \in [0, \hat{\beta}_B). \end{cases}$$

Figures 3 and 4 illustrate the Regular Saver equilibrium with full sophistication. Figure 3 shows the savings path starting from high autarky savings, while Figure 4 starts from low autarky savings. It is critical for the welfare implications of the regular saver product whether $\hat{\beta}_B \leq \hat{\beta}$. In other words, is the nondivisible achievable for a larger range of preferences when the regular saver product is used? The answer is yes, given a sufficiently large penalty. Since the period 0 agent chooses the penalty himself, $\hat{\beta}_B \leq \hat{\beta}$ is guaranteed to hold under full sophistication.[50] As a result, for $\beta \in [\hat{\beta}_B, \hat{\beta})$, the nondivisible is achievable with the regular-instalment product, but not without it. The threshold $\hat{\beta}_B$ decreases in the size of the chosen penalty D (a corollary of Proposition 9). Furthermore, for the region $\beta \in [\hat{\beta}, 1)$, the Regular Saver product weakly smoothes savings contributions (and thus consumption) towards $\bar{s}$.

Period 0 Adoption Decision and Penalty Choice In principal, any agent with $\beta \in [0, 1)$ can benefit from commitment. Given a sufficiently large penalty, it makes the nondivisible achievable and smoothes savings: Absent shocks, the contract is trivially enforceable in period 1 if $D > \bar{s}$, and in period 2 if $D > 2\bar{s}$. Even with $\beta = 0$, it is cheaper for the agent to make the contracted-upon savings contribution than to pay the penalty. As a result, the threshold $\hat{\beta}_B$ can be moved to an arbitrarily low

[50]Formally, $\hat{\beta}_B \leq \hat{\beta}$ holds regardless of penalty size if starting from a high-savings autarky scenario with $s_{min}(\hat{\beta}) = s_{max}(\hat{\beta}) \geq \bar{s}$. For scenarios with low autarky savings $s_{min}(\hat{\beta}) = s_{max}(\hat{\beta}) < \bar{s}$, the penalty D needs to be large enough to make an agent with $\beta = \hat{\beta}$ willing to jump to $s^B_{max}(\hat{\beta}) \geq \bar{s}$ to ensure $\hat{\beta}_B \leq \hat{\beta}$.

β. The downside of commitment is the risk of "rational default" due to shock frequency λ: The penalty not only acts to discipline the agent when income is available, it also needs to be paid when the agent no longer finds it welfare-maximising (or feasible) to save for the nondivisible. Limited liability implies that this risk is limited to shocks in period 2: If a shock hits in period 1, the agent has no assets or income, thus the penalty cannot be enforced. In period 3, the contract is no longer active. In contrast, if a shock hits in period 2, the agent's savings of $s_1 \geq \bar{s}$ may be lost to the penalty D, leaving the agent worse off than if he had not adopted commitment.

The resulting decision is a two-step problem: The period 0 agent first decides which penalty D offers the optimal trade-off between commitment and flexibility. He then makes a binary choice between adopting the regular saver product with the optimal penalty, or not adopting the product. Unfortunately, the choice of the optimal penalty is non-monotonic in β, and sensitive to the autarky scenario, due to the consumption smoothing motive: Consider starting from a low $\beta < \hat{\beta}$ in Figure 3 (high autarky savings). Increasing D will first shift the upper part of the s^B_{min}- and s^B_{max}-curve to the left, until $s^B_{max}(\beta) = s^B_{min}(\beta)$ holds for the agent's β (in other words, until $\hat{\beta}_B = \beta$). The nondivisible is now achievable, but at a skewed savings schedule $s_1 = s^B_{min}(\hat{\beta}_B) > \bar{s}$. The agent may choose to increase D further, in order to decrease s^B_{min} and smooth savings towards $\bar{s}$. However, the benefit associated with smoother savings contributions is a discrete drop from the benefit associated with achieving the nondivisible, and the agent may not deem it worthwhile to increase D further in the face of shock frequency λ. To see why the optimal penalty is non-monotonic in β, consider starting from a high $\beta \geq \hat{\beta}$ in Figure 4 (a scenario with low autarky savings). In autarky, the nondivisible is achievable, and the agent saves $s_1 = max(s_{min}, s_{opt}) \ll \bar{s}$. While the Regular Saver product is not needed to achieve the nondivisible, it can help to smooth consumption: As D increases, $\hat{s}(\beta)$ falls, and the dominated region $s_1 \in (\hat{s}(\beta), \bar{s})$ becomes larger, until it eventually includes $max(s^B_{min}, s^B_{opt})$.[51] The period 0 agent would like to choose the penalty at the minimum level which will make him jump to $s_1 = \bar{s}$.[52] Thus, he will choose D such that $\hat{s}(\beta) = max(s^B_{min}, s^B_{opt})$ holds exactly. As $max(s^B_{min}, s^B_{opt})$ first decreases in β along with s^B_{min}, and then increases in β along with s^B_{opt}, the penalty required to make $\hat{s}(\beta) = max(s^B_{min}, s^B_{opt})$ first increases and then decreases in β. Depending on $u''(c)$ and λ, the take-up decision for $\beta \in [\hat{\beta}, 1)$ (where the Regular Saver product is exclusively used for consumption smoothing) may be similarly non-monotonic.

For the sake of simplicity, I will abstract from the consumption smoothing motive, and focus on the range of $\beta \in [0, \hat{\beta})$. For this range of β, the nondivisible is not achievable in autarky, and obtaining it constitutes the primary benefit of the Regular Saver product. This focus is empirically meaningful: It restricts the analysis to the part of the population who are not able to save for lump-sum consumption expenditures by themselves, i.e., without the use of commitment. This is consistent with data from my sample population.[53] Define $D_{eff}(\beta)$ to be the minimum effective penalty which achieves $s^B_{max}(\beta) \geq s^B_{min}(\beta)$. Given full sophistication, a Regular Saver contract with a penalty D_{eff} will enable the agent to save for the nondivisible (absent shocks). By construction, $D_{eff} = 0$ for $\beta \geq \hat{\beta}$.

[51] Note that the s^B_{min}-curve is unaffected by D in the region $s_1 < \bar{s}$, as the contract is no longer active in period 2. Meanwhile, s^B_{opt} decreases in D, as it factors in the default on the contract. Therefore, the only possibility to smooth consumption via penalty D is through its effect on the dominated region $(\hat{s}(\beta), \bar{s})$.

[52] This is a simplification: The first-best is to save slightly below $\bar{s}$, which is not feasible with a regular-saver product. However, for small λ, this difference is small, and the agent is better off with a smooth savings schedule $\bar{s}$, compared to leaving the savings decision entirely at the discretion of period 1.

[53] 35 percent of the study population reported zero savings of any form, the median level of liquid assets (bank and home savings) was 500 pesos (U.S.$ 12), and the most common way to afford lump-sum expenditures was through high-interest borrowing (which includes a commitment to fixed instalments).

Proposition 9. *For a given λ, the minimum effective penalty D_{eff} weakly decreases in β.*

Proposition 10. *The optimal Regular Saver contract for a fully sophisticated agent with $\beta < \hat{\beta}$ depends on the effect of the minimum effective penalty, $D = D_{eff}$: Where D_{eff} results in $s^B_{min}(\beta) \leq s^B_{max}(\beta) \leq \bar{s}$ (as in Figure 4), the optimal contract is to choose D_{eff}. This achieves perfectly smooth equilibrium savings contributions $s_1 = \bar{s}$. Where D_{eff} results in $s^B_{max}(\beta) \geq s^B_{min}(\beta) > \bar{s}$ (as in Figure 3), the optimal contract involves $D \geq D_{eff}$, and achieves equilibrium savings $s_1 \geq \bar{s}$.*

For plausible ranges of the parameters, the case where D_{eff} guarantees perfect consumption smoothing $s_1 = \bar{s}$ (and thus eliminates the need to choose a higher penalty) coincides with the "low autarky savings" scenario.[54] The specific parameter restrictions needed are the subject of current research.

Having determined the optimal penalty for the Regular Saver product, the period 0 agent then faces the binary decision of whether or not to take up the product. The following inequality is sufficient for take-up to be optimal:

$$
\begin{aligned}
&(1-\lambda)^3[u(1-s_1) + u(2+s_1-p) + u(b)] \\
&+ (1-\lambda)^2\lambda[u(1-s_1) + (u(2+s_1-p) + u(p-1)] \\
&+ (1-\lambda)\lambda[u(1-s_1) + u(s_1 - D_{eff} - s_2^{No}) + E(u(y_3 + s_2^{No}))] \\
&+ \lambda[u(0) + E(u(y_2 - s_2^{No}) + u(y_3 + s_2^{No}))] \\
&\geq E[u(y_1 - s_1^{No}) + u(y_2 + s_1^{No} - s_2^{No}) + u(y_3 + s_2^{No})]
\end{aligned}
\tag{8}
$$

where $s_1 = max\{\bar{s}, s^B_{min}(\hat{\beta}_B)\}$ and $y_t = \{0, 1\}$ depending on the realisation of shocks. The rows of inequality 8 describe the different cases of shock occurence: The savings plan could be undisturbed by shocks until the end of the agent's life (first row), it could fail in period 3 (second row: period 3 lacks the income to buy the nondivisible), a shock in period 2 could lead to costly default (third row), or a shock in period 1 could prevent saving for the nondivisible altogether (fourth row).

This leads to the following results for period 0's adoption decision: Whether the agent adopts commitment will critically depend on shock frequency λ, nondivisible benefit b, price p, and required penalty $D_{eff}(\beta)$. However, ceteris paribus, those with the lowest values of β will require the highest penalties D_{eff}. Consequently, those with the lowest values of β are the *least* likely to adopt the product. To see this, realise that the benefit of an effective commitment contract (obtaining the nondivisible with a smooth schedule $s_1 = \bar{s}$) is independent of β: The period 0 agent bases his adoption decision on the welfare function $W = E[u(c_1) + u(c_2) + u(c_3)]$, which does not directly depend on β. Put simply, the time-inconsistency parameter β determines how difficult it is for the agent to save for the nondivisible, but not how much he benefits from achieving it.[55] As a result, agents with low β (and therefore a high required penalty D_{eff}) will find that commitment harms them in expectation, and will not adopt it.

The result on welfare is straightforward: Given full sophistication, everyone who adopts the commitment product is made better off in expectation. Agents perfectly anticipate their own behaviour,

[54]As discussed in the autarky section, "low autarky savings" refers to a situation where $s_{min}(\hat{\beta}) = s_{max}(\hat{\beta}) < \bar{s}$. It does not refer to the specific savings made by the agent in autarky.

[55]Strictly speaking, the benefit from commitment is only independent of β for $\lambda = 0$. With $\lambda > 0$, the period 0 agent has to rely on his future selves to make precautionary savings. The lower is β, the larger is the disagreement between the selves over how much should be saved for shocks. The commitment contract insures the agent against shocks at least in period 3 (no savings are available if a shock hits in period 1, and a shock in period 2 would leave the agent with $s_1 - D \geq 0$). Since precautionary savings decrease in β, the insurance effect of commitment is slightly more valuable for lower β. However, this effect is unlikely to quantitatively dominate the offputting effect of a higher required penalty D_{eff}.

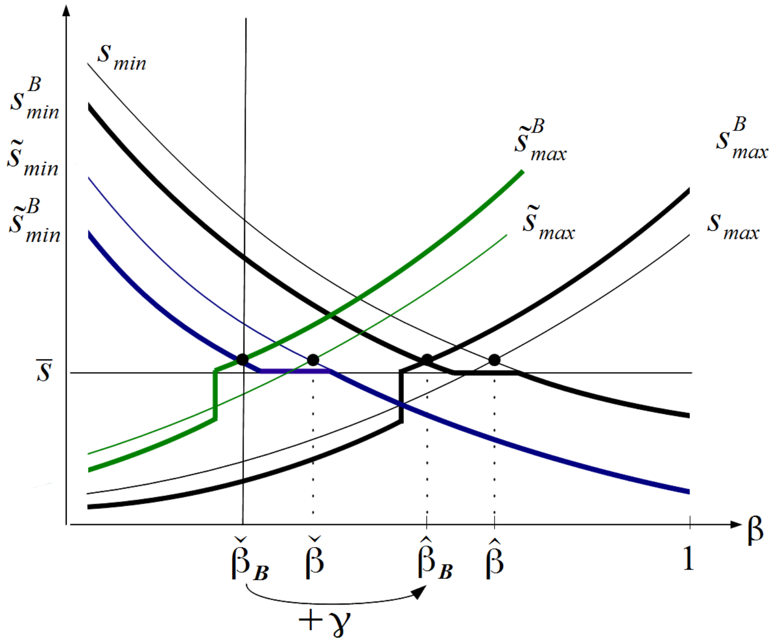


Figure 5: The Regular Saver Equilibrium with Partial Sophistication

and assess the required degree of commitment (D_{eff}) correctly. The only reason for contract defaults are shocks: A fraction λ of adopters defaults each period. In summary, commitment through a regular-instalment product will be weakly welfare-increasing for sophisticated hyperbolic discounters.

Equilibrium and Contract Choice: Partial Sophistication

The derivations for penalty choice and adoption decision for partially sophisticated agents are analogous to those for full sophistication – except that the period 0 agent systematically applies an incorrect belief $\tilde{\beta} > \beta$. This results in a biased perception not only of $\tilde{s}_{min} < s_{min}$ (as for period 1), but also of $\tilde{s}_{max}(\beta) \equiv s_{max}(\tilde{\beta}) > s_{max}(\beta)$, as period 0 is overconfident about the patience he will have in period 1. Since the same belief $\tilde{\beta}$ is used to assess $\tilde{s}_{min}$ and $\tilde{s}_{max}$, the partial sophistication bias in period 0 graphically corresponds to a shift in the entire schedule by a constant $\tilde{\beta} - \beta \equiv \gamma$ (see Figure 5).

The analysis will focus on those agents with $\beta < \tilde{\beta} < \hat{\beta}$: The part of the population who is not only unable to save without commitment, but who is also aware of this fact (for instance, because they have not observed themselves save in the past). Given a large benefit b of the nondivisible good, the primary motivation of such agents for adopting the regular-instalment commitment savings product will be to achieve the nondivisible. The minimum penalty which is perceived to be effective in making the nondivisible achievable (in other words, the penalty which results in $\tilde{s}^B_{min} = \tilde{s}^B_{max}$) is then $D_{eff}(\tilde{\beta})$, denoted $\tilde{D}_{eff}$. By construction, $\tilde{D}_{eff} = 0$ for $\tilde{\beta} \geq \hat{\beta}$.

The optimal penalty choice for partially sophisticated agents is a corollary of Proposition 10: An agent who believes to have $\tilde{\beta} < \hat{\beta}$ will unambigously choose the perceived minimum effective penalty, $D = \tilde{D}_{eff}$, whenever he anticipates that this will result in perfect consumption smoothing, i.e., when $\tilde{s}^B_{min}(\beta) \leq \tilde{s}^B_{max}(\beta) \leq \bar{s}$ at $\tilde{D}_{eff}$. This is likely to happen under low autarky savings scenarios. Where $\tilde{D}_{eff}$ results in $\tilde{s}^B_{max}(\beta) \geq \tilde{s}^B_{min}(\beta) > \bar{s}$ (typically in high autarky savings scenarios), the agent chooses

$D \geq \tilde{D}_{eff}$, and anticipates equilibrium savings $s_1 = \tilde{s}^B_{min} \geq \bar{s}$. By Proposition 9, D_{eff} decreases in β. Therefore, $\tilde{\beta} > \beta$ implies $\tilde{D}_{eff} < D_{eff}$.

The take-up decision is determined in the same way as for fully sophisticated agents, and captured in equation 8. The equation does not (directly) depend on β, but it depends on the (perceived) effective penalty D_{eff}. Partially sophisticated agents differ from fully sophisticated agents precisely in the fact that they perceive a lower $\tilde{D}_{eff} < D_{eff}$ to be sufficient. As a response, for the range of $\beta < \tilde{\beta} < \hat{\beta}$, the regular-instalment product is more attractive to partially sophisticated agents than to fully sophisticated agents: Conditioning on β, and holding λ, p, b and $u''(c)$ constant, those with higher sophistication gaps γ have a lower $\tilde{D}_{eff}$, and are thus more likely to adopt the product.

The savings outcome is a function of the chosen penalty. In addition, it may critically depend on the degree of *learning* which the agent undergoes during his life: In period 0, he believes he will use the parameter $\tilde{\beta}$ in all future periods. In period 1, he realises his true current β. In a static model where the agent does not update his beliefs after he observes his behaviour, period 1 will continue to believe that he will use $\tilde{\beta}$ in the future (much like a dieter who observes himself eating chocolate, but repeatedly plans to be more disciplined tomorrow). The other extreme is full updating: As period 1 learns his true current β, he updates his belief to $\tilde{\beta} = \beta$ for all future periods. Ali (2011) characterizes conditions under which agents' beliefs converge to full sophistication, presuming Bayesian updating. This paper discusses the two extreme assumptions: The case without updating, and the case of full updating.

Suppose the minimum effective penalty is perceived sufficient to guarantee perfect consumption smoothing, i.e., $\tilde{s}^B_{min}(\beta) \leq \tilde{s}^B_{max}(\beta) \leq \bar{s}$ at $\tilde{D}_{eff}$. The period 0 agent chooses $\tilde{D}_{eff}$ and expects $s_1 = \bar{s}$. In period 1, the agent learns his true current β, and thus $s^B_{max} < \tilde{s}^B_{max}$. Without updating, period 1 still believes in $\tilde{s}^B_{min}$. With full updating to $\tilde{\beta} = \beta$, the agent also learns the true $s^B_{min} > \tilde{s}^B_{min}$. In this scenario, updating is of no consequence for the savings outcome: $\tilde{D}_{eff}$ is constructed to make $\tilde{s}^B_{min} = \tilde{s}^B_{max}$ hold exactly. The realisation that $s^B_{max} < \tilde{s}^B_{max}$ is sufficient to inform the agent that saving is not feasible: Whether he believes in $\tilde{s}^B_{min}$ or s^B_{min} only determines the size of the gap $s^B_{max} < \tilde{s}^B_{min} < s^B_{min}$ which keeps him from saving (see Figure 5). As a response, he abandons his savings plan in period 1, pays the penalty, and saves s^{No}_1.

Starting from a situation where $\tilde{D}_{eff}$ provides incomplete consumption smoothing, i.e., $\tilde{s}^B_{max}(\beta) \geq \tilde{s}^B_{min}(\beta) > \bar{s}$, will generally produce the same result: Agents choose their penalty at $\tilde{D}_{eff}$ or slightly above. For most parameter specifications, period 1's realisation that $s^B_{max} < \tilde{s}^B_{max}$ will result in $s^B_{max} < \tilde{s}^B_{min} < s^B_{min}$, which leads to immediate contract default irrespective of learning behaviour. For illustration, consider an agent with $\beta = \check{\beta}_B$ in Figure 5, who believes that his future selves will use $\tilde{\beta} = \hat{\beta}_B$. The agent's β and $\tilde{\beta}$ are at the banking thresholds $\check{\beta}_B$ and $\hat{\beta}_B$ by construction of the penalty $\tilde{D}_{eff}$ and the assumed function $\tilde{\beta} = \beta + \gamma$.

However, the result may differ in cases where the agent has a particularly strong motive for consumption smoothing: The agent may voluntarily increase the penalty beyond $\tilde{D}_{eff}$ in order to reduce $\tilde{s}^B_{min} \geq \bar{s}$ and get closer to $\bar{s}$. The success of this endeavour depends on the size of the penalty, and on learning: The agent may increase D until $\tilde{s}^B_{min} = \bar{s}$ holds exactly (higher penalties cannot be optimal, since their only effect is to increase the cost of default in case of a shock). In period 1, the agent realises $s^B_{max} < \tilde{s}^B_{max}$. Without updating, if $\tilde{s}_{min} = \bar{s} < s_{max}$, the agent still believes he is able to save, and transfers $s_1 = \tilde{s}_{min} = \bar{s}$ to period 2 (this follows from $s_1 = max(\tilde{s}^B_{min}, s^B_{opt})$ and $s^B_{opt} < \bar{s}$). However, in reality, $s^B_{min} > \tilde{s}^B_{min} = \bar{s}$. The penalty that is sufficient to reduce $\tilde{s}^B_{min}$ to $\bar{s}$ is not sufficient to reduce s^B_{min} to $\bar{s}$. Comparable to the coordination failure in autarky, once period 2 arrives, the agent eats his savings, and fails to save for the nondivisible. The situation is welfare-reducing relative to autarky, as the effect of an uneven consumption path is exacerbated by the loss of the penalty D. Instead, consider the case with full updating: In period 1, he learns that $s^B_{min} > \tilde{s}^B_{min} = \bar{s}$. If the chosen penalty is large enough to guarantee $s^B_{max} \geq s^B_{min}$ for his true preferences, the agent is willing to save $s_1 = s^B_{min} > \bar{s}$. While the agent fails to achieve consumption smoothing, updating his beliefs enables him to avoid contract default, and obtain the nondivisible. When do such cases occur? The motive for consumption smoothing must be large, and the sophistication gap low. Therefore, successful saving under partial sophistication is most likely to occur for high autarky savings, small sophistication gaps γ, small shock frequency λ (so the agent is less averse to big penalties), and large nondivisible prices p (increasing the benefits to consumption smoothing).

The resulting welfare implications are discouraging: For $\beta < \tilde{\beta} < \hat{\beta}$, and without updating of beliefs, all partially sophisticated adopters default. Agents are particularly likely to adopt the contract if they have a high $\tilde{\beta}$, as is the case for those with large sophistication gaps γ. Default always occurs in period 1 when choosing $D = \tilde{D}_{eff}$. Welfare is unambigously reduced: It decreases from $W_A = E[u(y_1 - s^{No}_1) + u(y_2 + s^{No}_1 - s^{No}_2) + u(y_3 + s^{No}_2)]$ in autarky to $W_{RS} = E[u(y_1 - s^{No}_1 - D) + u(y_2 + s^{No}_1 - s^{No}_2) + u(y_3 + s^{No}_2)]$ with the commitment product. When choosing $D > \tilde{D}_{eff}$, default in period 2 is possible under some parameter specifications. Finally, with full updating of beliefs, the agent may be able to fulfill the contract and obtain the nondivisible under parameter specifications which strongly encourage consumption smoothing.

7 Alternative Explanations for Default

Previous sections have focused on partially sophisticated hyperbolic preferences in explaining why a majority of individuals who choose to adopt a regular-instalment commitment product will default soon after opening their accounts. This section will consider alternative explanations: Income optimism, aggregate shocks, and limited attention.

7.1 Income Optimism

As suggested by Browning and Tobacman (2007), the consumption behaviour of someone who is overoptimistic about his future income distribution cannot be distinguished from someone who is impatient – both will overconsume in the present. Overoptimistic beliefs about future income could explain the observed measure of time-inconsistency (from MPL questions): If individuals expect their future income to be higher than their current income, they may select the smaller, sooner reward when presented with the 'now vs. 1 month' frame, but choose the larger, later reward when presented with the '1 month vs. 2 months' frame. As a result, they would be falsely classified as present-biased. Income optimism could further explain default incidence: If people were overoptimistic about their income when they adopted the Regular Saver product, and realised this upon starting their savings plan, default may have become an optimal response.

Using data on predicted and realised incomes, I construct a measure which plausibly captures income optimism for groups. It is impossible to identify optimism on an individual level – an individual who reports to have lower income than predicted may either experience a bad draw from a correct income distribution (the 'bad luck' explanation), or he may have systematically biased beliefs about his income distribution ('optimism'). However, the law of large numbers implies that individuals should

TABLE VIII: INCOME OPTIMISM

	Not Present-Biased	Present-biased	All	T-stat P-value
Prediction Gap (growth)	3.290378	3.677686	3.357041	0.81
	(0.6976)	(1.2298)	(0.6146)	
Prediction Gap (level)	1.269759	-2.22314	0.6685633	0.22
	(1.1704)	(2.6292)	(1.0698)	
Observations	582	121	703	
	No Take-Up	**Take-Up**	**All**	**T-stat P-value**
Prediction Gap (growth)	1.738007	5.043011	2.582418	0.08
	(0.9255)	(1.8137)	(0.8325)	
Prediction Gap (level)	1.140221	-2.569892	0.1923077	0.30
	(1.8676)	(2.6941)	(1.552)	
Observations	271	93	364	
	Successful	**Default**	**All**	**T-stat P-value**
Prediction Gap (growth)	4.227273	5.77551	5.043011	0.67
	(2.6477)	(2.5106)	(1.8137)	
Prediction Gap (level)	-5.318182	-0.1020408	-2.569892	0.34
	(3.7117)	(3.8799)	(2.6941)	
Observations	44	49	93	

Standard deviations in parentheses. All numbers are group averages.

correctly predict their income *on average* if their beliefs about income are unbiased. On the other hand, if the present bias measure captures income optimism rather than time-inconsistency, then individuals classified as present-biased should have higher predicted-minus-realised income gaps than those classified as not present-biased. Further, if defaults were caused by individuals systematically misjudging their future income, then defaulting clients should have higher prediction gaps than those who successfully completed their contract.

Table VIII presents group averages of prediction gaps across three dimensions: The observed measure of present bias, take-up of the Regular Saver product ASA, and default on ASA. Prediction gaps are measured as follows: During the baseline survey in September and October 2012, individuals were asked to predict their average weekly household income for each month from October 2012 to March 2013. To make this task easier, individuals chose one of 31 income brackets, numbered from 1 for '0-50 pesos per week' to 31 for 'more than 10,000 pesos per week'. Six months later, in late March and April 2013, this exercise was repeated during the endline survey, except that individuals now stated their realised weekly income for the same time period. Two measures of optimism (or bad luck) are obtained: $Prediction\ Gap\ (growth)_i$ is the difference between predicted income growth and realised income growth, where growth is measured as $Growth_i = \sum_{m=Nov}^{Mar} (bracket_m - bracket_{October})$. In other words, income growth is proxied by the sum of deviations from October income, in units of income brackets. This approach is conservative, in the sense that it is robust to individuals using different income benchmarks for their October income in baseline and endline survey.[56] An alternative measure of optimism is $Prediction\ Gap\ (level)_i$, obtained by the simple difference between predicted and realised income levels

[56]For instance, individuals might have referred to the household income of their core household in the baseline survey, and their extended household in the endline survey, or vice versa. Clear definitions of what constitutes a household were provided, but some grey areas were unavoidable (e.g., where families lived with uncles or cousins, and shared a common budget for food, but not for other household expenses).

(summed), $Prediction\, Gap\, (level)_i = \sum_{m=Oct}^{Mar} (bracket_m^{pred} - bracket_m^{real})$. Consistent with noise in benchmark income levels, $Prediction\, Gap\, (level)_i$ exhibits more variation than $Prediction\, Gap\, (growth)_i$. Note that these measures cannot be included as covariates in take-up or default regressions – both because they are not meaningful on an individual level, and because they use data from the endline survey, and may thus not be orthogonal to treatment.

The sample for Table VIII are those individuals who participated in both the baseline and endline survey. The average prediction gap for income growth across the sample was 3.36 brackets, suggesting that moderate income optimism may be common. However, the average prediction gap is not higher for individuals classified as present-biased – if anything, the level measure suggests they may have been more pessimistic. In contrast, the average prediction gap is significantly higher for individuals who adopted the ASA product compared to those who did not, suggesting that those entering commitment contracts may have been more optimistic about their future income. Finally, individuals who defaulted on ASA did not report significantly higher prediction gaps for income growth than did clients who successfully completed their contract. However, it is worth noting that the level measure points to a possible pessimism of successful clients.

Summing up, there is mixed evidence that those who adopted ASA were optimistic about the growth of their income, relative to those who rejected the offer. The evidence does not suggest a connection between optimism and the observed measure of present bias. In addition, income optimism alone cannot explain why individuals demand commitment. Further, it does not provide a rationale for the observed link with the sophistication measure (which is based on self-reported temptation).

Similar arguments apply for optimism regarding the shock frequency λ (as discussed in Section 6): For instance, individuals could have heterogeneous shock frequencies λ_i, where shocks may refer to income shocks, consumption emergencies, and more generally the risk that saving may no longer be optimal. With rational expectations about λ_i, individuals with high shock frequencies are ceteris paribus less likely to select into commitment. However, if individuals have biased beliefs about λ_i (such as the belief that one's shock frequency rate corresponds to the average shock frequency in the population), then the consequence of a commitment contract may be a bulk of defaults soon after opening (as those individuals with the highest λ_i are likely to drop out first). Therefore, biased beliefs about the shock frequency provide another potential explanation for default occurence. Its limitation is similar to that of income optimism: Biased beliefs about λ alone do not predict a demand for commitment. Neither do they explain a correlation with measures of present-bias or sophistication.

Less parsimonious explanations may involve a combination of different factors, such as fully sophisticated hyperbolic preferences in combination with income optimism. This combination may predict both a demand for commitment and subsequent default. However, it fails to explain why measures of sophistication are *negatively* associated with take-up and default. In this sense, partial sophistication provides a parsimonious explanation that is consistent with the evidence.

7.2 Aggregate Shocks

Idiosyncratic and independent shocks are unlikely to cause the default timing pattern apparent in Figure 1. However, if an aggregate shock hit the sample population around the time of account opening, this may help to explain why 55 percent of clients defaulted shortly after adopting the product. The Philippines is a well-known area for earthquakes and tropical storms, and had recently been hit by tropical storm Washi (Philippine name 'Sendong') in December 2011, causing 1,268 casualties (more than half of them in Cagayan de Oro, a city 126km west of the study location).[57] The risk of such shocks was thus well-known at the time of marketing in September 2012, possibly affecting take-up rates. Indeed, tropical storm Bopha (Philippine name 'Pablo') hit the Mindanao region between December 2 and December 9, 2012. As opposed to storm Washi, storm Bopha did not cause flash flooding, and the main effect on the study location was a six-day power outage. While this may have affected on large businesses, power outages of several hours each day were common in the study area even before the storm, and provisions against power outages were widespread. Because of its limited effect on the area, storm Bopha was not locally classified as a natural disaster (which would have invoked both ASA's and Gihandom's emergency provisions). In the endline survey, 20.5 percent of the sample population reported some damage to their house or crops, with a median damage value of 1400 pesos (U.S. $33, conditional on non-zero damage). Within the sample of defaulting ASA clients, the percentage affected by the storm was 20.4. Asked whether they suffered reductions in income because of the power outages, only 3 out of 732 endline survey respondents answered in the affirmative.

While some negative effects of the storm cannot be ruled out, the timing of the storm does not match the timing of the defaults: The ASA accounts were opened between 20 September and 28 October. Out of 63 defaults, 35 made no further deposit after their opening balance, resulting in contract default upon the third missed deposit, three weeks after opening.[58] An additional 15 clients made one or two deposits after opening (see Figure 1 for the distribution of transactions). By the time of the storm in early December, most of the contract defaults had already occurred.

7.3 Limited Attention

An intuitive explanation for default suggests that clients may have simply forgotten to make their weekly deposits. Limited attention models such as that of Banerjee and Mullainathan (2008) suggest that attention is a scarce resource, which needs to be divided between home and work in order to catch emerging problems before they cause damage. In their model, the amount of damage an individual suffers from problems ocurring at home or at work (such as a child's sickness, or running out of stock for one's business) is a function of the attention which the individual invests into home life and workplace. Given the relatively low stakes of the Regular Saver account (with default penalties roughly equivalent to a day's household income), it would be understandable if individuals prioritised their attention on their home and work lives, rather than on their bank accounts. However, this explanation predicts that individuals would not take up the Regular Saver product in the first place: During the marketing stage, ASA was clearly presented as attention-intensive: Clients were presented with an explicit savings plan including due dates for each week, and given the instruction to physically deposit their instalments at the bank. Most respondents received their income in cash, and bank transfers were uncommon. In the Banerjee and Mullainathan (2008) model, not investing attention in one aspect of one's life incurred a risk that a costly problem would go unnoticed. In contrast, not investing attention in the ASA product (by adhering to the weekly schedule) resulted in *certain* default. As a result, if the returns to investing attention at work or at home exceeded the returns to investing attention in the ASA schedule, then individuals should not have adopted the product. The data suggests that this indeed reduced take-up: Among the clients who were assigned to the Regular Saver treatment but chose not to adopt the

[57] Statistics from the Philippine National Disaster Risk Reduction and Management Council (NDRRMC).
[58] 85 percent of ASA clients opted for weekly deposits, 15 percent opted for bi-weekly deposits.

product, being "too busy to go to the bank" was a common reason for rejecting. Among those clients who accepted the offer, "distance to the bank branch" does not predict default (as measured by GPS coordinates, see Table V).

8 Conclusion

Commitment devices are receiving increasing attention both in the academic literature and in the public eye, and are generally portrayed as a promising way to overcome intrapersonal conflict. Using the example of a commitment savings product in the Philippines, I present evidence that people may fail at choosing commitment contracts which are suitable for their preferences. I argue that an individual's ability to correctly choose a welfare-improving commitment contract depends on his degree of sophistication, i.e., on the individual's awareness of the nature of his time-inconsistency. I observe that a majority of individuals who takes up a commitment product chooses very low stakes for this commitment, and then defaults on it. Both take-up and default decisions are systematically linked to low measures of sophisticated time-inconsistency, suggesting that imperfect (or partial) levels of sophistication are widespread. By the nature of commitment, a tendency to choose unsuitable contracts is costly. Implications reach beyond commitment savings, and may extend to rich country applications such as gym contracts (as shown by DellaVigna and Malmendier (2006)), diet clubs, and long-term pension savings plans.

From a policy perspective, the presented comparison between a (harder) regular-instalment commitment and a (softer) withdrawal-restriction commitment may suggest a possible trade-off between efficacy and risk of offering commitment products: Offering stronger commitments with more pressure may provide greater benefits on *average* – as observed by a fourfold effect of the ASA product on average bank savings, and an increased likelihood to purchase one's savings goal using own funds. However, offering stronger commitments may also involve an increased risk of adverse effects on welfare for partially sophisticated agents. In the present study, a 'softer' commitment contract is exemplified by the date-based Gihandom account: At the end of the savings period, individuals simply received their savings back, and 'undesirable' behaviour went unpenalized. While the absence of penalties may keep welfare risks to a minimum, beneficial effects of the product may be similarly limited: Offering the account had a comparatively small effect on average savings, and an even smaller effect on the median.

The welfare risks suggested in this study are not singular – a closer look at heterogeneity behind average treatment effects in the literature may reveal that adverse effects of commitment products are widespread. As a consequence, research on new commitment products should carefully consider possible risks to welfare, with particular view to partially sophisticated time-inconsistent agents.

References

Ali, S Nageeb, "Learning Self-Control," *The Quarterly Journal of Economics*, 2011, *126(2)*, 857–893.

Amador, Manuel, Iván Werning, and George-Marios Angeletos, "Commitment vs. Flexibility," *Econometrica*, 2006, *74(2)*, 365–396.

Ambec, Stefan and Nicolas Treich, "ROSCAs as Financial Agreements to Cope with Self-Control Problems," *Journal of Development Economics*, 2007, *82(1)*, 120–137.

Ameriks, John, Andrew Caplin, John Leahy, and Tom Tyler, "Measuring Self-Control Problems," *The American Economic Review*, 2007, *97(3)*, 966–972.

Anderson, Siwan and Jean-Marie Baland, "The Economics of Roscas and Intrahousehold Resource Allocation," *Quarterly Journal of Economics*, 2002, *117(3)*, 963–995.

Ariely, Dan and Klaus Wertenbroch, "Procrastination, Deadlines, and Performance: Self-control by Precommitment," *Psychological Science*, 2002, *13* (3), 219–224.

Ashraf, Nava, Dean Karlan, and Wesley Yin, "Deposit Collectors," *The B.E. Journal of Economic Analysis & Policy*, 2006, *5(2).*

_ , _ , and _ , "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *Quarterly Journal of Economics*, 2006, *121(2)*, 635–672.

Banerjee, Abhijit and Sendhil Mullainathan, "The Shape of Temptation: Implications for the Economic Lives of the Poor," Working Paper, National Bureau of Economic Research 2010.

Banerjee, Abhijit V. and Sendhil Mullainathan, "Limited Attention and Income Distribution," *American Economic Review*, 2008, *98(2)*, 489–493.

Basu, Karna, "Commitment Savings in Informal Banking Markets," 2012.

Bauer, Michal, Julie Chytilova, and Jonathan Morduch, "Behavioral Foundations of Microcredit: Experimental and Survey Evidence from Rural India," *American Economic Review*, 2012, *102(2)*, 1118–1139.

Benartzi, Shlomo and Richard H. Thaler, "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy*, 2004, *112(1)*, 164–187.

Besley, Timothy, "Savings, Credit and Insurance," in Jere Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics*, Vol. 3A, Elsevier, 1995.

_ , Stephen Coate, and Glenn Loury, "The Economics of Rotating Savings and Credit Associations," *American Economic Review*, 1993, *83(4)*, 792–810.

Binswanger, Hans P, "Attitudes Toward Risk: Experimental Measurement in Rural India," *American Journal of Agricultural Economics*, 1980, *62(3)*, 395–407.

Browning, Martin and Jeremy Tobacman, "Discounting and Optimism Equivalences," 2007.

Brune, Lasse, Xavier Giné, Jessica Goldberg, and Dean Yang, "Commitments to Save: A Field Experiment in Rural Malawi," Policy Research Working Paper Series 5748, World Bank 2011.

Bryan, Gharad, Dean Karlan, and Scott Nelson, "Commitment Devices," *Annual Review of Economics*, 2010, *2(1)*, 671–698.

DellaVigna, Stefano and Ulrike Malmendier, "Contract Design and Self-Control: Theory and Evidence," *The Quarterly Journal of Economics*, 2004, *119(2)*, 353–402.

_ and _ , "Paying Not To Go To The Gym," *The American Economic Review*, 2006, *96(3)*, 694–719.

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

_ **, Rachel Glennerster, and Michael Kremer**, "Using Randomization in Development Economics Research: A Toolkit," *Handbook of Development Economics*, 2007, *4*, 3895–3962.

Dupas, Pascaline and Jonathan Robinson, "Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya," *American Economic Journal: Applied Economics*, 2013, *5(1)*, 163–192.

_ **, Sarah Green, Anthony Keats, and Jonathan Robinson**, "Challenges in Banking the Rural Poor: Evidence from Kenya's Western Province," Working Paper 17851, National Bureau of Economic Research 2012.

Fischer, Greg and Maitreesh Ghatak, "Repayment Frequency in Microfinance Contracts with Present-Biased Borrowers," Economic Organisation and Public Policy Discussion Paper 21, London School of Economics 2010.

Frederick, Shane, George Loewenstein, and Ted O'donoghue, "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 2002, *40(2)*, 351–401.

Fudenberg, Drew and David K Levine, "A Dual-Self Model of Impulse Control," *The American Economic Review*, 2006, *96(5)*, 1449–1476.

Giné, Xavier, Dean Karlan, and Jonathan Zinman, "Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation," *American Economic Journal: Applied Economics*, 2010, *2(4)*, 213–235.

Giné, Xavier, Jessica Goldberg, Dan Silverman, and Dean Yang, "Revising Commitments: Field Evidence on the Adjustment of Prior Choices," Working Paper, National Bureau of Economic Research 2012.

Gugerty, Mary Kay, "You Can't Save Alone: Commitment in Rotating Savings and Credit Associations in Kenya," *Economic Development and Cultural Change*, 2007, *55(2)*, 251–282.

Gul, Faruk and Wolfgang Pesendorfer, "Temptation and Self-Control," *Econometrica*, 2001, *69(6)*, 1403–1435.

_ **and** _ , "Self-Control and the Theory of Consumption," *Econometrica*, 2004, *72(1)*, 119–158.

Hofmann, Anett, "Just A Few Cents Each Day: Can Fixed Regular Deposits Overcome Savings Constraints?," 2013.

Imbens, Guido W and Joshua D Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62(2)*, 467–475.

Johnson, Simon, John McMillan, and Christopher Woodruff, "Property Rights and Finance," *American Economic Review*, 2002, *92(5)*, 1335–1356.

Karlan, Dean and Jonathan Zinman, "Price and Control Elasticities of Demand for Savings," 2013.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, *112(2)*, 443–478.

Loewenstein, George, "Anticipation and the Valuation of Delayed Consumption," *The Economic Journal*, 1987, *97(387)*, 666–684.

Mullainathan, Sendhil, Bindu Ananth, and Dean Karlan, "Microentrepreneurs and Their Money: Three Anomalies," Working Paper, Financial Access Initiative 2007.

O'Donoghue, Ted and Matthew Rabin, "Doing It Now or Later," *American Economic Review*, 1999, *89(1)*, 103–124.

Sayman, Serdar and Ayse Öncüler, "An Investigation of Time Inconsistency," *Management Science*, 2009, *55(3)*, 470–482.

Shefrin, Hersh M and Richard H Thaler, "The Behavioral Life-Cycle Hypothesis," *Economic Inquiry*, 1988, *26(4)*, 609–643.

Sinn, Miriam, "Risk and Time-Inconsistency: Evidence from a Field Experiment in West Bengal," 2012.

Strotz, Robert Henry, "Myopia and Inconsistency in Dynamic Utility Maximization," *The Review of Economic Studies*, 1955, *23(3)*, 165–180.

Takeuchi, Kan, "Non-Parametric Test of Time Consistency: Present Bias and Future Bias," *Games and Economic Behavior*, 2011, *71(2)*, 456–478.

Tarozzi, Alessandro and Aprajit Mahajan, "Time Inconsistency, Expectations and Technology Adoption: The Case of Insecticide Treated Nets," ERID Working Paper 105, Duke University 2011.

Thaler, Richard, "Mental Accounting and Consumer Choice," *Marketing Science*, 1985, *4(3)*, 199–214.

A Appendix: Supplementary Figures and Tables

TABLE IX: PERSONAL SAVINGS GOALS

	All	All (%)	ASA clients	Gihandom clients
Education	163	21.79	18	21
General Savings/Not specified	148	19.79	37	21
House/Lot purchase/construction/repair	106	14.17	20	12
Christmas/Birthday/Fiesta/Baptism	91	12.17	12	16
Capital for Business	69	9.22	9	5
Household Item (Appliance/Furniture)	41	5.48	5	4
TV/DVD Player/Laptop/Cellphone	33	4.41	3	2
Emergency Buffer	31	4.14	1	0
Health/Medical	26	3.48	3	2
Agricultural/Lifestock	19	2.54	2	6
Motorbike/Car/Boat	17	2.27	4	2
Travel/Vacation	4	0.53	0	1
Total	748	100	114	92
Median Goal Amount (pesos)	2400		2400	2400
Median Time until Goal Date (days)	137		138	133
Median Termination Fee (pesos, if ASA)	–		150	–
Date-Based Goal (if Gihandom)	–		–	53
Amount-Based Goal (if Gihandom)	–		–	39

TABLE X: TAKE-UP RATES

	Assigned	Reached	Take-Up	Take-Up (% assigned)	Take-Up (% reached)
Regular Saver (ASA)	457	423	114	25%	27%
Withdrawal Restriction (Gihandom)	228	219	92	40%	42%
Standard Account (OSA) with P100	913	852	788	86%	92%

TABLE XI: AVERAGE TREATMENT EFFECTS - ROBUSTNESS TO UNBALANCED COVARIATES

Dependent	Change in Bank Savings		Change in Other Savings		Change in Outstanding Loans	
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Regular Saver	585.465***	588.830***	426.811	247.250	-840.258	-871.795
Treatment (ASA)	(129.251)	(132.821)	(671.844)	(688.523)	(1,180.168)	(1,197.380)
Withdrawal Restr.	148.243***	163.561***	-328.159	-613.247	-308.549	-678.111
Treatment	(40.927)	(53.661)	(705.461)	(704.891)	(1,258.139)	(1,274.365)
(Gihandom)						
Impatience		-121.374		843.994		1,635.189
		(106.121)		(536.597)		(1,039.716)
Risk Aversion		24.084		-93.407		329.584
		(39.676)		(133.289)		(230.476)
HH Income		24.678		-377.650*		-88.329
		(21.932)		(192.628)		(244.735)
Constant	27.160***	-113.780	63.451	1,377.945	1,882.729**	286.813
	(9.399)	(135.998)	(531.028)	(930.966)	(920.828)	(1,391.783)
R^2	0.02	0.03	0.00	0.03	0.00	0.01
Observations	748	745	603	600	720	715

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All coefficients measure Intent-to-Treat (ITT) effects, i.e., the effect of being offered the commitment product. Columns (1)-(2) use administrative data from the partner bank. Columns (3)-(6) use survey-based data. The dependent variable in columns (3)-(4) is the change in reported savings at home and at other banks (other than the partner bank) between the baseline survey and the personal savings goal date (3-6 months later). The dependent variable in columns (5)-(6) is the change in the reported outstanding loan balance between the baseline survey and the endline survey, 6 months later. Both loan and survey-based savings data are truncated at 1 percent.

TABLE XII: ASA TAKE-UP & DEFAULT: ROBUSTNESS

Dependent	ASA Take-Up				ASA Default (R-Sample)			
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Present Bias	0.0805	-0.0403	0.0623	0.0001	0.1084*	0.0293	0.0933	0.0765
	(0.0864)	(0.0614)	(0.0825)	(0.0684)	(0.0651)	(0.0453)	(0.0616)	(0.0511)
Soph. Present Bias	-0.0620**		-0.0570**		-0.0429*		-0.0371	
(Pres.Bias*Temptation)	(0.0294)		(0.0288)		(0.0238)		(0.0237)	
Perceived Temptation	-0.0050	-0.0139	-0.0095		-0.0212**	-0.0279***	-0.0179*	
	(0.0124)	(0.0115)	(0.0123)		(0.0105)	(0.0097)	(0.0105)	
Pres.Bias*Self-Control				-0.0392				-0.0667*
				(0.0394)				(0.0384)
Perceived Self-Control				-0.0277				-0.0221
				(0.0216)				(0.0183)
Full Controls	YES	YES	NO	YES	YES	YES	NO	YES
Marketer FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	402	402	408	402	402	402	408	402

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Entries in the table represent the marginal coefficients of the corresponding probit regression. In columns (3) and (7), the set of control variables has been limited to age and gender (all other control variables may directly or indirectly represent choice variables). Columns (2) and (6) omit the interaction term $tempted - ideal * presentbias$ ("Pres.Bias*Temptation"), which is used as a measure of sophisticated time-inconsistency (see Section 3). Columns (4) and (8) instead use $expected - ideal * presentbias$ ("Pres.Bias*Self-Control") as an alternative measure of sophistication. Note that 316 out of 402 (79%) individuals in the R-sample report zero (or in 13 cases, negative) values of Perceived Self-Control. Interacted with the measure of present bias, this implies that only 21 out of 402 values of Pres.Bias*Self-Control are non-zero. While the relationship with Take-Up is not significant (possibly due to a lack of variation), the coefficient on Pres.Bias*Self-Control is roughly comparable in magnitude and sign to the coefficient on Pres.Bias*Temptation. See Section 3 for a description and discussion of the measurement of sophistication.

TABLE XIII: QUANTILE REGRESSIONS

		(1) Change in Bank Savings	(2) Change in Other Savings	(3) Change in Outstanding Loans
10th	Regular Saver	0.00	252.00	-4,000.00*
Percentile	(ASA)	(0.00)	(2,353.66)	(2,282.56)
	Withdrawal Restr.	0.00	-148.00	-345.00
	(Gihandom)	(0.00)	(2,670.35)	(2,598.30)
20th	Regular Saver	0.00	-271.00	-2,000.00*
Percentile	(ASA)	(0.00)	(630.63)	(1,021.07)
	Withdrawal Restr.	0.00	-1,071.00	-1,000.00
	(Gihandom)	(0.00)	(715.48)	(1,162.30)
30th	Regular Saver	0.00	-150.00	-800.01**
Percentile	(ASA)	(0.00)	(261.67)	(394.72)
	Withdrawal Restr.	0.00	-240.00	-700.00
	(Gihandom)	(0.00)	(296.88)	(449.32)
40th	Regular Saver	0.00	0.00	0.00
Percentile	(ASA)	(5.45)	(53.89)	(129.39)
	Withdrawal Restr.	0.00	0.00	0.00
	(Gihandom)	(6.29)	(61.15)	(147.28)
50th	Regular Saver	0.00	0.00	0.00
Percentile	(ASA)	(5.23)	(97.89)	(41.80)
	Withdrawal Restr.	100.00***	56.67	0.00
	(Gihandom)	(6.03)	(111.06)	(47.58)
60th	Regular Saver	0.00	85.00	50.00
Percentile	(ASA)	(0.00)	(229.72)	(261.24)
	Withdrawal Restr.	100.00***	-135.00	-100.00
	(Gihandom)	(0.00)	(260.62)	(297.38)
70th	Regular Saver	0.00	110.00	-234.00
Percentile	(ASA)	(17.91)	(389.19)	(711.40)
	Withdrawal Restr.	100.00***	-343.44	-800.00
	(Gihandom)	(20.64)	(441.56)	(809.80)
80th	Regular Saver	200.00	-208.00	840.00
Percentile	(ASA)	(181.42)	(587.84)	(1,226.00)
	Withdrawal Restr.	150.00	-865.96	340.00
	(Gihandom)	(209.10)	(666.93)	(1,395.59)
90th	Regular Saver	2,051.87***	-635.00	925.00
Percentile	(ASA)	(329.68)	(1,290.76)	(3,737.72)
	Withdrawal Restr.	280.00	-1,050.00	-489.00
	(Gihandom)	(379.97)	(1,464.43)	(4,254.74)
Observations		748	603	720

Estimated standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Survey-based data (columns (2) and (3)) is truncated at 1 percent. All reported coefficients are Intent-to-Treat effects.

The effect of offering the Regular Saver (ASA) product on total bank savings (ordinary plus commitment savings accounts) is not apparent until the 90th percentile. This is consistent with a large effect on the 51 ASA clients who successfully completed their contract, and a limited effect on non-adopters. The ASA product was offered to 423 individuals, of whom 114 adopted the product. The 63 ASA clients who defaulted largely achieved a zero change in savings - a majority of defaulters stopped depositing soon after opening their account (see Figure 1), and their opening balance was consumed by the default penalty.

The effect of offering the Withdrawal Restriction (Gihandom) product on bank savings is 100 pesos at the median - this is likely the mechanical result of a 42 percent take-up rate and a 100 pesos minimum opening balance. In contrast to ASA clients, those Gihandom clients who stopped depositing after their opening balance (79 percent) did not lose their savings to a default penalty, but their savings remain frozen in their account (up to a goal date or amount, see Section 5.2).

The regressions in columns (2) and (3) are based on survey responses on individuals' outstanding loan balance, as well as on savings at home and at other banks. While there is a large amount of noise in the survey data, there is no systematic evidence of a substitution from other sources of savings into savings at the partner bank. However, offering the Regular Saver product may have facilitated the biggest reductions in loan demand (at 10th, 20th, and 30th percentile).

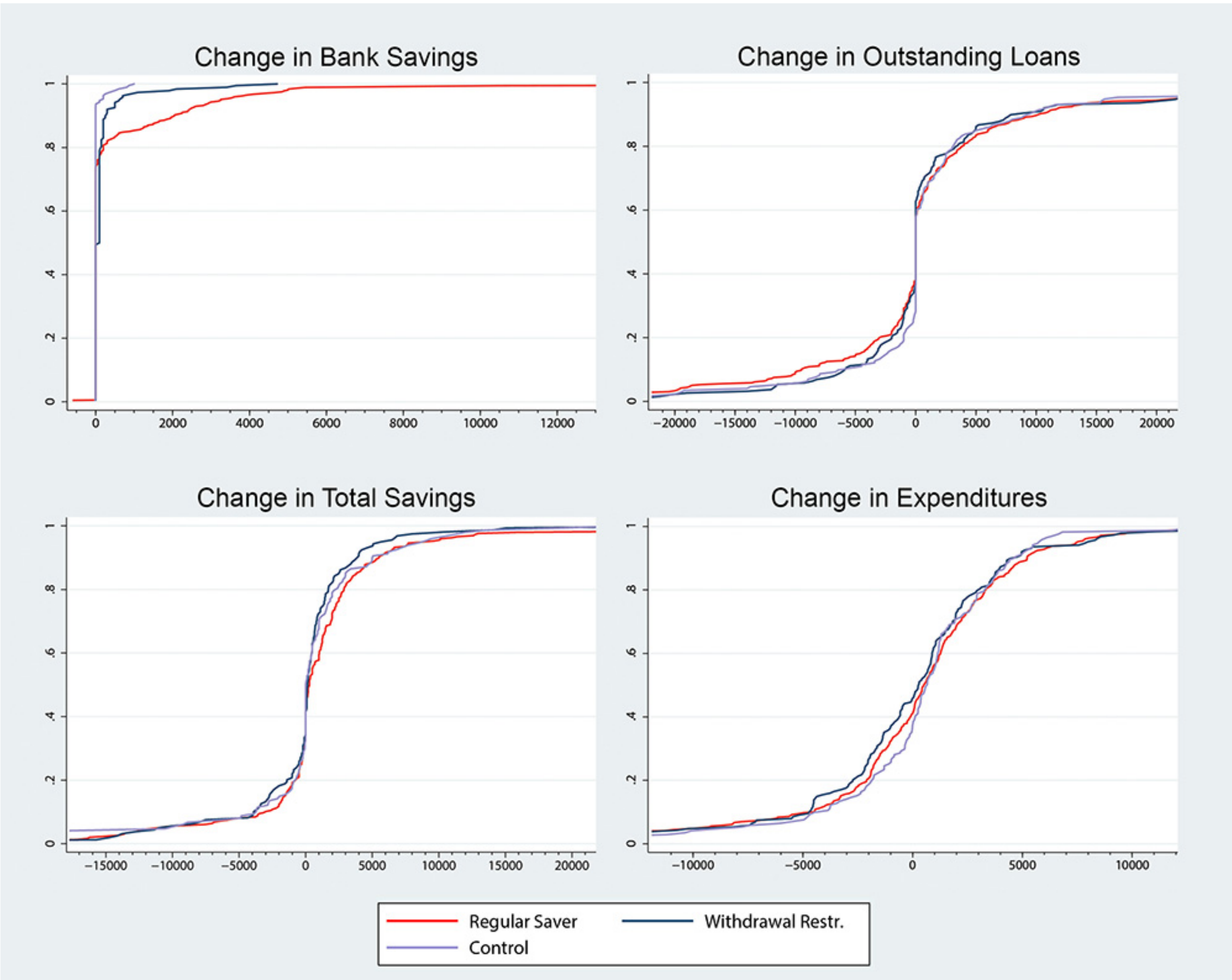


Figure 6: Distributional Effect of Treatment on Savings, Loans and Expenditures

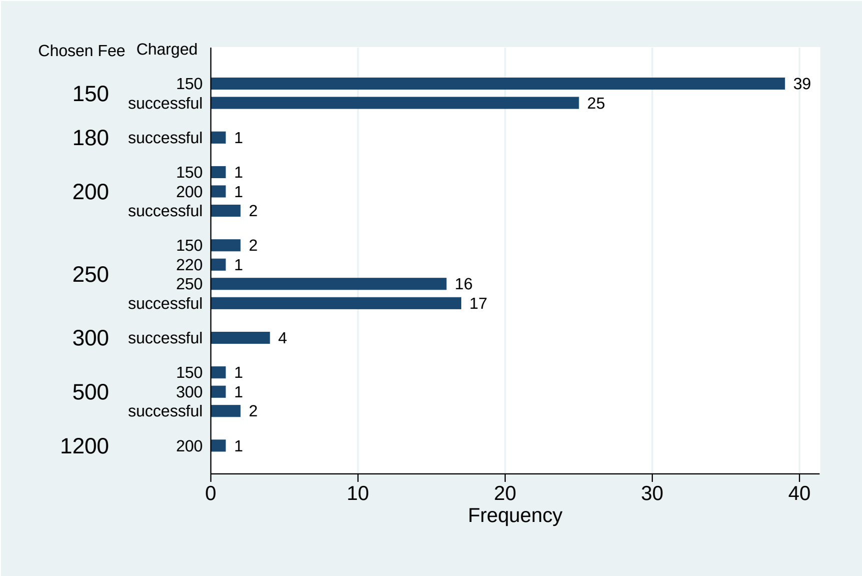


Figure 7: Termination Fees (Chosen & Charged)

B Appendix: Survey Measurement and Marketing Material

The ad-hoc randomization to determine who would receive real rewards for the time-preference questions was implemented as follows: At the start of the survey, enumerators verified respondents' ID as a part of the screening process. Enumerators then performed a calculation based on an individual's birth day, month and year. If the calculated number was odd, the respondent received a survey containing questions with real rewards. If the calculated number was even, the survey was administered with hypothetical questions.[59] Individuals were not informed about this randomisation when starting the survey, but the nature of rewards was transparent at the time of asking the questions. Serious consideration was given to the possibility of an uncertainty bias: In the presence of uncertainty about whether they would receive a promised future payment, even time-consistent agents would have an incentive to always pick the immediate reward. Choices in the future time frame would be unaffected, resulting in an upward bias on the present bias measure. To assure individuals that all payments were guaranteed, both cash and official post-dated bank cheques were presented during the game.

In addition to the measures for present bias and sophistication, the baseline survey obtained measures of other covariates of interest: A measure of the strength of financial claims from others is obtained using a methodology similar to that in Johnson et al. (2002): Individuals were presented with a hypothetical scenario in which they keep 3000 pesos in their house, set aside for a particular expenditure that is due in one month. If the people around them knew about this money, how many would ask for assistance, and how much would they ask? This hypothetical framing avoids the endogeneity inherent in asking respondents directly about actual transfers made to others (actual transfers were also observed, but not used in the analysis). The 'Financial Claims' variable used in this paper is an indicator for individuals who reported to face above-median claims from others (the median was 500 pesos, which was also the mode). Risk aversion is a score in [1, 6], and represents the individual's choice when

[59] The calculation was designed to give an odd number if the individual's birth year was odd, and even otherwise. The survey team was unaware of this connection. Given the availability of verified IDs which included birthdays, it was possible to check ex-post that the correct type of survey had been administered.

Figure 8: Test of Risk Aversion (Methodology: Binswanger (1980))

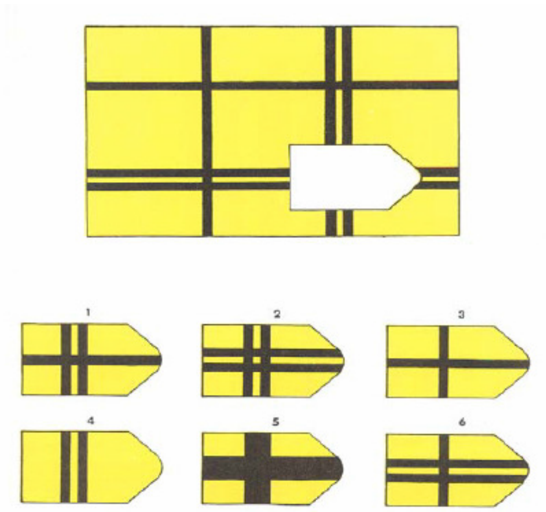


Figure 9: Illustration: Test of Cognitive Ability

Personal Savings Plan

Name: Sample

Address: Gingoog City, Mindanao

Purpose of Savings: Tuition Fees

Goal Date: 4 July

Goal Amount: 5000PhP

wk	Date Due	Deposit Due	Date of Deposit	Deposit made? (tick!)
1	22 Feb	250		
2	29 Feb	250		
3	07 Mar	250		
4	14 Mar	250		
5	21 Mar	250		
6	28 Mar	250		
7	04 Apr	250		
8	11 Apr	250		
9	18 Apr	250		
10	25 Apr	250		

wk	Date Due	Deposit Due	Date of Deposit	Deposit made? (tick!)
11	2 May	250		
12	9 May	250		
13	16 May	250		
14	23 May	250		
15	30 May	250		
16	6 Jun	250		
17	13 Jun	250		
18	20 Jun	250		
19	27 Jun	250		
20	4 Jul	250		

Figure 10: Personal Savings Plan (All Treatment Groups)

faced with a set of lottery options with increasing expected value and increasing variance (see Figure 8). Choosing the 'no-risk' lottery A yielded a score of 6, for extreme risk aversion (this option was chosen by 48 percent of the sample). Cognitive ability is proxied by the number of correct answers (out of five possible) from a culture-free intelligence test (see Figure 9 for a sample question). A financial literacy score is given by the number of correct answers (again, out of five possible) to basic numeracy questions. Household bargaining power is measured as follows: Individuals were asked who was the main decisionmaker for five types of household expenses (market purchases, durable goods, transfers to others, personal recreation, and schooling of children). For each type of expense at their discretion, their bargaining score increased by one, resulting in a measure with a range [0, 5]. 95 percent of respondents were female; thus the variable measures predominantly female bargaining power. Distance to the bank branch is measured as the linear geographic distance to the partner bank, obtained using GPS coordinates. An existing savings account indicates that the individual reported to have an existing savings or checking account at any bank (not necessarily the partner bank) at the time of the baseline survey. Finally, the shock arrival rate is proxied by the number of unexpected emergencies (such as death or illness of a household member, redundancy, natural desasters, damage to house and crops, theft, and a range of others) that a household suffered in the last 12 months before the start of the treatment ('Emergencies last yr').

C Appendix: Proofs of Propositions in Section 6

Proposition 1. *The following equation holds if s_1 is bigger than some threshold value, $s_1 \geq s_{min}$.*

$$u(1+s_1-(p-1))+\beta[(1-\lambda)u(b)+\lambda u(p-1)] \geq u(1+s_1-s_2^{No})+\beta E[u(y_3+s_2^{No})]$$

Proof. It is sufficient to prove that once s_1 is high enough to satisfy the equation above (i.e., buying the good is optimal), the equation will also be satisfied for all higher values of s_1. Consider a value s_1' such that buying the good is optimal, then

$$u(2+s_1'-p)+\beta[(1-\lambda)u(b)+\lambda u(p-1)] \geq u(1+s_1'-s_2)+\beta E[u(y_2+s_2)].$$

The equation holds for all $s_2 < p-1$, thus it also holds for $s_2^{No}(s_1'')$, the s_2 that is optimal at a higher level $s_1'' > s_1'$, conditional on the nondivisible not being bought. Due to strict concavity of $u(c_t)$,

$$u(1+s_1'-s_2^{No}(s_1''))-u(1+s_1'-(p-1)) \geq u(1+s_1''-s_2^{No}(s_1''))-u(1+s_1''-(p-1)),$$

i.e., the consumption gain $(p-1)-s_2^{No}$ from deciding not to save for the good in period 2 gives a higher utility gain when starting from the lower consumption level $1+s_1'$ than when starting from consumption level $1+s_1''$. Since

$$\beta[(1-\lambda)u(b)+\lambda u(p-1)]-\beta E[u(y_2+s_2^{No}(s_1'')] \geq u(1+s_1'-s_2^{No}(s_1''))-u(1+s_1'-(p-1))$$

holds by the optimality of buying the good at s_1', substitution and rearranging yields

$$u(1+s_1''-(p-1))+\beta[(1-\lambda)u(b)+\lambda u(p-1)] \geq u(1+s_1''-s_2^{No})+\beta E[u(y_2+s_2^{No})] \quad \text{for all } s_1'' > s_1'.$$

Therefore, when s_1 has reached some threshold s_{min}, saving for the nondivisible is optimal for all $s_1 \geq s_{min}$. $\square$

Proposition 2. *s_{min} is strictly decreasing in β. The effect of λ on s_{min} is ambiguous.*

Proof. For a given β, evaluate equation 3 at $s_1 = s_{min}$. If β is increased to $\beta' > \beta$, the inequality still holds: $u(b) > u(1+s_2)$ and $u(p-1) > u(s_2)$ for all $s_2 < p-1$ given $b > p$. Intuitively, the weight of the reward of saving increases relative to the cost. Since $u'(c) > 0$, the inequality becomes more slack, and will still be satisfied for $s_1' = s_{min} - \epsilon$. Therefore, s_{min} decreases in β.

Investigating the sign of $\delta s_{min}/\delta\lambda$, note that an increase in λ makes it less attractive to save for the nondivisible (which will not be obtained in case of a shock), increasing s_{min}. However, a stronger motive for precautionary savings on the right-hand side decreases the savings difference $(p-1) - s_2^{No}$, which decreases s_{min}. Which effect dominates is a function of $(b-p)$ and $u''(c)$. Formally, both sides of the inequality decrease in λ. As the shock hits, the right-hand side loses 1, at a consumption level $1 + s_2^{No} < p$. The left-hand side loses $(b-p)+1 > 1$, at a higher consumption level $b > 1 + s_2^{No}$. $\square$

Proposition 3. *s_{max} is strictly increasing in β.*

Proof. Evaluate inequality 4 at $s_1 = s_{max}$. For each side separately, take the derivative w.r.t. β. By the envelope condition, $\frac{dU}{d\beta} = \frac{\delta U}{\delta\beta} + \frac{\delta U}{\delta s_1^{No}}\frac{\delta s_1^{No}}{\delta\beta} = \frac{\delta U}{\delta\beta}$. For a time-inconsistent period 1 agent with $\tilde{\beta} < 1$, only s_1 is a choice variable – $\tilde{s}_2^{No}$ is inferred by backward induction, and depends on his belief $\tilde{\beta}$ (rather than on β). The resulting derivative of the left-hand side is bigger than the derivative of the right-hand side:

$$\begin{aligned}
&(1-\lambda)^2(u(2+s_1-p)+u(b)) \\
&+(1-\lambda)\lambda(u(2+s_1-p)+u(p-1)) \\
&+\lambda(u(s_1-\tilde{s}_2^{No})+E[u(y_3+\tilde{s}_2^{No})]) \\
&> E[u(y_2+s_1^{No}-\tilde{s}_2^{No})+u(y_3+\tilde{s}_2^{No})]
\end{aligned}$$

This inequality follows from inequality 4, noting that $u(1-s_{max}) < u(1-s_1^{No})$ holds by definition of s_{max}. As a result, when s_1 is held constant at s_{max}, and β is increased, the left-hand side increases more than the right-hand side does, so the original inequality is maintained and becomes more slack. The inequality will still hold for $s_1 = s_{max} + \epsilon$. Thus, s_{max} is strictly increasing in β. $\square$

Proposition 4. *s_{max} weakly decreases in the amount of naiveté, $\tilde{\beta} - \beta$.*

Proof. For a given β, an increase in $\tilde{\beta} > \beta$ is associated with a less sophisticated agent. The parameter $\tilde{\beta}$ enters the s_{max}- function through period 1's expectation of period 2's precautionary savings, $\tilde{s}_2^{No}(s_1) = argmax(u(y_2+s_1-s_2)+\tilde{\beta}E[u(y_3+s_2)]$. An increase in $\tilde{\beta}$ causes expected precautionary savings $\tilde{s}_2^{No}$ to increase. This brings savings closer to period 1's ideal: Since period 1 discounts period 2 and 3 at the same rate, he would like his future self to save more than he actually does. As $\tilde{\beta}$ increases, period 1 is more optimistic that period 2 will follow his preferences. As a result, both sides of inequality 4 increase in $\tilde{\beta}$. However, the agent is more dependent on precautionary savings if he does not save for the nondivisible good, since savings for the nondivisible act as an insurance against shocks. Thus, the left-hand side of the inequality increases less than the right-hand side, and the inequality may no longer hold at the original s_{max}. Hence, s_{max} weakly decreases in $\tilde{\beta}$. $\square$

Proposition 5. *The optimal allocation of savings from period 1's perspective, denoted $s_1 = s_{opt}$, is characterized by*

$$u'(1 - s_{opt}) = \beta[(1-\lambda)u'(2 + s_{opt} - p) + \lambda u'(s_{opt} - \tilde{s}_2^{No})(1 + \frac{\delta \tilde{s}_2^{No}}{\delta s_1} \cdot \frac{1-\tilde{\beta}}{\tilde{\beta}})].$$

Proof. Maximising expected lifetime utility from period 1 perspective, conditional on purchase of the nondivisible (i.e., on $s_2 = p - 1$), yields the following first-order condition for $s_1 = s_{opt}$:

$$\begin{aligned} u'(1 - s_{opt}) &= \beta[(1-\lambda)u'(2 + s_{opt} - p) + \lambda u'(s_{opt} - \tilde{s}_2^{No})] \\ &+ \beta\lambda \frac{\delta \tilde{s}_2^{No}}{\delta s_1}[-u'(s_{opt} - \tilde{s}_2^{No}) + Eu'(y_3 + \tilde{s}_2^{No})] \end{aligned}$$

Note that $\delta U_1/\delta \tilde{s}_2^{No} \neq 0$ given $\tilde{\beta} < 1$: Period 1 self does not expect his future self to share his preferences, thus the envelope condition does not apply for $\tilde{s}_2^{No}$. The first-order condition for s_{opt} can be simplified using the first-order condition from $\tilde{s}_2^{No}$: $\beta Eu'(y_3 + \tilde{s}_2^{No}) = u'(s_1 - \tilde{s}_2^{No})$. Substituting this into the above and simplifying yields equation 5. □

Proposition 6. *s_{opt} is strictly increasing in β, and always smaller than s_{max}.*

Proof. s_{opt} is determined by equation 5. Increasing β unambigously increases the right-hand side of the equation (note $\delta \tilde{s}_2^{No}/\delta s_1 > 0$). To clear, the marginal utility of period 1 consumption must increase, implying an increase in s_{opt}. Thus, s_{opt} increases in β. The second part of the proposition, $s_{opt} \leq s_{max}$, follows by the definition of s_{max}. □

Proposition 7. *For small λ, and in the region $s_1 \geq \bar{s}$, adopting a regular-instalment product increases the maximum the agent is willing to save, i.e., $s^B_{max} > s_{max}$. A sufficient constraint on the shock frequency is $\lambda < \frac{u'(1)}{u'(0.5)}$. In the region $s_1 < \bar{s}$, adopting the regular-instalment product unambiguously decreases s_{max}.*

Proof. In the region $s_1 \geq \bar{s}$: From inequality 7, the introduction of a penalty D will increase s_{max} whenever $\beta\lambda[u(s_1 - \tilde{s}_2^{No}) - u(s_1 - D - \tilde{s}_2^{No})] < u(1 - s_1^{No}) - u(1 - D - s_1^{No})$. To a first-order approximation, this is equivalent to $\beta\lambda u'(s_1) \cdot D < u'(1) \cdot D$, which holds whenever $\lambda < u'(1)/u'(s_1)$. Given $s_1 \geq \bar{s} > 0.5$, it is sufficient that $\lambda < u'(1)/u'(0.5)$. Therefore, inequality 7 always holds using the original $s_{max}(\beta)$, and it still holds for $s_{max}(\beta) + \epsilon$. For the special case where $D > s_1$, limited liability applies: The left-hand side stays constant as D increases, while the right-hand side decreases in D. The positive effect of D on s_{max} is reinforced. The resulting $s^B_{max}(\beta)$ will be strictly higher than $s_{max}(\beta)$ for $s_1 \geq \bar{s}$.

In the region $s_1 < \bar{s}$: The agent compares

$$\begin{aligned} u(1 - s_1 - D) &+ \beta(1-\lambda)^2(u(2 + s_1 - p) + u(b)) \\ &+ \beta(1-\lambda)\lambda(u(2 + s_1 - p) + u(p-1)) \\ &+ \beta\lambda(u(s_1 - \tilde{s}_2^{No}) + E[u(y_3 + \tilde{s}_2^{No})]) \\ &\geq u(1 - D - s_1^{No}) + \beta E[u(y_2 + s_1^{No} - \tilde{s}_2^{No}) + u(y_3 + \tilde{s}_2^{No})] \end{aligned}$$

With a strictly concave utility function, the utility loss from D when starting at consumption level $1 - s_1$ is bigger than the utility loss from D when starting at consumption level 1: $u(1 - s_1) - u(1 - s_1 - D) > u(1) - u(1 - D)$ for $s_1 > 0$. In other words, the penalty D hurts the agent more when he is saving for the nondivisible than when he is not. With the left-hand side decreasing more than the right-hand side, willingness to save will decrease, shifting the $s^B_{max}(\beta)$-curve below the original $s_{max}(\beta)$-curve for $s_1 < \bar{s}$. Further note $\tilde{s}_2^{No}$ is affected by D, but only through s_1^{No}. For s_1^{No}, the envelope condition applies. $\square$

Proposition 8. *The threshold $\hat{s}(\beta)$ weakly decreases in β. Equivalently, as β increases, a larger range $s_1 \in (\hat{s}(\beta), \bar{s})$ is strictly dominated by $\bar{s}$.*

Proof. The threshold $\hat{s}(\beta)$ is the lowest value of s_1 which satisfies

$$\begin{aligned}
&u(1-\hat{s}-D) + \beta(1-\lambda)(u(2+\hat{s}-p)) + \beta\lambda(u(\hat{s}-\tilde{s}_2^{No}(\hat{s})) + E[u(y_3+\tilde{s}_2^{No}(\hat{s}))]) \\
&\quad \leq u(1-\bar{s}) + \beta(1-\lambda)(u(2+\bar{s}-p)) + \beta\lambda(u(\bar{s}-D-\tilde{s}_2^{No}(\bar{s})) + E[u(y_3+\tilde{s}_2^{No}(\bar{s}))]).
\end{aligned}$$

By construction, $\hat{s}(\beta) < \bar{s} - D$ for all $\beta > 0$. Given $u(1-\hat{s}-D) > u(1-\bar{s})$, it must be that

$$\begin{aligned}
&\beta(1-\lambda)(u(2+\hat{s}-p)) + \beta\lambda(u(\hat{s}-\tilde{s}_2^{No}(\hat{s})) + E[u(y_3+\tilde{s}_2^{No}(\hat{s}))]) \\
&<\beta(1-\lambda)(u(2+\bar{s}-p)) + \beta\lambda(u(\bar{s}-D-\tilde{s}_2^{No}(\bar{s})) + E[u(y_3+\tilde{s}_2^{No}(\bar{s}))]).
\end{aligned}$$

This inequality will still hold for $\beta' > \beta$, and become more slack. All values of s_1 which were strictly dominated at β are also strictly dominated at β'. The dominated region $(\hat{s}(\beta), \bar{s})$ becomes weakly larger. $\square$

Proposition 9. *For a given λ, the minimum effective penalty D_{eff} weakly decreases in β.*

Proof. For a given β, $D_{eff}(\beta)$ is defined as the minimum penalty such that $s^B_{min}(\beta) \leq s^B_{max}(\beta)$. Holding the penalty D fixed, and increasing β to $\beta' > \beta$, Proposition 2 and 3 assert that $s^B_{min}(\beta') < s^B_{min}(\beta) \leq s^B_{max}(\beta) < s^B_{max}(\beta')$. Thus, penalty $D_{eff}(\beta)$ is effective for all $\beta' \geq \beta$. $\square$

Proposition 10. *The optimal Regular Saver contract for a fully sophisticated agent with $\beta < \hat{\beta}$ depends on the effect of the minimum effective penalty, $D = D_{eff}$: Where D_{eff} results in $s^B_{min}(\beta) \leq s^B_{max}(\beta) \leq \bar{s}$, the optimal contract is to choose D_{eff}. This achieves perfectly smooth equilibrium savings contributions $s_1 = \bar{s}$. Where D_{eff} results in $s^B_{max}(\beta) \geq s^B_{min}(\beta) > \bar{s}$, the optimal contract involves $D \geq D_{eff}$, and achieves equilibrium savings $s_1 \geq \bar{s}$.*

Proof. First, note that a fully sophisticated agent will never adopt a contract with $D < D_{eff}$: This results in $s^B_{max}(\beta) < s^B_{min}(\beta)$, and thus in certain default in period 1, which is dominated by not adopting the product. It then trivially follows that when D_{eff} results in $s^B_{max}(\beta) \geq s^B_{min}(\beta) > \bar{s}$, the optimal contract involves $D \geq D_{eff}$, and achieves equilibrium savings $s_1 \geq \bar{s}$.

Second, when D_{eff} results in $s^B_{min}(\beta) \leq s^B_{max}(\beta) \leq \bar{s}$, choosing D_{eff} necessarily results in equilibrium savings $\bar{s}$. To see this, recall that $s^B_{min}(\beta) = s_{min}(\beta)$ in the region $s_1 < \bar{s}$: Period 1 has already defaulted on the contract, implying the contract is no longer active in period 2. Further, by Proposition 7, $s^B_{max}(\beta) < s_{max}(\beta)$ in the region $s_1 < \bar{s}$. Therefore, starting from $\beta < \hat{\beta}$ and thus $s_{max}(\beta) < s_{min}(\beta)$, introducing a penalty will never lead to an intersection $s^B_{max} = s^B_{min}$ in the region $s_1 < \bar{s}$. The only possibility for $s^B_{min}(\beta) \leq s^B_{max}(\beta) \leq \bar{s}$ to occur is an intersection of the curves on the vertical (dominated) part of the s^B_{max}- curve, where $s^B_{min} \in [\hat{s}(\beta), \bar{s})$, and $s^B_{max} = \bar{s}$. This happens when the penalty is sufficiently high to make the agent willing to jump to $\bar{s}$. From the equilibrium savings schedule, $s_1 = \bar{s}$ if $max(s^B_{min}, s^B_{opt}) \in [\hat{s}, \bar{s})$. $\square$

Impact of Seasonality Adjusted Flexible Microcredit on Repayment and Food Consumption: Experimental Evidence from Rural Bangladesh [#]

December 9, 2013

Abu Shonchoy and Takashi Kurosaki[*]

Abstract:

The mismatch between credit repayments and income seasonality implies a challenge for microfinance institutions (MFIs) working in developing countries. For instance in northern Bangladesh, income and consumption downfalls during the lean season after the transplanting of major paddy crops are a serious threat to the household economy. Poor landless agricultural wage laborers suffer the most due to this seasonality as they face difficulty to smooth their consumption. In designing microcredit products, MFIs do not usually provide any flexibility or seasonal adjustment during the lean season, however. This is mainly because MFIs are afraid of the possibility that such flexibility might break the repayment discipline of borrowers, resulting in higher default rates. We thus conducted a randomized controlled trial in 2011–12 in northern Bangladesh to test empirically whether seasonality adjusted flexible microcredit leads to an increase in repayment problems for MFIs and whether it can increase and stabilize consumption of borrower households. Our results suggest no statistically discernible difference among the treatment arms in case of default, overdue amount, or repayment frequency. On the other hand, we find no positive impact of the repayment flexibility on immediate food consumption during the period of seasonality. After a year of initial intervention, however, we start to see positive changes in the food intake during the lean season. Our preliminary results are in favor of seasonality adjusted flexible design of microcredit.

JEL codes: G21, O16, D12.
Keywords: Microcredit, Default, Seasonality, Consumption Smoothing, Bangladesh.

[#] This is a preliminary version. Please do not quote without permission.

[*] Shonchoy: Institute of Developing Economies, IDE-JETRO; e-mail: parves.shonchoy@gmail.com. Kurosaki: Institute of Economic Research, Hitotsubashi University; e-mail: kurosaki@ier.hit-u.ac.jp.

1. Introduction

Given the current global move to fight poverty and hunger, it is important to understand the seasonal dimension of the poverty and hunger nexus, which affects the poor of developing countries regularly and repeatedly. Agriculture-dependent rural poverty could be linked to such distinct crop-cycle-based seasonality, and it becomes more severe with adverse periodic climatic conditions that could lead to poor-quality harvests or outright crop failure (Chambers et al. 1981). Moreover, inadequate access to formal credit and insurance products further traps people in chronic and inter-generational poverty—poverty that is very difficult to tackle with general public policy measures and social safety net approaches.

For example, in Bangladesh, the term "seasonality" is associated with a seasonal food deprivation phenomenon known locally as *monga*; it is mostly common in northern Bangladesh (Khandker and Mahmud 2012). Rural life in Bangladesh revolves around the agricultural cycle, which is characterized by three crop seasons that are in turn based on three categories of rice: *aus* (April to August), *aman* (July/August to November/December; traditionally the most important paddy crop), and *boro* (December/January to April). As a consequence of this cycle, two major seasonal deficits occur: one from late September to early November, and the other from late March to early May. With the widespread expansion of *boro* cultivation in recent years, the incidence of the lean period in March–May has significantly declined. However, the lean season in September–November that follows the transplantation of the *aman* crop still affects most parts of the country, and especially the northwest part of Bangladesh (Khandker and Mahmud 2012). Almost no alternative agricultural activity takes place in that period, and the nonagricultural sector cannot sufficiently absorb the seasonally unemployed labor.

During *monga*, drastic drops in employment-led income constitute the major reason behind reduced food consumption; such a phenomenon has been well documented in the literature (e.g., Rahman and Hossain 1995). Such a lack of income and alternative means for earnings limit the purchasing power of the people, and this situation cannot be mitigated with the minuscule amounts of assets and savings that poor households typically carry. Anecdotal evidence suggests that, on average, the number of meals consumed is significantly reduced during *monga*, and that the families of young and elderly members suffer the most. The absence of a functional credit market obstructs households from smoothing their consumption (Pitt and Khandker 2002). As a result, many individuals borrow from landlords or informal money lenders—both of which tend to charge very high interest rates—and they subsequently fall into a debt trap.

Given this status quo, various coping strategies have emerged among the *monga*-affected people of northern Bangladesh. Other than borrowing from informal sources that charge high interest rates, coping strategies common among them include advance sales of labor (Khandker and Mahmud 2012), the purchase of household essentials on credit, skipping meals during the lean season (Berg and Emran 2011), and seasonal migration (Shonchoy 2011). Of these coping strategies, temporary seasonal migration to urban areas appears to be a relatively practical and rational strategy, as individuals can move from rural areas to nearby urban areas or cities for a short period of time, in an attempt to earn a livelihood during the lean season. However, such a migration strategy is not suitable for everyone, due to family constraints (especially among households with female heads or disabled heads that may not be able to migrate during the lean season); additionally, credit and financing constraints, a lack of networking, and asymmetric information problems limit individuals' ability to migrate (Bryan et al. 2012).

One recent policy development in developing countries has been the emergence of microfinance institutions (MFIs) that focus on poverty alleviation. It is argued that, given access to even small amounts of credit, entrepreneurs from poor households will find opportunities to engage in viable income-generating activities (IGA)—many of which will be secondary to their primary occupations—and thus ameliorate poverty on their own. According to the microcredit Summit Campaign, as of December 2007, MFIs had 154,825,825 clients; of these, more than 100 million were women. In 2006, Mohammad Yunus and the Grameen Bank were awarded the Nobel Prize for Peace, for their contributions to poverty reduction, especially in Bangladesh. However, among academics, there is thus far no consensus on the impact of microcredit on income improvement and poverty reduction (Banerjee et al. 2009). On one hand, various studies on the impact of microcredit in developing countries have found evidence of consumption-smoothing, asset-building (Pitt and Khandker 1998), and poverty reduction (Khandker 2005). Conversely, using the same dataset of Pitt and Khandker (1998), Morduch (1999) found that the average impact of microfinance is "nonexistent."

A major drawback of the microcredit framework is its rigid loan repayment rules (Karlan and Mullainathan 2007). Nearly all loan contracts are fixed in their repayment schedules, which involve equal weekly payments, along with a high interest rate. However, MFIs work with poor rural people who most often have uncertain and infrequent incomes, and these circumstances make it very difficult for them to maintain such rigid weekly loan repayments. Especially during the lean period—when there are no jobs available in the rural agricultural sector—it can be very difficult for the poor to generate income, let alone comply with a loan repayment scheme; indeed, to say that rigid weekly repayments during the time of seasonal hardship exacerbates their misery is an understatement. It was found that during *monga*, households take extreme measures—like selling productive assets (Khandker and Mahmud 2012) or borrowing from loan sharks who charge extraordinarily high interest rates—in order to maintain a clean record of repayment and be assured access to future microcredit loans from MFIs.

Using primary data from rural households in Bangladesh, Shonchoy (2009, 2011) shows that during the lean season, access to microcredit does not increase the income levels of individuals, compared to those with no access to credit, ceteris paribus. Additionally, Shonchoy (2009, 2011) at the time of survey found no MFI that operates any well-targeted microfinance program solely dedicated to tackling seasonality issues such as *monga*. Given that seasonality in northern Bangladesh is historically well known, it is particularly puzzling to find that no leading microcredit product—save for PRIME intervention by PKSF[1]—has been designed to mitigate the effects of seasonality by providing some form of moratorium of loan repayment during *monga*.

The mismatch between credit repayments and income can create serious distortions that, for some people, deepen the debt trap, especially if they take extreme measures to repay loans on a weekly basis during the lean period. In this study, we examine whether these distortions are inevitable. If MFIs could allow some flexibility in the microcredit repayment schedules in periods of uncertain income during lean periods, this may improve the livelihood of the poor, provide them with greater flexibility and mobility, and in turn improve their capacity to repay the loan. Currently, MFIs are reluctant to relax their loan repayment rules; it seems that they fear that allowing people a moratorium on a weekly repayment scheme during the lean period may adversely affect their debt repayment discipline. It is possible that borrowers, if they are given seasonal adjustment in repayment, could become behaviorally accustomed to making lower or no repayments when those payments are nonetheless required, ultimately leading to lower recovery rates or even higher default rates.

Given this trade-off, it appears that an appropriate way of addressing these issues is the introduction of a field experiment that features a randomized controlled trial (RCT). A large number of RCT studies have been undertaken in microfinance-related research; such research covers a wide range of subjects, including the impact of microfinance (Banerjee et al. 2009), weekly versus monthly repayment (Field and Pande 2008, Field et al. 2012), group versus individual liability (Giné and Karlan 2011), random variations in meeting frequency (Feigenberg et al. 2011), and variance in a loan's term structure (Field et al. 2013), to name a few.

Despite this potential, rigorous evaluation of the impact of such seasonality adjusted flexibility in microcredit design is lacking in the literature. Among the few existing studies, Shoji (2010) evaluates the effectiveness of Bangladeshi microfinance in introducing a

[1] PRIME (Programmed Initiatives for Monga Eradication) was introduced in 2006 by PKSF (Palli Karma–Sahayak Foundation), a microcredit wholesaler and umbrella organization in Bangladesh. Under the PRIME scheme, individual nongovernment organizations (NGOs) receive credit facilities that have "flexible" terms—under which those NGOs are free to negotiate the credit amount, repayment schedule, and frequency of meetings with the beneficiary, and impose completely different sets of schemes with various borrowing groups. While this is ideal for beneficiaries to some extent, it is not easy to evaluate flexibility in terms that improve the accessibility of beneficiaries to microfinance, performance in IGA, or the livelihoods of their families.

contingent repayment system, beginning in 2002; this system allowed for the rescheduling of savings and installments for affected members during times of natural disaster. Using evidence pertaining to flooding in 2004 and based on an instrumental variable approach, Shoji found that rescheduling played the role of a safety net by substantially decreasing the probability that borrowers would skip meals in response to negative shocks; the effect was even more pronounced on the landless and women. Furthermore, if we restrict our attention to studies in the context of *monga*-related seasonal deprivation in northern Bangladesh, we find there to be a similar dearth of qualitative research. Khandker and Mahmud (2012) analyze the correlates of seasonal deprivation while focusing on social protection programs and microcredit, using nonexperimental data. In India, the neighboring country of Bangladesh, Czura et al. (2011) examine the impact of repayment flexibility by undertaking a randomized experiment with dairy farmers; they show that repayment flexibility contributed to consumption-smoothing and also enhanced demand for credit. With the exception of this study by Czura et al. (2011), we are unaware of any rigorous study on the impact of seasonality adjusted repayment flexibility in South Asia based on an RCT design.

We thus initiated RCT experiments in northern Bangladesh in early 2011. The aim of this study is to elucidate the mismatch between seasonality and the terms of microcredit, and to understand the impact of seasonality-adjusted microcredit. In our RCT design, our counterpart NGO first formed typical microfinance groups from randomly chosen villages. Borrowers were then provided with credit and began making weekly repayments after a short, two-week grace period. For a random subsample of these borrower groups, the repayment schedule was relaxed in two ways during the designated *monga* period. Under the first treatment, the borrower was temporarily given a moratorium, while under the second flexibility treatment, the repayment scheme was changed into a monthly repayment during *monga*.

We surveyed 1,440 households belonging to the borrower groups both before (baseline) and after one year of intervention (endline). We also executed a short *monga* survey during the time of *monga* in 2011 and in 2012, to understand the severity of the seasonal conditions. Making use of both survey and experimental methods, we empirically analyze the impact of the flexibility schemes on repayment and consumption. As a preview of the results, we find no statistically discernible difference among the treatment arms in case of default, overdue amount, or repayment frequency, while we find strong positive impact of the repayment flexibility on food consumption, among other seasonality-affected variables. We believe that our study contributes a new insight on the consequences of flexible microcredit that is both geographically and seasonally adjusted to help the vulnerable and lean season-affected poor cope better with periods of hardship.

The rest of the paper is organized as follows. Section 2 describes our RCT design and field surveys. Section 3 investigates the impact of the repayment flexibility on repayment behavior of borrowers, while Section 4 investigates its impact on consumption of borrower households. Section 5 concludes the chapter.

2. Experimental Design for Flexible Microcredit Trials

2.1 RCT Strategy

(1) Inflexible Microcredit as the Control

A typical Grameen-style microcredit scheme proceeds as follows (Armendariz and Morduch 2010): individuals eligible for microcredit first form a group wherein its members are expected to help each other in times of difficulty. Not all members can borrow immediately. It is usually the case that only some of them are offered credit after all members have saved a small amount of money on a regular basis; the rest of them are given credit after the first borrowers successfully repay several installments and all members have continued to save the same small amount on a regular basis. Weekly repayments begin without a long grace period. With typical Grameen-type microcredit, the first lent amount is small, and it is to be repaid in 50 weekly installments within a 12-month period.

Several rationales have been offered for this rigidly designed repayment schedule (Armendariz and Morduch 2010). The success of frequent repayment in minimizing default and delay could be attributed to the early warning mechanism, the lender's capture of information vis-à-vis the income flow of the borrower, and the borrower's commitment to save regularly. Repayment in group meetings in front of others also drives regular repayment by those borrowers who would like to maintain their reputation within the village.

Probably on account of these mechanisms, classic Grameen-type microcredit has been successful in maintaining high repayment rates.[2] However, attending weekly meetings regularly puts a high burden on the borrowers in terms of the opportunity costs of their time and financial stress (Field et al. 2012). Relaxing several of the classic Grameen-type features is thus being demanded from borrowers. Academic research has responded to this request, to identify the key element that was the most critically important in guaranteeing high repayment rates. For example, using a field experiment approach, Giné and Karlan (2011) evaluate the impact of removing group liability in the Philippines; they find there was no adverse impact on repayment, as long as public and frequent repayment systems were maintained. On the other hand, recent studies comparing weekly versus monthly installments and based on RCT designs show mixed results. In India, Field and Pande (2008) show no difference between microfinance schemes with weekly and monthly repayment frequencies, as long as repayments were made in public

[2] See Kurosaki and Khan (2012) for an exceptional case where an MFI suffered from high default rates, despite adopting a Grameen-type credit scheme. In their case, due to weak enforcement of the contingent renewal rule, strategic default prevailed among borrowers.

meetings. The same RCT also shows that the change from weekly to monthly repayment greatly reduced borrowers' financial stress (Field et al. 2012). In contrast, in Indonesia, Feigenberg et al. (2011) find that repayment performance was better when repayments were collected weekly rather than monthly.

Given this background, we adopted the following borrowing and repayment scheme as the control. Borrowers obtain credit of BDT 3,000[3] and begin repayment after a short, two-week grace period. Repayments are made in 45 installments, each of which is BDT 75 (except for the last one, which is BDT 60), implying a gross interest payment of BDT 360 that is spread throughout the borrowing period of approximately one year. Each of the weekly installments is to be repaid by the borrower at a weekly meeting. The borrower is obliged to attend the weekly meeting, even during the *monga* period. This design of a traditional or inflexible microcredit scheme is denoted as the "Control."

(2) Flexible Microcredit as the Treatment

During the *monga* period, microcredit borrowers may face difficulties in preparing the money needed for regular repayment. To facilitate the demand for repayment flexibility within this context, the treatment relaxes the repayment schedule in two ways during the *monga* period, which for this purpose is designated as September 20–December 20.

Under the first treatment, "Flexible 1," a moratorium is temporarily applied to repayments during the designated *monga* period. During that moratorium, households within the Flexible 1 groups do not pay any installment. After the *monga* period, the borrowers begin to pay BDT 100 per week, so that their total repayment amount and repayment period would be identical to those of the Control group.

As a variant of the first treatment, one-third of those treated with Flexible 1 are also given income generation activities (IGA) support. We refer to this treatment as "Flexible 1 + IGA." Under IGA support, instead of providing cash, we provide microcredit borrowers with a productive asset of their choice, within the credit amount, along with advice for utilizing the asset; no further subsidy is provided.

Under the second flexibility treatment, the repayment schedule is changed to feature three monthly installments of BDT 300 each during the designated *monga* period, instead of 12 weekly repayments of BDT 75 each. After the *monga* period, borrowers resume paying BDT 75 per week, so that their total repayment amount and repayment period would be the same as those of the control group. We refer to this treatment as "Flexible 2." This treatment arm provides less flexibility than Flexible 1 (in terms of loan repayment obligation), while it

[3] BDT 100 is equivalent to approximately JPY 99 or USD 1.22. BDT 3,000 therefore equals approximately USD 37.

provides better loan collection discipline than Flexible 1.

(3) Randomization of Treatment Arms

To preclude unequal treatment among members within a group, we randomized the four treatment statuses at the borrower-group level. Since our counterpart NGO usually forms one group in one village, our randomization took place at the village level.

Of the list of 90 villages that were under potential treatment by the counterpart NGO, we randomly selected 12 villages for "Control," 24 for "Flexible 1," 12 for "Flexible 1 + IGA," and 24 for "Flexible 2." In the randomization, we stratified the villages based on their distance from the closest bus station and the location type of the village (see the next subsection).

The reason for the larger number of villages under "Flexible 1" and "Flexible 2" than under "Flexible 1 + IGA" and "Control" was that our initial design had another experiment dimension, distinguished by the timing of when the borrower groups would be delivered the information that the repayment schedule would be relaxed. The intention was to create exogenous variation in the information structure, as implemented by Karlan and Zinman (2009) in the context of consumer credit in South Africa. However, due to delays in group formation and loan disbursement (our schedule of experiment unfortunately overlapped with the holy month of Ramadan), the exact timing of the announcement became similar across all groupings. Therefore, in analyzing the impact of our experiment, we eventually merged the two types of treatments (initially designed as "surprise" and "preannounced flexibility").

In each village, our counterpart NGO formed a borrower group known as *samity*, which comprised 20 members who satisfied the NGO's microcredit criteria and had voiced an interest in receiving microcredit. The member names were then recorded in the *samity* formation book by the loan officers. In the book, each *samity* member was assigned a number in ascending order; the members who happened to hold numbers 1–15 were to be offered credit, while those holding numbers 16–20 were kept in the group as observers. This design of randomization was not known to the *samity* members before the announcement of the treatments. This randomization thus implies the following sample distribution: there are 72 sample villages and 1,440 sample households, one-sixth or one-third of which falls into one of the four treatment arm categories; three-fourths of the sample households (1,080 households) were actual borrowers of microcredit.

2.2 Implementation of Surveys and RCT Interventions

(1) Counterpart NGO and Study Area

Our counterpart NGO is Gono Unnayan Kendra (GUK), which operates in the greater Gaibandha area, comprising five districts in northern Bangladesh: Gaibandha, Kurigram, Rangpur, Lalmonirhat, and Nilphamari. It has offices in all 32 *upazillas* (subdistricts) in Gaibandha district and five offices in the Kurigram district. Prior to this study, GUK had a limited experience in running traditional microfinance; on the other hand, it had already been a promoter of flexible microfinance in combination with its reportedly successful "asset transfer" program, which was financed by international donors. However, since its asset transfer program contains a large subsidy component, it is not clear how much of its success vis-à-vis outreach to the ultrapoor can be attributed to the flexibility in their repayment design *per se*. For instance, under one of GUK's programs, ultrapoor beneficiaries were provided with a livestock animal and required to return the offspring or an equivalent monetary value. This design also implies a much longer grace period than traditional microcredit.

In the study area, poverty is concentrated in so-called *char* areas. *Char* literally means "river island," and it is an area of land regularly formed from river bed sediment that has been eroded by the major rivers of Bangladesh. People living on *char* islands tend to be poorer and more vulnerable to various types of natural disasters (Khandeker and Mahmud 2012). For this reason, in our experiments, we distinguished the *char*, river basin, and inland areas where our target group—i.e., the poor and vulnerable—live. More concretely, in the randomization, we stratified villages based on the distance from the closest bus station, and on the village location types (*char*, river basin, or inland). The distribution of our final sample villages is shown in Table 1. Forty-five of the 72 sample villages (62.5% of the sample) were in Gaibandha district; the rest (37.5%) were in the Kurigram district. Eighteen of the 72 sample villages (25.0% of the sample) were in *char* areas, 42 villages (58.3%) were in inland areas, and the remaining 12 villages (16.7%) were in river basin areas.

(2) Schedule of Surveys and Experiments in the Field

Figure 1 shows the timeline of our surveys and experiments. In the first half of 2011, we visited Gaibandha and GUK to undertake preparatory investigations and make logistical arrangements. Following our agreement with GUK regarding the research design, village-level randomization was implemented, followed by the formation of *samity*. The baseline survey (Panel 1) of 1,440 households was executed in July–September 2011; it captured detailed information on the household roster; education; health, including the weights of the children; occupation; assets; income; migration experiences; agricultural production; nonagricultural enterprises; saving; credit; debt; *monga* coping; and the like.

[Figure 1 about here]

In the first three weeks of September 2011, microcredit of BDT 3,000 was issued to each of three-fourths of our sample households. Our initial plan was to issue the microcredit earlier. However, due to the holy month of Ramadan and the subsequent festival of Eid-ul-Fitr, the disbursement was delayed. As a result, those households who were given flexible microcredit entered the designated *monga* period before the due date of their first repayment installment. Nevertheless, GUK was able to collect monthly installments (Flexible 2) and larger weekly installments in the post-*monga* period (Flexible 1), without experiencing serious delays or nonrepayment problems. As designed, in all villages, the number of *samity* members who were issued credit was 15 (i.e., three-fourths of the *samity* members).

After the RCT experiments began, three more surveys were executed: the first *monga* survey (Panel 2) in November 2011, the follow-up survey (Panel 3) in July–August 2012, and the second *monga* survey (Panel 4) in November–December 2012. Panel 1 (the baseline survey) and Panel 3 were based on the long questionnaire, which covers all aspects of the household economy; Panel 2 and Panel 4, meanwhile, were based on the short questionnaire, which focused on how the household was coping with ongoing *monga* difficulties. Panel 1 was meant to capture the state of affairs *before* our interventions, Panel 2 describes the household economy *during* our interventions, and Panel 3 and Panel 4 were designed to collect information *after* our RCT experiments. In Panels 1 and 2, 1,440 households were surveyed. In Panels 3 and 4, 1,422 of the initial 1,440 households were resurveyed, implying an attrition rate of 1.25%.

In addition to these surveys, administrative data for all borrowers (i.e., 1,080 borrowers) were obtained from GUK. This dataset provides us with detailed and precise information on repayment behavior.

The distribution of our final sample households is shown in Table 1. Data for the full set of 1,440 household observations surveyed in Panel 1 are utilized as the baseline information. Data for the subset of 1,080 borrowers are utilized in Section 3 in which the impact of flexibility on repayment behavior is investigated. Data for the subset of 1,422 Panel-3 households are utilized in Section 4 in which the impact of flexibility on food consumption is investigated.

[Table 1 about here]

2.3 Validity of Randomization

As our randomization was implemented properly, we expect to observe no systematic difference in pre-intervention characteristics at the village level across the various treatment arms. To test this expectation, we estimated the following village-level regression model, using the baseline survey data:

$$X_v = b_0 + b_1 D_{1v} + b_2 D_{2v} + b_3 D_{3v} + u_v, \qquad (1)$$

where X_v is a pre-intervention variable for village v, D_{jv} is a dummy variable for treatment j (j = 1, 2, 3; i.e., Flexible 1, Flexible 1 + IGA, and Flexible 2, respectively), and u_v is a zero mean error term. If the null hypothesis that $b_1 = b_2 = b_3 = 0$ is not rejected, the balance test is passed.

Similarly, we expect to observe no systematic difference in pre-intervention characteristics at the household level across the four treatment arms, either.[4] To test this, we estimated the following household-level regression model, using the baseline survey data:

$$X_h = b_0 + b_1 D_{1h} + b_2 D_{2h} + b_3 D_{3h} + b_4 D_{4h} + u_h, \tag{2}$$

where X_h is a pre-intervention variable for household h, D_{jh} (j = 1, 2, 3) is a dummy variable indicating that household h was provided with flexible microcredit under treatment arm j (j = 1, 2, 3; i.e., Flexible 1, Flexible 1 + IGA, and Flexible 2, respectively), D_{4h} is a dummy for nonborrower households, and u_h is a zero mean error term. If the null hypothesis that $b_1 = b_2 = b_3 = 0$ is not rejected, the balance test is passed. If there was no selection bias in assigning borrower vs. nonborrower households within each *samity*, we expect b_4 to be zero as well. Because the randomization had been implemented at the village level and sample households were drawn using the village as the primary sampling unit, we used robust standard errors for b's clustered at the village level, in order to test the null hypotheses using equation (2).

Appendix Table 1 shows the results for village-level variables. At the village level, the distance from the closest bus station to the village, the dummy for a *char* village, and the dummy for an inland village were perfectly orthogonal to the treatment, confirming our randomization strategy. For all six variables that represent village-level public facilities (bazar, college, Hindu temple, town, bus stand,[5] and railway station), the null hypothesis that $b_1 = b_2 = b_3 = 0$ was not rejected at the 5% level. In this sense, the balance test at the village level was passed, suggesting that our randomization strategy at the village level had been implemented properly. Nevertheless, the null hypothesis was rejected at the 10% level for the case of Hindu temples, and the individual coefficient on D_{2v} was significant at the 5% level for the case of distance to the nearest town. As we had randomized the treatment status, we assessed them as having occurred by chance. As will be shown in Section 4, these nonrandom components do not affect our impact analysis; see the results of the robustness check, undertaken by controlling for these baseline village-level variables.

Appendix Table 2 shows the regression results for household-level variables using four variables characterizing the household head, six variables characterizing household members,

[4] It might be possible for a difference to occur at the household level across treatment arms, as treatments had been randomized at the village level. For example, Czura et al. (2011) state that "Differences in client characteristics are due to the fact that randomization occurred at the group level and groups form according to socioeconomic characteristics" (p.10).

[5] The "bus stand" here refers to the availability of any bus stand in the village, while the "bus station" used in our randomization strata refers to the distance from the closest bus station where medium- and long- distance bus services are available.

five variables characterizing land holdings, and five variables characterizing liquid asset ownership. All of these variables were compiled from the baseline survey data.[6] Of the 20 variables analyzed in Appendix Table 2, in only two cases (i.e., the ratio of adults in the household roster and the literacy rate of adult females) was the null hypothesis that $b_1 = b_2 = b_3 = 0$ rejected at the 5% level; in only three cases (i.e., the household size, average age of members, and the ratio of adults) was the null hypothesis that $b_1 = b_2 = b_3 = b_4 = 0$ rejected at the 5% level. If we individually assess the significance of b_1, b_2, and b_3, again, only one of them (i.e., b_3 for the ratio of adults in the household roster) was statistically significant at the 5% level. At most, the balance check only marginally failed for the four variables of the household size, average age, adult ratio, and adult female literacy rate. We can therefore safely conclude that these rejections occurred by chance and that randomization had been properly implemented. As will be shown in Section 4, the nonrandom components at the household level do not affect our impact analysis (see the robustness check undertaken by controlling for these baseline household-level variables).

2.4 Summary of the Experimental and Survey Design

This section explained the experimental design of our RCT in northern Bangladesh, which had been undertaken to examine the impact of flexible microcredit that targets the ultrapoor. After describing our experimental design, this section also compared the means of sample villages' and households' characteristics across the various treatment arms. It was found that most of the observable characteristics prior to our intervention were very similar across the treatment arms, indicating that randomization had been implemented properly. Means of the baseline survey data also showed that our sample households owned very few liquid assets (such as household appliances or livestock) and managed very small land holdings. These findings indicate that our sample households belong to the poorest section of rural Bangladesh.

3. Impact of Flexibility on Repayment Behavior

In this section, we examine repayment behavior to test whether seasonal adjustment in microcredit affects the default rate and repayment delays. Through this examination, we assess the general claims by the NGOs vis-à-vis a moratorium during *monga*.

3.1 Extent of Default and Absence in Weekly Meetings

(1) Definition and Summary Statistics of Empirical Variables

We compiled two sets of empirical variables that characterize the extent of repayment

[6] To be more precise, due to data entry problems, we used Panel 3 data for the household demography variables (age was adjusted by one year), supplemented by Panel 1 data for the 22 attrition households. For land and assets, we used Panel 1 data.

problems. Table 2 shows definitions and summary statistics of these variables.

The first set of empirical variables is based on the information on a borrower's payment due at the end of a loan cycle. The first variable, *default*, is defined as a dummy variable taking the value of 1 if the overdue amount was positive, and 0 otherwise. On average, 73% of borrowers had a positive overdue amount at the end of the loan cycle. The second variable, *overdue*, is a continuous variable for the absolute amount of delinquency. Its mean was BDT 564.56. We can convert this number into a relative number by dividing it by the total due amount (BDT 3,360). On average, the overdue amount was equivalent to 17% of their required accumulated amount at the end of the loan cycle. However, if you do the same calculation adjusted with savings (where total saving amount will be adjusted with the due amount, 6 months after the schedule loan cycle) then on average around 51% had due after the end of loan cycle with 9.52% of required accumulated amount. Therefore, although the incidence of default was frequent, the overdue amount was small in both absolute and relative terms on average.

The second set of empirical variables is based on the information on the number of weekly meetings missed by borrowers. MFIs typically impose a strict loan collection regime, where each borrower must pay weekly loan installments of equal amounts. However, in our experimental design, we instructed GUK not to impose any strict loan repayment discipline. Instead, we instructed GUK to conduct household visits each week, hold weekly meetings, and inform each borrower of the cumulative amount due. This was done, in particular, to observe the loan collection pattern and behavior of loan repayment among borrowers. In our definition, "missed weeks" considers only those cases where the borrowers did not pay at all[7] and had not earned any credit toward one or more missed weeks of payments. On average, borrowers missed payments by about 6 weeks under this definition. The average ratio of total missed weeks to the total loan collection weeks (variable *rmiss*) was 0.18, or 18%. Therefore, although the overdue amount was small on average, borrowers missed meetings quite frequently at the average rate of one in six.

As discussed in Section 2, our experimental design used as randomizations strata three distinct geographical properties: the *char*, river basin, and inland areas. Across three regions, borrowers in *char* areas had more difficulty in repayment than those in other two areas if we focus on two variables defined on the overdue amount at the end of a loan cycle (*default*, and *overdue*). As *char* households typically face greater difficulties in ensuring a regular flow of income, and they recurrently suffer on account of seasonal adversity, we expected that *char* households had more difficulty in regular repayment than other households. This expectation was met regarding the overdue amount and default status as depicted in figure 2.

[Figure 2 about here]

[7] Any partial payment would not result in missed weeks.

(2) Seasonality

One important aspect of this loan repayment analysis is understanding the impact of seasonality on total collection and weekly repayment. To examine any pattern of seasonality, Figure 3 plots monthly loan collection and missed weeks information.

Most of the underpayments occurred in the off-harvest periods (e.g., September–October and March–April). This reflects the income-smoothing problem faced by borrowers during these months. However, the drop in the repayment ratio during these months was not very large in magnitude. In contrast, months of December–January and May–June were associated with higher repayment on average. In December, overpayment was recorded on average. This seasonality pattern was found in all three regions of *char*, inland, and river basin.

To understand the discipline framework imposed by the MFIs, seasonality in the number of weekly meetings missed is informative. As shown in Figure 3, borrowers tended to miss more weekly payments as they reached the end of the loan cycle, compared to the beginning of the loan collection period. One interesting observation to note is that the ratio of missed weeks to the total monthly due weeks was lower in November–December and in May, which could be attributed to the paddy harvest cycle, as previously observed. An almost similar pattern and trend are observed for all three regions.

3.2 Impact of Flexibility on Default and Absence in Weekly Meetings

(1) Econometric Model

Since our treatment assignment was distributed randomly (see Section 2), to empirically complement our discussion of the repayment analysis, we could simply use ordinary least squares (OLS) regressions to evaluate the impact of various treatments on a number of outcomes. More precisely, we estimated:

$$Y_h = b_0 + b_1 D_{1h} + b_2 D_{2h} + b_3 D_{3h} + u_h, \quad (3)$$

where Y_h is the outcome variable for household h, D_{jh} (j = 1, 2, 3) is a dummy variable indicating that household h was provided flexible microcredit under treatment arm j (j = 1, 2, 3; i.e., Flexible 1, Flexible 1 + IGA, and Flexible 2, respectively), and u_h is a zero mean error term. Equation (3) was applied to all borrowers in the sample so that the number of observations was 1,080. Because the randomization was implemented at the village level and sample households were drawn using the village as the primary sampling unit, we used robust standard errors for b's cluster at the village level, in order to test the null hypothesis.

The coefficient b_0 indicates the repayment behavior of control borrowers who were under the traditional, inflexible microcredit scheme. If the null hypothesis $b_1 = b_2 = b_3 = 0$ is rejected, we will investigate which flexibility scheme was more effective than others by comparing the three parameters of b_1, b_2, and b_3. If the null is not rejected, the coefficient b_0 indicates the repayment behavior of all borrowers on average. Therefore, in the tables that show regression results, the estimate for the intercept is presented in the first row, which is readily interpreted as the estimate for the overall mean if all coefficients on the dummy variables are zero, for the sake of convenience.

(2) Regression Results

Table 3.1 shows the regression results for micro-credit repayment behavior of borrowers under different treatment groups, using indicator variable, namely the total overdue amount, overdue as a percentage of total due amount, overdue amount after 6 months of the loan cycle and Overdue adjusted with the savings, after six months of schedule loan cycle. For each variable, the odd column reports the basic regression results controlling for stratification, while the even column reports the results from a specification with additional control for household observables. As we can see, there is no statistically significant difference among the treatment groups compared with the control group of “traditional microcredit” and "Flexible 1" groups show relatively more favorable point estimations, compared with other groups, albeit not statistically significant.

[Table 3.1 and 3.2 about here]

To understand the repayment discipline and commitment behavior of various groups, equation (3) was re-estimated using binary indicator variable of loan discipline where the variable will take the value of one for having one missing week and zero otherwise. Table 3.2 shows that on an average, the Control group (a traditional, rigid weekly repayment scheme) had difficulties with repayment discipline framework along with “Flexible 2” and "Flexible 1 + IGA" repayment groups (although statistically marginal and weakly significant). However, if we emphasis on number of total missed weeks (column 3-4) and number of total missed weeks as a percentage of total due weeks (column 5-6), our results do not show any statistically significant differences among the groups. This seems to suggest that flexibility of full moratorium during the *monga* period did not result in reducing the repayment discipline, which is the opposite to MFIs’ fear.

Now, we would like to focus our discussion by highlighting the indicator variable *default*, a dummy variable equal to 1 if a borrower’s due payment at the end of the loan cycle is positive, and 0 otherwise. As shown in column (1) and (2) of Table 3.3, the difference among the groups was not statistically significant at all. The null hypothesis that $b_1 = b_2 = b_3 = 0$ was not rejected at the 10% level, indicating that the flexibility in our RCT did not result in higher default rate.

[Table 3.3 about here]

We found that neither the seasonality nor the spatial heterogeneity (*char*, river-basin, and inland) affected the regression results reported in column (1) and (2). The rejection of the null hypothesis that $b_1 = b_2 = b_3 = 0$ was found robust to other specifications that allow for the seasonality or the spatial heterogeneity.[8] We also looked at the same default indicator using the repayment adjustment technique used by our partner MFI - adjustment of overdue amount with weekly savings amount at the shamity, 6 months after the scheduled loan cycle, reported in column (3) and (4). These results are entirely consistent with earlier findings of column (1) and (2). In the last two columns of Table 3.3 (column 5-6), we estimated the impact of flexibility on total saving amount of the borrowing groups. It turned out that "Flexible 1" borrowers, on an average, saved significantly lesser amount than other two groups, which probably indicate that "Flexible 1" groups invested their credit on those investments that does not immediately generate enough revenue to save on weekly meetings.

One loan cycle period for experimentation was from September 2011 to July 2012. Unfortunately, during the period of May 2012, our survey area suffered from periodic flash floods, which affected almost all the geographical areas, however, the effect was much more pronounced in *char* and river basis areas. As a result, it might be possible that flood affected borrowers have been finding it difficult to maintain the repayment discipline of our partner MFI and have more delinquency amount, which resulted in no significant differences among the borrowing groups. To check whether this is the case, we estimated regression with the loan discipline and default indicators of our borrowing groups, before the occurrence of flood (upto May 2012 repayment records) in Table 3.4. These results are lastly consistent with earlier findings and we have not found any statistical differences among the groups and across the estimations.

[Table 3.4 about here]

During our experiment period, some shamity resisted the weekly repayment after the *monga* period and forced loan officer to agree with monthly repayment after the flexible period, rather than weekly repayment as designed. It turned out that "Flexible 2" groups have made systematically more resistance than any other groups (See Table 3.5), perhaps due to their behavioral adjustment with monthly repayment during the *monga.*

The conclusion from this section is thus clear. As far as the delinquency is concerned, we did not find any systematic difference among the treatment arms. In terms of loan discipline, Flexible 2 borrowers showed some statistically significant discipline problem, though weakly, among all treatment groups. Moreover, we find evidence that "Flexible 2" groups made statistically significant resistant for monthly repayment than other groups.

[8] The robustness check results are not reported here, but are available upon request.

(3) Subjective Evaluation of Flexible Microcredit by Borrowers

To understand borrowers' reactions to the current repayment flexibility experiment, and their feedback with respect to it, we executed a satisfaction survey that followed the work of Devoto et al. (2012), who asked existing clients whether they had any complaints, problems, or difficulties with the assigned treatment schedule of repayment. The survey was conducted as a part of the first *monga* survey (Panel 2) in November 2011. In the current study, if the borrower responded negatively, then we categorized such an answer as "not satisfied" in the satisfaction index, and 0 otherwise.

The regression result based on equation (3) is presented in Table 4. It clearly shows that borrowers under the Flexible 1 repayment scheme (complete moratorium of repayment during *monga*) were more likely to report positively than the typical microcredit repayment scheme (regular weekly repayment). Among the treatment arms, Flexible 1 had a higher level of satisfaction than the other groups; this finding is consistent with our hypothesis. Our conjecture is that because of this satisfaction, borrowers maintained their discipline in repayment under flexible schemes.

[Figure 4 about here]

3.3 Summary and Discussion

In this section, we empirically analyzed the repayment behavior among borrowers with access to various microcredit products assigned to them under the RCT-based field experimental framework. Using an RCT-based field experiment in northern Bangladesh, we randomly assigned seasonality-adjusted flexible microcredit and traditional rigid microcredit to various borrowing groups. Our results suggest there are no statistically discernible differences among the treatment arms in terms of default or overdue amounts, and these findings thus support the provision of a flexible microcredit design.

As mentioned in the introduction, our main motivation in introducing seasonality-adjusted flexible microcredit was to verify the rationales of the MFIs working in northern Bangladesh in not providing flexibility in loan repayment during *monga*. The reluctance of MFIs in providing flexibility or seasonal adjustments during *monga* is mainly due to their worry that the flexibility might break the borrowers' loan collection discipline so that it might increase the rate of loan default. When we introduced this experimental design, GUK, our counterpart NGO, strongly argued that the loan default rate would increase significantly in the moratorium group (Flexible 1): they thought that it would hamper loan discipline and also affect their financial behavior vis-à-vis the making of regular installment payments. Some GUK executives also said that the loan borrowers from the moratorium group might "run away" with the money. Our regression results convincingly show that this worry is baseless. Unlike the claims of MFIs in Bangladesh, we saw no statistically significant differences among the treatment arms in terms of seasonality-adjusted flexible microcredit. With the treatment arm featuring a complete moratorium of weekly repayment during *monga* (high-risk credit) and monthly repayment during *monga* (low-risk credit), we found that borrowers did not show any statistically significant pattern of delinquency or lower frequency repayment that was in line with the claims of the MFIs of i) discipline problems or ii) repayment problems. It appears that even when imposing a high level of credit risk (Flexible 1) on our counterpart MFI, GUK did not face a level of delinquency that was statistically different from the delinquency amount seen among traditional groups (the delinquency rates were 3.77% and 3.75% of the total due amount in the cases of traditional and Flexible 1 borrowing, respectively). In other words, even after allowing a moratorium during *monga*, we found that our counterpart NGO managed to regain more than 95% of its targeted amount of credit with interest, and so this can be considered a successful business microfinance model.

4. Impact of Flexibility on Household Consumption

In this section, we examine whether seasonal adjustment in microcredit affects the food consumption level of borrower households. Through this examination, we assess the welfare impact of moratorium or less frequent repayment meetings during *monga*.

4.1 Data on Household Food Consumption

For the impact analysis regarding consumption, we use microdata collected in the resurvey (Panel 3, July–August 2012) and the second *monga* survey (Panel 4, November–December 2012) of the 1,440 households that were covered in the baseline survey. We were able to resurvey 1,422 households, implying an attrition rate of 1.25%. Although this rate is low, we need to be attentive to the possibility of attrition bias, if the attrition happened in a nonrandom manner. In the third panel of Table 1, we show the distribution of resurveyed households across various treatment arms. As shown in the table, attrition occurred among households in villages under Flexible 1, Flexible 1 + IGA, and Flexible 2 treatment arms, while attrition did not occur among households in the Control villages. According to chi-squared tests, the dropout dummy and the treatment status was independent.[9] Furthermore, the village and household characteristics that were found marginally correlated with the treatment did not have any explanatory power when we regressed the dropout dummy on these variables (see Appendix

[9] We first tested the independence between the attrition dummy and five household status (Flexible 1, Flexible 1 + IGA, Flexible 2, Control, and Nonborrower). The chi-squared statistics with the degree of freedom (d.o.f.) at 4 was 4.257, whose *p*-value was 0.370. We then tested the same null excluding Control households as there was no attrition among this group. The chi-squared statistics with d.o.f. at 3 was 1.4654, whose *p*-value was 0.690.

Table 3). Therefore, we conclude that the resurvey data can be used in the impact evaluation without concerns vis-à-vis attrition bias.

Table 5 describes the qualitative measures of food consumption,[10] which we will analyze in this section. During *monga* 2011,[11] many households were not able to have three stomach-full meals each day. The average number of *num_mong1* was 2.1 meals per day; this became as low as 1.7 meals a day, if we specifically focus on the worst days during *monga* (variable named *num_mong2*). A dummy variable that takes the value of 1 if the household could afford two or three meals per day, even during the worst period, is used as a measure of food safety (denoted as *safe_mong* in the table). Using this measure, 68% of the households were food-secure during *monga* 2011. As another measure of food security, we will analyze a dummy variable for meat consumption within a month during *monga* 2011 (denoted as *meat_mong*), indicating that 76% of sample households were able to eat some meat.[12]

As shown in the middle panel of Table 5, food consumption situations recovered substantially after *monga*. The average number of stomach-full meals in a day during the normal, non-*monga* time in 2012 (*num_norm1*) was 2.9 meals a day; that number was slightly reduced to 2.1 meals a day, if we specifically focus on the worst days during the same period (*num_norm2*). Using *safe_norm*, a dummy variable that takes the value of 1 if the household could afford two or three meals per day, even during the worst period, 89% of the households were food-secure during normal non-*monga* times in 2012. For the purpose of impact analysis of food consumption during this period, we will use only *num_norm2* and *safe_norm* as dependent variables.

The last panel of Table 5 shows food consumption situations during *monga* 2012. Households again suffered from consumption irregularities as in *monga* 2011 as the average number of *monga_foodHH* again drops to 2.1 meals per day and becomes even lower as 1.3 meals a day, if we specifically focus on the worst days during *monga* 2012 (variable named *minimum_monga*). Seventy percent of the respondents reported that they suffered from food shortage during the *monga* when surveyed again in 2012, one year after the initial intervention.

Similar to the case for the repayment behavior, food consumption variables too are systematically correlated with geographical categories: *char*, inland, and river basin. Inland households had the highest mean for all of the six variables. This is as expected as households

---

[10] Quantitative information on household consumption—such as total expenditure, including the imputed value of self-produced foods—is not available in our dataset.

[11] Information on food consumption during *monga* 2011 was collected in the Panel 3 survey, which covered the entire *monga* period; this information, therefore, is not the same as that on food consumption, which was collected during *monga* 2011—i.e., in the Panel 2 survey in November 2011. The results reported in this paper remain qualitatively the same, if we use the Panel 2 survey data instead.

[12] In the questionnaire, we also asked about fish consumption. The absolute majority of sample households were able to eat fish in a month, even during *monga*. Given this lack of variation, we use meat as a measure of protein security.

living in inland areas away from rivers have better access to food markets than households living in *char* areas or areas close to rivers. Against our prior expectation, *char* households had higher means for five of the six variables than river-basin households, although the difference was small.

4.2 Impact of Flexibility on Household Food Consumption

(1) Econometric Model

Because the intervention was randomly assigned (see Section 2), we simply regressed the Panel 3 outcomes on the dummy variables for various treatments, to evaluate the impact. More precisely, we estimated:

$$Y_h = b_0 + b_1 D_{1h} + b_2 D_{2h} + b_3 D_{3h} + b_4 D_{4h} + u_h, \quad (4)$$

where Y_h is a post-intervention outcome variable for household h, D_{jh} (j = 1, 2, 3) is a dummy variable indicating that household h was provided with flexible microcredit under treatment arm j (j = 1, 2, 3; i.e., Flexible 1, Flexible 1 + IGA, and Flexible 2, respectively), D_{4h} is a dummy for nonborrower households, and u_h is a zero mean error term. If the null hypothesis that $b_1 = b_2 = b_3 = 0$ is not rejected, it is indicated that the flexibility within our RCT had no impact. If this null hypothesis is not rejected while another null hypothesis that $b_4 = 0$ is rejected, it is indicated that microcredit provision had an impact, regardless of flexibility. If the null hypothesis that $b_1 = b_2 = b_3 = 0$ is rejected, we will investigate which flexibility scheme was more effective than others by comparing the three parameters of b_1, b_2, and b_3. Because the randomization was implemented at the village level and sample households were drawn using the village as a primary sampling unit, we use robust standard errors for b's clustered at the village level, in order to test the null hypotheses.

Although randomization is likely to result in the treatment and control households being similar across all variables in expectation, within any particular sample, there can be small baseline differences (see Appendix Tables 1-2). To address this issue, we added to equation (4) a control for baseline variables that were associated with significant differences across treatment arms. We will report on this as a robustness check. Other specifications using changes in outcomes between Panels 3 and 1 as dependent variables are left for future research.

As other robustness checks, we estimated two further models. In the first one, the last term in equation (4), $b_4 D_{4h}$, was allowed to have various slopes, depending on the village-level treatment type. If there existed spillover effects from borrowers to nonborrower households within a *samity*, and the spillover effects were systematically different, depending on the treatment arm assigned to the *samity*, nonborrower households could be heterogeneous across the village-level treatment arms. The extended model can accommodate this possibility. Second, we dropped the last term in equation (4), $b_4 D_{4h}$, and estimated the contracted model using only data on borrower households.

(2) Expected Signs of Parameter Estimates

To examine the impact of repayment flexibility on food consumption, we estimated equation (4) using each the variables listed in Table 5, first tow panels (*num_mong1*, *num_mong2*, *safe_mong*, *num_norm2*, and *safe_norm*) as dependent variables. As stated previously, the variable *num_norm1* in Table 5 was not analyzed, due to a lack of variation therein.

Theoretically speaking, the impact of repayment flexibility on food consumption is indirect. The flexibility does not directly affect the ways in which households choose consumption. On the other hand, it indirectly affects consumption through income, price, and credit constraint effects.

We begin the discussion of the likely sign of b_4. We expect it to be negative, i.e., we expect that the provision of microcredit increases food consumption. The first channel is the income effect. If microcredit enhances permanent household income by allowing households to allocate resources more efficiently, the resulting increase in income should be reflected in higher levels of food consumption. This route should apply to each of the six dependent variables. The second channel is the price effect. If microcredit enhances the productivity of self-employment businesses and there is imperfection in labor markets, the shadow price of family labor should increase, which is in turn likely to lead to the allocation of more household resources to food (as the major input to human capital). However, it is also possible that an increase in shadow wage could work in the *opposite* direction regarding food consumption demand. The net impact can be either positive or negative theoretically, but in either case, the absolute value of the net impact is not likely to be large. The third channel is the credit constraint effect. By definition, the provision of microcredit to a household enhances its ability to smooth resource allocation across time. Since *monga* suffering is anticipated by households, it is possible that reducing food consumption during *monga* is a symptom of a binding liquidity constraint. If this is the case, we expect b_4 to be more negative when the dependent variables are food consumption during *monga* than during the normal time following *monga*.

If the flexibility arrangements examined in our experiments have similar magnitudes of income, price, and credit effects, we expect each of b_1, b_2, and b_3 to be zero. Alternatively, if Flexible 1 + IGA makes it more likely for borrower households to engage in self-employment businesses that yield immediate gains, the income and price effects are likely to be larger for this treatment than for others. If this is the case, we expect b_2 to be positive and larger than each of b_1 and b_3. Regarding the liquidity effect, we expect Flexible 1 and Flexible 1 + IGA to have additional gains over Flexible 2, and Flexible 2 to have additional gains over Control. This is because the repayment moratorium gives households greater freedom to allocate money across 60 days of *monga* than can the inflexible, traditional microcredit scheme; similarly, monthly repayments give households more freedom to allocate money across 30 days in a month during *monga* than can traditional microcredit. If this is the case, we expect $b_1 = b_2 > b_3 > 0$.

(3) Regression Results using Panel 3 Data

The results using Panel 3 data regarding the impact of our RCT on food consumption are reported in Table 6. Regarding food consumption during *monga* 2011 (columns (1)–(4), Table 6), the null hypothesis that $b_1 = b_2 = b_3 = 0$ was not rejected at the 10% level for all four consumption variables. This indicates that the flexibility in our RCT had no impact on household-level food consumption behavior during *monga* 2011. Looking at individual parameters, none of them is statistically significant if we use the traditional cut-off threshold at the 5% level. In the equation for *num_mong1* (number of stomach-full meals per day during *monga*), parameter b_2 (the impact of Flexible 1 + IGA) is positive and statistically significant at the 10% level. The estimated parameter suggests that such borrowers were 12.4 percentage points more likely to have one more stomach-full meal.

Parameter b_4 was estimated with a negative sign (as expected) in three of the four equations, but its absolute value was small; it was also statistically insignificant in all four equations if we use the traditional cut-off threshold at the 5%. The results regarding the impact of our RCT on food consumption during normal, non-*monga* times in 2012 are reported in columns (5)–(6), Table 6. When the number of minimum stomach-full meals per day during these normal times (*num_norm2*) was used as the dependent variable, all coefficients on the four dummy variables were small in terms of absolute values, and the null hypothesis that $b_1 = b_2 = b_3 = b_4 = 0$ was not rejected at the 20% level. On the other hand, when the same variable was transformed as a dummy for food safety during normal times (*safe_norm*), b_3, the impact of Flexible 2, was negative and significant at the 5% level. The estimated parameter suggests that such borrowers were 9.2 percentage points less likely to be food-secure (versus the sample average of 89%). This marks a welfare loss associated with flexible microcredit, which remains a puzzle. When the dependent variable is *safe_norm*, the point estimate for b_4 is –0.082 and statistically significant at the 10% level. The estimated parameter indicates that nonborrowers were 8 percentage points more likely to be food-insecure (versus the sample average of 89%). This evidence weakly supports the favorable impact of credit provision in enhancing consumption.

The results reported in Table 6 were robustly found from other specifications.[13] We tried (i) extending model (4) with baseline village and household attributes as additional explanatory variables, (ii) extending the last term in equation (4), b_4D_{4h}, to have different slopes depending on the treatment arms, (iii) re-estimating equation (4) without the last term, while using only borrower households, and (iv) using the limited dependent variable models, considering the truncation or integer nature of the dependent variables. The robustness check results from extension (i) are reported in Table 7. From the nine village-level variables analyzed in the balance check in Appendix Table 1, the distance to the nearest town and the distance to a Hindu temple were included as village-level controls, since these two variables were associated with a marginal failure of the balance check. Similarly, from the 20 household-level variables analyzed in Appendix Table 2, the household size, average age of members, ratio of adults, and the literacy rate of adult females were included as household-level controls, since they were associated with a marginal failure of the balance check. The addition of these six controls did not alter the coefficients substantially and test results regarding the four parameters of interest: b_1, b_2, b_3, and b_4. One small change was that the three coefficients that were significant at the 10% level in Table 6 became significant at the 20% level only.

(4) Regression Results using Panel 4 Data

Regression results using Panel 4 data turn out to be more favorable to the flexible microcredit scheme. In other words, after a year of initial intervention, we start to see positive changes in the food intake during the lean season. Tables 8-9 shows the regression results regarding food consumption during *monga* 2012. In these two tables, we report two sets of regression results for each of the six variables listed in Table 5 (last panel) as the dependent variable. For each variable, the odd column reports the basic regression results controlling for stratification, while the even column reports the results from a specification with additional control for household observables.

Overall, the impact of microcredit has been positive on food consumption. Column (1) and (2) of Table 8 show that the number of stomach-full meal consumption during the *monga* 2012 has increased for all the treatment groups, however, the impact is much pronounced for Flexible 1 + IGA group and Flexible 2 groups compared with other treatment groups. Similarly, column (3) and (4) of Table 8 indicate strong positive impacts of Flexible 1 + IGA group as this group consumed about 38 percent more stomach-full meal on the worst day of *monga* in 2012 followed by Flexible 2 group (18 percent). In column (5) and (6) of Table 8, the estimated parameter suggests that Flexible 1 + IGA and Flexible 2 borrowers were about 39 and 19 percentage points less likely to be food-insecure, respectively, versus the sample average of 70

---

[13] The robustness check results are not reported here, but are available upon request.

percentage point.

[Table 8 and 9 about here]

We further inquired about this patter on improvement of food consumption to know more about the improvement in overall protein consumption in the *monga* period of 2012. To explore it, we collected information on protein intake - in terms of fish, meat (chicken, lamb or beef) and eggs consumption in dinner during a typical week in *monga*. The result of such estimations are depicted in Table 9 where we can clearly see that among all the category of protein, fish consumption among the treatment groups (column 3 and 4) has substantially increased where Flexible 1 + IGA borrowers have significantly improved their fish consumption during the *monga* 2012 followed by other two treatment groups. Regarding egg consumption (column 6, Table 9) , parameter b_4 is negative (marginally significant at the 10% level), indicating that any type of microcredit is useful in securing protein consumption. This confirms the finding in Tables 6-7 regarding the meat consumption during *monga* 2011.

From these sets of results, it appears that after a year of initial intervention, some positive changes in the food intake during the lean season is happening among the treatment groups. We also see that Flexible 1 + IGA borrowers have been more successful in using the credit to translating into better welfare outcomes, when captured through food consumption behavior during the lean season. This could be consistent with our previous finding of no or slightly adverse impact of Flexible + IGA on food consumption during *monga* 2011. It took time for microenterprise income to increase under Flexible + IGA. Flexible 2 borrowers also showed some indication for improvements in food consumption, however these improvements are rather marginal and in most cases weekly significant. However, Flexible 1 groups hardly showed any improvements in food consumption compared with control group. One explanation could be that most of the Flexible 1 group members spend their credit on ambitious business projects due to no immediate pressure of repayment. These business probably do not generate enough income within a short-run. As a result, during the *monga* period of 2012, we do not see any significant impact of flexible 1 groups on food consumption.

4.3 Summary and Discussion

This section empirically assessed whether a flexible repayment design for microcredit could enhance food consumption among the ultrapoor. We used two rounds of cross-sectional datasets collected in 2012, after an RCT was implemented in 2011−12 in northern Bangladesh. We found repayment flexibility to have no immediate positive impact on food consumption during the intervention as well as during the normal period after the intervention. During this period, all microcredit borrowers tended to have more secured food consumption than nonborrowers, although the difference was marginal. After a year of initial intervention, the impact started to become larger and the difference started to appear among the treatment groups compared with control groups.

In the context of the current study, we could suggest several possible explanations for the insignificance of the flexibility impact until the mid 2012. First, if the main route through which the provision of microcredit enhances consumption is the reduction of liquidity constraints, our finding is consistent with the view that the main problem for the ultrapoor is consumption-smoothing between the *monga* and non-*monga* seasons, as they were already able to smooth consumption within each season in the absence of microcredit. If this is the case—and both income and price effects are negligible—there should be no difference across microcredit types, but nonborrowers' consumption should be smaller than that of borrowers. Our empirical results broadly support this pattern. The unexpected negative coefficient of the impact of a repayment moratorium with IGA support in the regression for meat consumption during *monga* is consistent with this view as well: the borrowers under this scheme experienced difficulty in smoothing resources between the future and the current *monga* period, and they were compelled to spend more on their IGA. Second, the insignificance of the repayment flexibility impact could be due to the insignificant difference in income changes across the four credit schemes studied. This was likely when the borrowed money was invested in a business that did not generate immediate income gains.

Our finding of finding positive impact using the Round 4 data, especially among Flexible 1 + IGA borrowers, confirm the above speculation. Under the context of this study where the poor borrowers do not have sufficient entrepreneurship ability, Flexible + IGA has higher potential in raising the income level of borrowers. At the same time, Flexible + IGA also requires initial sacrifice of resources to be invested in such enterprise. During *monga* 2012, however, higher income generated from IGA enabled the borrowers to increase their consumption relatively more than the other borrowers. Furthermore, during *monga* 2012, protein consumption was more secure among microcredit borrowers of any time than nonborrowers, as found during *monga* 2011. This indicates the continuation of the difficulty faced by the ultrapoor in *Char* areas in smoothing consumption between the *monga* and non-*monga* seasons.

5. Conclusion

In this paper, we empirically examine whether flexible microcredit leads to an increase in repayment problems for MFIs and whether it can increase and stabilize consumption of borrower households. The empirical analysis is based on data collected through a randomized controlled trial in 2011-12 in northern Bangladesh. Our results suggest no statistically discernible difference among the treatment arms in case of default and overdue amount. This is in favor of flexible design of microcredit. However, in terms of loan discipline, we find weak evidence of discipline problem for "Flexible 2" borrowers and evidence that "Flexible 2" groups made statistically significant resistant for monthly repayment than other groups after the designated *monga* period for repayment adjustment. This is an important lesson for designing seasonality adjusted flexible microcredit and perhaps it is better to implement similar repayment pattern throughout the loan cycle with flexibility during the period of seasonality, rather than mixing different repayment patter during the course of the loan cycle.

On the other hand, we find that it took time for such seasonality adjusted flexible microcredit to have an impact on food consumption. There was no positive impact of the repayment flexibility on immediate food consumption during the period of seasonality. After a year of initial intervention, however, we start to see positive changes in the food intake during the lean season. All microcredit borrowers tended to have more secure food consumption than nonborrowers. This could be due to the possibility that the main problem for the ultrapoor is consumption smoothing between the lean and non-lean seasons and the length of time required for income changes induced by credit schemes to realize. The findings of this study will help MFIs optimize their credit schemes; they could also help other interested parties, including governmental institutions, advocate a relaxation of microcredit rules, or search for alternative policy instruments.

References

Armendariz, B. and J. Morduch (2010). *The Economics of Microfinance*. Second edition. Cambridge, MA: MIT Press.

Banerjee, A. V. and E. Duflo (2011). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*, Public Affairs.

Banerjee, A., E. Duflo, R. Glennerster and C. Kinnan (2009). "The Miracle of Microfinance? Evidence from a Randomized Evaluation." J-PAL Working Paper.

Berg, C. and M. S. Emran (2011). "Does Microfinance Help the Ultrapoor Cope with Seasonal Shocks? Evidence from Seasonal Famine (*Monga*) in Bangladesh." SSRN eLibrary.

Bryan, G., S. Chowdhury and A. Mobarak (2012). "Seasonal Migration and Risk Aversion." CEPR Discussion Paper No. DP8739. Available at SSRN: http://ssrn.com/abstract=1988671.

Chambers, R., R. Longhurst, and A. Pacey (1981). *Seasonal Dimensions to Rural Poverty*. London: Frances Pinter.

Czura, K., D. Karlan and S. Mullainathan (2011). "Does Flexibility in Microfinance Pay Off? Evidence from a Randomized Evaluation in Rural India." Working Paper, Goethe University Frankfurt, October 2011.

Devoto, F., E. Duflo, P. Dupas, W. Parienté, and V. Pons (2012). "Happiness on Tap: Piped Water Adoption in Urban Morocco." *American Economic Journal: Economic Policy* 4(4): 68–99.

Feigenberg, B., E. Field, and R. Pande (2011). "The Economic Returns to Social Interaction: Experimental Evidence from Microfinance." Harvard University.

Field, E. and R. Pande (2008). "Repayment Frequency and Default in Microfinance: Evidence from India*." Journal of the European Economic Association* 6(2–3): 501–509.

Field, E., R. Pande, J. Papp, and Y. J. Park (2012). "Repayment Flexibility Can Reduce Financial Stress: A Randomized Control Trial with Microfinance Clients in India." *PLoS One* 7(9): 1-7. e45679.

Field, E., R. Pande, J. Papp, and N. Rigol (2013). "Does the Classic Microfinance Model Discourage Entrepreneurship Among the Poor? Experimental Evidence from India." *American Economic Review* 103(6): 2196–2226.

Giné, X. and D. Karlan (2011). "Group versus Individual Liability: Short and Long Term Evidence from Philippine Microcredit Lending Groups." Economic Growth Center, Yale University, New Haven, CT, USA.

Karlan, D. and J. Zinman (2009). "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment." *Econometrica* 77(6):

1993–2008.

Karlan, D. and S. Mullainathan (2007). “Is Microfinance Too Rigid?” VoxEU December 2007.

Khandker, S. R. and W. Mahmud (2012). *Seasonal Hunger and Public Policies: Evidence from Northwest Bangladesh*. World Bank Publications, Washington DC, USA.

Khandker, S. R. (2005). “Microfinance and Poverty: Evidence Using Panel Data from Bangladesh.” *World Bank Economic Review* 19(2): 263–86.

Khandker, S. R. (2012). “Seasonality of Income and Poverty in Bangladesh.” *Journal of Development Economics* 97: 244–56.

Kurosaki, T. and H. U. Khan (2012). “Vulnerability of Microfinance to Strategic Default and Covariate Shocks: Evidence from Pakistan,” *The Developing Economies*, 50(2): 81–115.

Morduch, J. (1999). “The Microfinance Promise.” *Journal of Economic Literature* 37(4): 1569–1614.

Pitt, M. M. and S. R. Khandker (1998). “The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter? *Journal of Political Economy* 106(5): 958–96.

Pitt, M. M. and S. R. Khandker (2002). “Credit Programmes for the Poor and Seasonality in Rural Bangladesh.” *Journal of Development Studies* 39(2): 1-24.

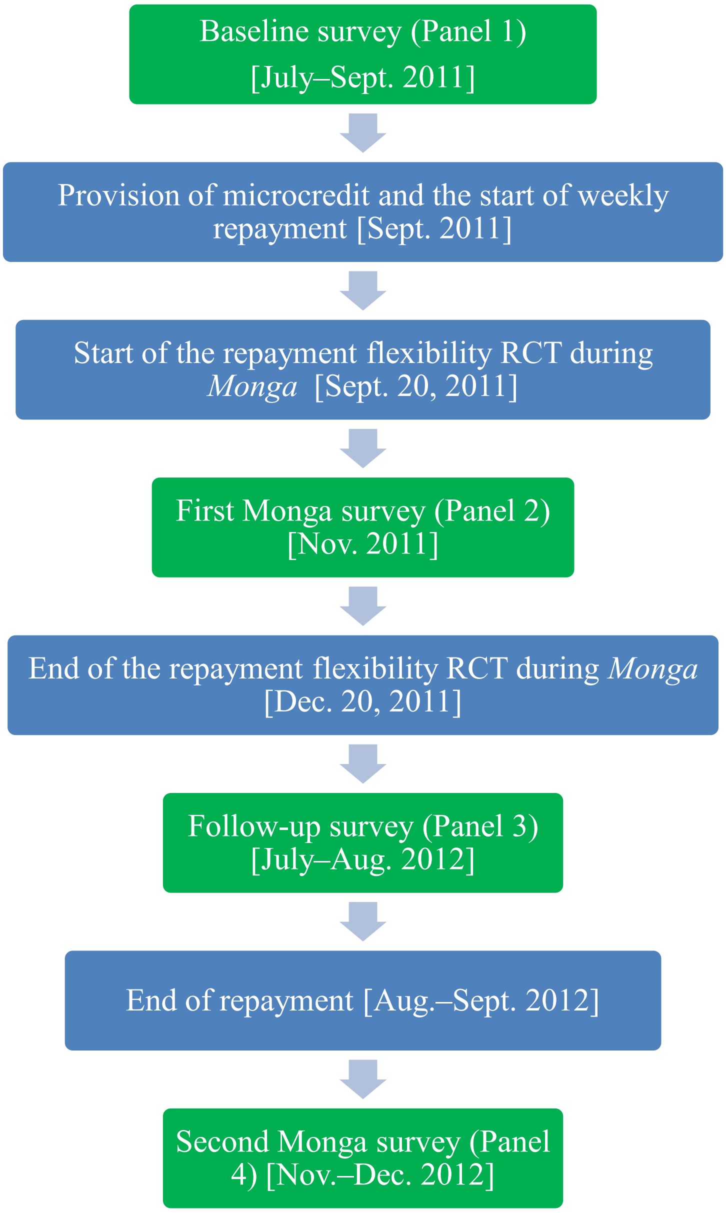
Rahman, H. Z. (1995). “Mora Kartik: Seasonal Deficits and the Vulnerability of the Rural Poor.” In H. Z. Rahman and M. Hossain (Eds.), *Rethinking Rural Poverty: Bangladesh as a Case Study*. New Dehli, India: Sage Publications.

Shoji, M. (2010). “Does Contingent Repayment in Microfinance Help the Poor During Natural Disasters?” *Journal of Development Studies* 46(2): 191–210.

Shonchoy, A. (2009). “Essays in Development Economics.” Unpublished PhD dissertation, University of New South Wales.

Shonchoy, A. (2011). “Seasonal Migration and Microcredit in the Lean Period: Evidence from Northwest Bangladesh.” IDE Discussion Paper No. 294, Japan.

Figure 1: Timeline of Interventions and Surveys



Source: Prepared by the authors. The blue panels show events regarding interventions, and the green panels show events regarding surveys.

Figure 2: Difference in repayment pattern across three geographic areas

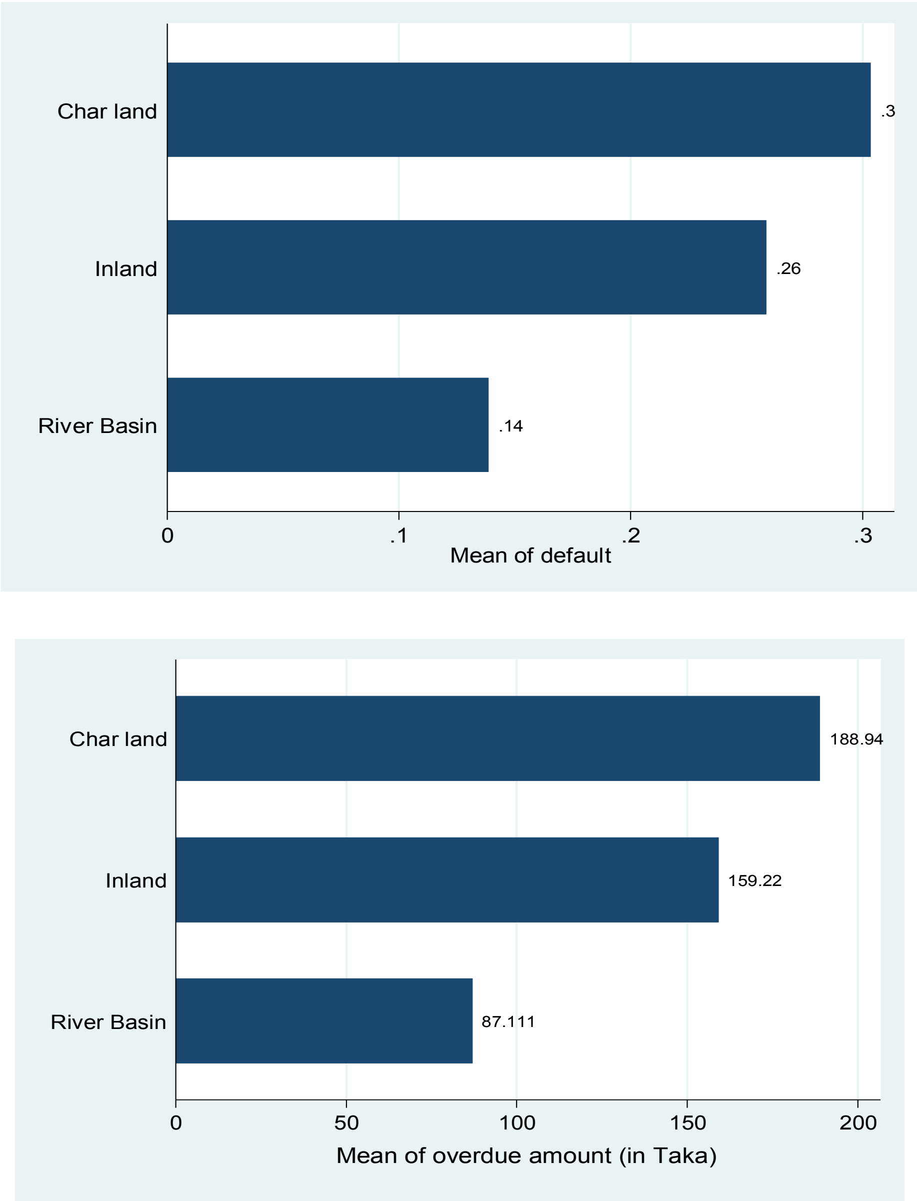


Figure 3: Seasonality of Repayment Behavior

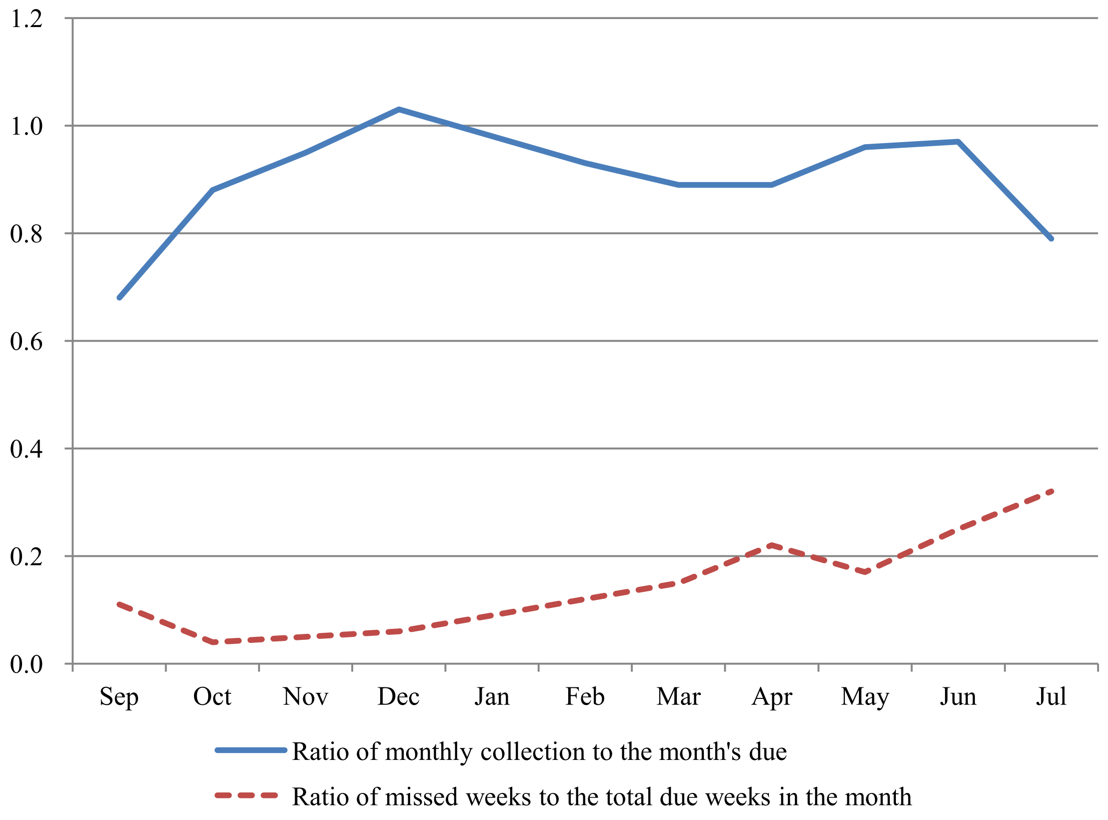


Table 1: Distribution of Sample Villages and Households by Treatment Type, Northern Bangladesh, 2011-12

	Treatment Allocation at the Village Level				Total
	Control	Flexible 1	Flexible 1 + IGA	Flexible 2	
Number of villages	12	24	12	24	72
By district					
Gaibandha District	9	16	8	12	45
Kurigram District	3	8	4	12	27
By location type					
Char	3	6	3	6	18
Inland	7	14	7	14	42
River-basin	2	4	2	4	12
Number of households in the benchmark survey, 2011 (Panel 1)					
Borrower	180	360	180	360	1,080
Nonborrower	60	120	60	120	360
Total	240	480	240	480	1,440
Number of households in the resurvey, 2012 (Panel 3 and 4)					
Borrower	180	356	176	356	1,068
Nonborrower	60	117	59	118	354
Total	240	473	235	474	1,422

Source: Compiled by the authors.

Table 2: Definitions and Summary Statistics of Variables Related with Repayment Behavior, Northern Bangladesh, 2011-12

Variable	Definition	N	Mean	Std. Dev.	Min	Max
Overdue at the end of a loan cycle						
default	Dummy for default (1 if the due amount is positive at the end of a loan cycle)	1,080	0.73	0.444	0	1.00
overdue	Due amount at the end of a loan cycle (in BDT)	1,080	564.56	618.2	0	3285.00
loverdue	Log of Due amount at the end of a loan cycle (in BDT)	1,080	4.66	2.9	0	8.10
roverdue	Due as a percentage of Total due amount	1,080	0.17	0.2	0	0.98
due_ledger_closing	Due amount after 6 months of the loan cycle	1,080	159.64	501.9	0	3210.00
ledgerclosing_adjusted	Due adjusted with the savings, after six months of schedule loan cycle	1,080	114.81	423.4	0	2990.00
Overdue before the flood in June						
cummaydue	Cummulative due upto June	1,080	276.25	397.0	0	2550.00
lcummaydue	Log of cummulative due upto June	1,080	3.85	2.7	0	7.85
Number of weekly repayments missed						
weeksdeafult	Total number of missed weeks	1,080	5.92	6.601	0	41.00
lmiss	log of total number of missed weeks	1,080	1.53	0.924	0	3.74
evermiss	Dummy variable weeks default (1 if atleast missed one week of the required due)	1,080	0.86	0.345	0	1.00
rmiss	Number of missed weeks as a percentage of total due weeks	1,080	0.18	0.185	0	1.00
Number of weekly repayments missed before the flood in June						
cummaymiss	Cumulative total number of missed weeks upto June	1,080	3.13	4.530	0	34.00
lmissmay	log of cumulative total number of missed weeks upto June	1,080	0.99	0.880	0	3.56
evermiss_may	Dummy variable weeks default (1 if atleast missed one week of the required due)	1,080	0.69	0.462	0	1.00
rmissmay	Number of missed weeks as a percentage of total due weeks, upto June	1,080	0.13	0.180	0	1.00

Note: Mean and standard deviations are simple ones, without weighting.

Source: Compiled by the authors using the administrative information for borrowers.

Table 3.1: Impact of Flexible Microcredit on Repayment Behavior (Indicators for overdue amount)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Overdue amount at the end of a loan cycle (in BDT)	Overdue amount at the end of a loan cycle (in BDT)	Overdue as a percentage of Total due amount	Overdue as a percentage of Total due amount	Overdue amount after 6 months of the loan cycle	Overdue amount after 6 months of the loan cycle	Overdue adjusted with the savings, after six months of schedule loan cycle	Overdue adjusted with the savings, after six months of schedule loan cycle
Constant	430.369***	910.764***	0.128***	0.271***	219.603*	465.278*	143.275	345.604*
	(153.84)	(270.88)	(0.05)	(0.08)	(126.16)	(237.14)	(109.17)	(202.09)
Flexible 1	15.705	25.782	0.005	0.008	-88.487	-91.273	-64.336	-65.684
	(163.39)	(164.02)	(0.05)	(0.05)	(128.85)	(129.24)	(112.41)	(113.71)
Flexible 1+ IGA	93.745	96.939	0.028	0.029	46.983	42.161	38.854	36.855
	(208.09)	(206.99)	(0.06)	(0.06)	(175.42)	(173.45)	(146.79)	(145.78)
Flexible 2	133.988	139.503	0.04	0.042	-11.626	-16.803	0.199	-2.893
	(181.73)	(181.08)	(0.05)	(0.05)	(129.20)	(128.68)	(110.33)	(111.26)
Mean of the Dependent Variable	546.86	546.86	0.16	0.16	141.75	141.75	99.32	99.32
Observations	1068	1068	1068	1068	1068	1068	1068	1068
R-squared	0.039	0.027	0.039	0.055	0.066	0.044	0.052	0.027
Log Likelihood	-8310	355.4	361.8	-8030	-8024	-7846	-7841	-8316
Control of Stratification	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control for Load Distribution Dates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control for Household Characterestics	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the village level are shown in brackets. Significant at the 10% (*), 5% (**), and 1% (***). Household Chracterestics include age, age squared, sex, marital status, education qualification of the household head as well as size of the household, number of adults and childrens in the households and total land holdings. We dropped those ovservation that are attrited.

Source: Estimated by the authors using the microdata described in the text.

Table 3.2: Impact of Flexible Microcredit on Repayment Behavior (Repayment Discipline Indicators)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Indicator for having atleast one missed week	Indicator for having atleast one missed week	No. of missed weeks	No. of missed weeks	Missed weeks as a percentage of total due weeks	Missed weeks as a percentage of total due weeks
Constant	0.713***	0.685***	6.372***	10.090***	0.139***	0.231***
	(0.09)	(0.15)	(2.13)	(2.87)	(0.05)	(0.08)
Flexible 1	0.115	0.117	0.072	0.086	0.044	0.045
	(0.09)	(0.09)	(2.27)	(2.29)	(0.05)	(0.05)
Flexible 1+ IGA	0.145*	0.139	-1.314	-1.376	0.023	0.022
	(0.09)	(0.08)	(2.14)	(2.14)	(0.06)	(0.06)
Flexible 2	0.170*	0.169*	0.666	0.615	0.075	0.074
	(0.09)	(0.09)	(2.22)	(2.23)	(0.05)	(0.05)
Mean of the Dependent Variable	0.86	0.86	5.78	5.78	0.17	0.17
Observations	1068	1068	1068	1068	1068	1068
R-squared	0.062	0.071	0.114	0.123	0.078	0.087
Log Likelihood	-352.1	-346.9	-3432	-3427	372.3	377.4
Control of Stratification	Yes	Yes	Yes	Yes	Yes	Yes
Control for Load Distribution Dates	Yes	Yes	Yes	Yes	Yes	Yes
Control for Household Characterestics	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the village level are shown in brackets. Significant at the 10% (*), 5% (**), and 1% (***). Household Chracterestics include age, age squared, sex, marital status, education qualification of the household head as well as size of the household, number of adults and childrens in the households and total land holdings. We dropped those ovservation that are attrited.

Source: Estimated by the authors using the microdata described in the text.

Table 3.3: Impact of Flexible Microcredit on Repayment Behavior (Default Indicators)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Binary Default (if total due amount at the end of loan cycle > 0)	Binary Default (if total due amount at the end of loan cycle > 0)	Binary Default (if total due adjusted 6 month after the loan cycle with savings amount > 0)	Binary Default (if total due adjusted 6 month after the loan cycle with savings amount > 0)	Total Savings	Total Savings
Constant	0.585***	0.803***	0.271**	0.524*	578.274***	656.627***
	(0.11)	(0.19)	(0.13)	(0.27)	(48.71)	(99.41)
Flexible 1	0.048	0.052	0.12	0.131	-92.650**	-91.958**
	(0.11)	(0.11)	(0.13)	(0.13)	(43.20)	(42.67)
Flexible 1+ IGA	0.031	0.034	0.043	0.05	-79.066	-80.258
	(0.13)	(0.12)	(0.16)	(0.15)	(65.27)	(64.42)
Flexible 2	0.019	0.022	0.157	0.168	-39.795	-38.792
	(0.11)	(0.11)	(0.14)	(0.14)	(57.17)	(56.35)
Mean of the Dependent Variable	0.73	0.73	0.48	0.48	433.51	433.51
Observations	1,068	1,068	1,068	1,068	1,068	1,068
R-squared	0.24	0.101	0.118	0.234	0.194	0.206
Log Likelihood	-504.7	-717.4	-707.2	-509.1	-6979	-6971
Control of Stratification	Yes	Yes	Yes	Yes	Yes	Yes
Control for Load Distribution Dates	Yes	Yes	Yes	Yes	Yes	Yes
Control for Household Characterestics	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the village level are shown in brackets. Significant at the 10% (*), 5% (**), and 1% (***). Household Chracterestics include age, age squared, sex, marital status, education qualification of the household head as well as size of the household, number of adults and childrens in the households and total land holdings. We dropped those ovservation that are attrited.

Source: Estimated by the authors using the microdata described in the text.

Table 3.4: Impact of Flexible Microcredit on Repayment Behavior (Various Indicators upto the flood, May 2012)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Overdue amount upto May (in BDT)	Overdue amount upto May (in BDT)	Overdue as a percentage of total due amount (upto May)	Overdue as a percentage of total due amount (upto May)	No. of missed weeks upto May	No. of missed weeks upto May	Indicator for having atleast one missed week upto May	Indicator for having atleast one missed week upto May
Constant	300.815**	450.436**	0.090**	0.134**	3.989**	6.086***	0.598***	0.472*
	(129.54)	(176.85)	(0.04)	(0.05)	(1.68)	(2.08)	(0.12)	(0.24)
Flexible 1	-7.252	-8.013	-0.002	-0.002	-0.677	-0.693	0.122	0.119
	(129.17)	(129.55)	(0.04)	(0.04)	(1.71)	(1.71)	(0.12)	(0.12)
Flexible 1+ IGA	-29.529	-37.679	-0.009	-0.011	-1.257	-1.354	0.052	0.045
	(128.25)	(128.12)	(0.04)	(0.04)	(1.63)	(1.62)	(0.14)	(0.14)
Flexible 2	77.573	72.532	0.023	0.022	-0.008	-0.078	0.224*	0.219*
	(135.40)	(134.65)	(0.04)	(0.04)	(1.64)	(1.64)	(0.12)	(0.12)
Mean of the Dependent Variable	267.40	267.40	0.08	0.08	3.04	3.04	0.69	0.69
Observations	1,068	1,068	1,068	1,068	1,068	1,068	1,068	1,068
R-squared	0.117	0.13	0.117	0.13	0.129	0.143	0.115	0.121
Log Likelihood	-7801	-7793	871.3	879.2	-3026	-3018	-629	-625.6
Control of Stratification	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control for Load Distribution Dates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control for Household Characterestics	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the village level are shown in brackets. Significant at the 10% (*), 5% (**), and 1% (***). Household Chracterestics include age, age squared, sex, marital status, education qualification of the household head as well as size of the household, number of adults and childrens in the households and total land holdings. We dropped those ovservation that are attrited.

Source: Estimated by the authors using the microdata described in the text.

Table 3.5: Forced repayment pattern by groups

VARIABLES	(1) Forced Montly repayment by group after Monga	(2) Forced Montly repayment by group after Monga	(3) Forced Montly repayment by group after Monga	(4) Forced Montly repayment by group after Monga
Constant	0.006	-0.151	0.006	-0.152
	(0.03)	(0.12)	(0.03)	(0.12)
Flexible 1	-0.013	-0.015	-0.013	-0.014
	(0.02)	(0.02)	(0.02)	(0.02)
Flexible 1+ IGA	-0.02	-0.02	-0.019	-0.018
	(0.03)	(0.03)	(0.03)	(0.03)
Flexible 2	0.093*	0.093*	0.092*	0.092*
	(0.05)	(0.05)	(0.05)	(0.05)
Mean of the Dependent Variable	0.041	0.041	0.042	0.042
Observations	1,068	1,068	1,079	1,079
R-squared	0.155	0.161	0.153	0.159
Log Likelihood	288.3	292.4	295.5	299.6
Control of Stratification	Yes	Yes	Yes	Yes
Control for Load Distribution Dates	Yes	Yes	Yes	Yes
Control for Household Characterestics	No	Yes	No	Yes

Notes: Robust standard errors clustered at the village level are shown in brackets. Significant at the 10% (*), 5% (**), and 1% (***). Household Chracterestics include age, age squared, sex, marital status, education qualification of the household head as well as size of the household, number of adults and childrens in the households and total land holdings. We dropped those ovservation that are attrited for estimations in column (1) and (2).

Source: Estimated by the authors using the microdata described in the text.

Table 4: Regression Result for Satisfaction Survey

	Dependent variable: dummy for satisfaction
Intercept	0.456***
	[0.104]
D 1 (dummy for Flexible 1)	0.303**
	[0.124]
D 2 (dummy for Flex. 1+IGA)	0.106
	[0.163]
D 3 (dummy for Flexible 2)	0.206
	[0.128]
R^2	0.050
Adjusted R^2	0.047

Notes: The number of observations is 1,080. Robust standard errors clustered at the village level are shown in brackets. Significant at the 10% (*), 5% (**), and 1% (***).

Source: Compiled by the authors using the benchmark survey data.

Table 5: Definitions and Summary Statistics of Variables Related with Food Consumption, Northern Bangladesh, 2011-12

Variable	Definition	N	Mean	Std. Dev.	Min	Max
Food consumption during *monga* 2011						
num_mong1	Number of stomach-full meals in a day during Monga 2011	1,414	2.114	0.411	1	3
num_mong2	Number of minimum stomach-full meals a day during Monga 2011	1,412	1.693	0.498	1	3
safe_mong	Dummy for food safety during Monga 2011 (defined as *num_mong2* = 2 or 3)	1,412	0.676	0.468	0	1
meat_mong	Dummy for having meat within a month during Monga 2011	1,414	0.756	0.430	0	1
Food consumption during normal times in 2012						
num_norm1	Number of stomach-full meals in a day during normal time in 2012	1,416	2.859	0.362	1	3
num_norm2	Number of minimum stomach-full meals a day during normal time in 2012	1,415	2.127	0.586	1	3
safe_norm	Dummy for food safety during normal time in 2012 (defined as *num_norm2* = 2 or 3)	1,415	0.885	0.319	0	1
Food consumption during *monga* in 2012						
monga_foodHH	Number of Stomack full meals during the Monga 2012	1,400	2.100	0.397	1	3
minimum_monga	Number of Minimum Stomack full meals during the Monga 2012	1,400	1.358	0.504	1	3
foodshortage	Dummy for Foood Shortage: Is this Household suffering from food shortage during the monga of 2012	1,440	0.700	0.458	0	1
nmeat_monga	Number of times your dinner contains meat (chicken, beef or lamb) during a typical week in the Monga	1,440	0.001	0.037	0	1
nfish_monga	Number of times your dinner contains fish during a typical week in the Monga 2012	1,440	0.953	0.903	0	7
negg_monga	Number of times your dinner contains egg during a typical week in the Monga 2012	1,440	0.204	0.433	0	4

Note: Mean and standard deviations are simple ones, without weighting. The question of "Number of (minimum) stomach-full meals in a day" was asked of the respondents who had reported a typical number, and so that the answer took an integer value of either 1, 2, or 3.

Source: Compiled by the authors using the 2012 resurvey data (Panel 3) and 2012 Monga survey (Panel 4)

Table 6: Impact of Flexible Microcredit on Food Consumption

	Food consumption during *monga* 2011				Food consumption during normal times in 2012	
	num_mong1	*num_mong2*	*safe_mong*	*meat_mong*	*num_norm2*	*safe_norm*
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	2.096***	1.708***	0.697***	0.837***	1.708***	0.697***
	[0.043]	[0.074]	[0.075]	[0.040]	[0.074]	[0.075]
D 1 (dummy for Flexible 1)	0.009	0.025	0.028	-0.077	0.025	0.028
	[0.066]	[0.086]	[0.088]	[0.061]	[0.086]	[0.088]
D 2 (dummy for Flex. 1+IGA)	0.144*	0.012	-0.005	-0.146	0.012	-0.005
	[0.075]	[0.110]	[0.102]	[0.089]	[0.110]	[0.102]
D 3 (dummy for Flexible 2)	-0.020	-0.053	-0.070	-0.093	-0.053	-0.070
	[0.056]	[0.092]	[0.088]	[0.057]	[0.092]	[0.088]
D 4 (dummy for non-borrower)	0.012	-0.036	-0.039	-0.082*	-0.036	-0.039
	[0.053]	[0.075]	[0.075]	[0.045]	[0.075]	[0.075]
R^2	0.014	0.004	0.006	0.008	0.004	0.006
F -stat. for zero slopes of all dummies	1.43	0.52	0.75	1.10	0.52	0.75
F -stat. for zero slopes of *D* 1, *D* 2, and *D* 3	1.84	0.43	0.81	1.41	0.43	0.81
Number of observations	1,414	1,412	1,412	1,414	1,412	1,412

Notes: Robust standard errors clustered at the village level are shown in brackets. Significant at the 10% (*), 5% (**), and 1% (***).

Source: Estimated by the authors using the microdata described in the text.

Table 7: Impact of Flexible Microcredit on Food Consumption (Robustness check with baseline controls)

	Food consumption during *monga* 2011				Food consumption during normal times in 2012	
	num_mong1	*num_mong2*	*safe_mong*	*meat_mong*	*num_norm2*	*safe_norm*
Baseline village characteristics						
Mondir (Hindu temple)	0.078	0.145**	0.144*	0.049	0.018	0.027
	[0.092]	[0.065]	[0.074]	[0.053]	[0.075]	[0.035]
Town	-0.124	-0.045	-0.046	-0.002	0.049	0.036
	[0.086]	[0.087]	[0.094]	[0.078]	[0.103]	[0.033]
Baseline household characteristics						
Household size (number of members)	0.031***	0.019	0.020*	0.020**	0.058***	0.026***
	[0.012]	[0.012]	[0.011]	[0.009]	[0.013]	[0.007]
Average age of household members	-0.001	0.000	0.000	-0.001	-0.003	0.000
	[0.001]	[0.001]	[0.001]	[0.001]	[0.002]	[0.001]
Ratio of adults (age 15+)	0.230**	-0.016	0.013	0.079	0.222**	0.06
	[0.087]	[0.067]	[0.061]	[0.057]	[0.090]	[0.045]
No. of chickens and ducks owned	0.093***	0.012	0.013	0.032	0.001	-0.029
	[0.033]	[0.030]	[0.029]	[0.039]	[0.043]	[0.023]
Treatment status						
$D1$ (dummy for Flexible 1)	-0.006	0.028	0.031	-0.074	-0.040	-0.057
	[0.068]	[0.093]	[0.095]	[0.063]	[0.077]	[0.038]
$D2$ (dummy for Flex. 1+IGA)	0.109	0.024	0.006	-0.143	0.092	0.046
	[0.074]	[0.118]	[0.110]	[0.091]	[0.083]	[0.033]
$D3$ (dummy for Flexible 2)	-0.027	-0.022	-0.037	-0.080	-0.024	-0.060
	[0.063]	[0.100]	[0.097]	[0.058]	[0.087]	[0.042]
$D4$ (dummy for non-borrower)	0.002	-0.017	-0.020	-0.073	-0.006	-0.057
	[0.055]	[0.083]	[0.083]	[0.045]	[0.068]	[0.035]
Intercept	1.843***	1.590***	0.552***	0.715***	1.849***	0.792***
	[0.109]	[0.127]	[0.125]	[0.085]	[0.133]	[0.061]
R^2	0.040	0.015	0.019	0.014	0.023	0.030
F-stat. for zero slopes of all explan. variables	2.72***	1.23	1.36	0.92	2.74***	3.23***
F-stat. for zero slopes of $D1$, $D2$, $D3$, and $D4$	1.26	0.26	0.36	0.91	0.82	5.44***
Number of observations	1,414	1,412	1,412	1,414	1,415	1,415

Notes: Robust standard errors clustered at the village level are shown in brackets. Significant at the 10% (*), 5% (**), and 1% (***).

Source: Estimated by the authors using the microdata described in the text.

Table 8: Impact of Flexible Microcredit on Food Consumption during the monga 2012

	Number of Stomack full meals during the Monga 2012	*Number of Stomack full meals during the Monga 2012*	*Number of Minimum Stomack full meals during the Monga 2012*	*Number of Minimum Stomack full meals during the Monga 2012*	*Dummy for Foood Shortage: Is this Household suffering from food shortage during the monga of 2012*	*Dummy for Foood Shortage: Is this Household sufferin from food shortage during the monga of 2012*
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	2.051***	2.133***	1.245***	1.382***	0.772***	0.950***
	(0.06)	(0.13)	(0.10)	(0.20)	(0.09)	(0.16)
D 1 (dummy for Flexible 1)	0.120	0.121	0.156	0.160	-0.157*	-0.157*
	(0.07)	(0.07)	(0.11)	(0.11)	(0.09)	(0.09)
D 2 (dummy for Flex. 1+IGA)	0.305***	0.305***	0.380***	0.378***	-0.395***	-0.382***
	(0.11)	(0.11)	(0.12)	(0.12)	(0.11)	(0.11)
D 3 (dummy for Flexible 2)	0.158*	0.158**	0.182	0.187*	-0.195*	-0.196**
	(0.08)	(0.08)	(0.11)	(0.11)	(0.10)	(0.10)
D 4 (dummy for non-borrower)	0.01	0.01	0.02	0.02	0.00	(0.00)
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)
R^2	0.108	0.115	0.091	0.099	0.107	0.129
Number of observations	1,389	1,389	1,389	1,389	1,422	1,422
Control of Stratification and District	Yes	Yes	Yes	Yes	Yes	Yes
Control for Load Distribution Dates	Yes	Yes	Yes	Yes	Yes	Yes
Control for Household Characterestics	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the village level are shown in brackets. Significant at the 10% (*), 5% (**), and 1% (***). Household Chracterestics include age, age squared, sex, marital status, education qualification of the household head as well as size of the household, number of adults and childrens in the households and total land holdings. We dropped those ovservation that are attrited.

Source: Estimated by the authors using the microdata described in the text.

Table 9: Impact of Flexible Microcredit on Protein Intake during the monga 2012

	Number of times your dinner contains meat (chicken, beef or lamb) during a typical week in the Monga 2012	*Number of times your dinner contains meat (chicken, beef or lamb) during a typical week in the Monga 2012*	*Number of times your dinner contains fish during a typical week in the Monga 2012*	*Number of times your dinner contains fish during a typical week in the Monga 2012*	*Number of times your dinner contains egg during a typical week in the Monga 2012*	*Number of times your dinner contains egg during a typical week in the Monga 2012*
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	-0.003	-0.005	1.127***	1.089***	0.207***	0.095
	(0.00)	(0.01)	(0.10)	(0.29)	(0.06)	(0.17)
D 1 (dummy for Flexible 1)	0.002	0.002	0.212*	0.202*	-0.077	-0.079
	(0.00)	(0.00)	(0.12)	(0.12)	(0.06)	(0.06)
D 2 (dummy for Flex. 1+IGA)	0.004	0.005	0.538**	0.497**	-0.062	-0.066
	(0.00)	(0.00)	(0.21)	(0.21)	(0.07)	(0.07)
D 3 (dummy for Flexible 2)	0.000	0.001	0.251*	0.241*	-0.053	-0.051
	(0.00)	(0.00)	(0.15)	(0.14)	(0.07)	(0.07)
D 4 (dummy for non-borrower)	-0.001	-0.001	-0.050	-0.045	-0.040	-0.050*
	(0.00)	(0.00)	(0.05)	(0.05)	(0.03)	(0.03)
R^2	0.01	0.01	0.25	0.29	0.05	0.07
Number of observations	1,422	1,422	1,422	1,422	1,422	1,422
Control of Stratification and District	Yes	Yes	Yes	Yes	Yes	Yes
Control for Load Distribution Dates	Yes	Yes	Yes	Yes	Yes	Yes
Control for Household Characterestic	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the village level are shown in brackets. Significant at the 10% (*), 5% (**), and 1% (***). Household Chracterestics include age, age squared, sex, marital status, education qualification of the household head as well as size of the household, number of adults and childrens in the households and total land holdings. We dropped those ovservation that are attrited.

Source: Estimated by the authors using the microdata described in the text.

Appendix Table 1: Balance Test at the Village Level

	A. Dependent variable: Location (strata used in randomization)			B. Dependent variable: Minutes of travel to the nearest facility					
	Distance from the closest bus station (km)	Dummy for a *char* village	Dummy for an inland village	Bazar	College	Mondir (Hindu temple)	Town	Bus stand	Railway station
Intercept	32.167***	0.250*	0.583***	7.917**	27.083***	29.583***	34.167***	29.583***	61.667***
	[9.695]	[0.129]	[0.146]	[3.711]	[6.764]	[8.100]	[7.666]	[6.764]	[16.069]
D 1 (dummy for Flexible 1)	12.458	0.000	0.000	1.042	3.75	3.125	8.958	26.292	18.750
	[12.431]	[0.158]	[0.179]	[4.536]	[7.833]	[9.846]	[8.661]	[21.057]	[19.620]
D 2 (dummy for Flex. 1+IGA)	10.333	0.000	0.000	5.417	13.333*	25.000*	25.833**	24.583*	25.417
	[14.172]	[0.182]	[0.207]	[5.382]	[7.904]	[13.530]	[11.848]	[12.809]	[22.523]
D 3 (dummy for Flexible 2)	15.25	0.000	0.000	-0.417	1.875	16.458*	13.333	10.417	26.458
	[12.426]	[0.158]	[0.179]	[4.495]	[8.071]	[8.869]	[8.463]	[8.054]	[20.090]
R^2	0.020	0.000	0.000	0.026	0.047	0.109	0.109	0.028	0.028
F -stat. for zero slopes of all dummies	0.537	0.000	0.000	0.556	1.731	2.556*	1.807	1.549	0.649

Note: The number of observations is 72. Robust standard errors are shown in brackets. Significant at the 10% (*), 5% (**), and 1% (***). Dependent variables for B are measured in minutes if public transportation is used and the value of zero is assigned when the facility exists in the village.

Source: Estimated by the authors using the benchmark survey data.

Appendix Table 2: Balance Test at the Household Level

	A. Dep. variable: Characteristics of the head				B. Dependent variable: Charactesitics of household members					
	Age	Dummy for female	Dummy for literacy	Years of schooling	Household size	Average age	Female ratio	Ratio of adults (age 15+)	Literacy rate of adult males	Literacy rate of adult females
Intercept	38.672***	0.228***	0.239***	1.589***	3.722***	26.367***	0.557***	0.702***	0.277***	0.229***
	[1.142]	[0.063]	[0.038]	[0.246]	[0.211]	[1.238]	[0.021]	[0.026]	[0.038]	[0.037]
D 1 (dummy for Flexible 1)	-0.536	-0.036	-0.017	-0.186	0.328	-1.302	-0.022	-0.045	0.009	-0.014
	[1.296]	[0.081]	[0.045]	[0.287]	[0.257]	[1.404]	[0.026]	[0.030]	[0.047]	[0.042]
D 2 (dummy for Flex. 1+IGA)	-0.411	-0.117	0.006	-0.183	0.433	-2.166	-0.031	-0.039	0.030	0.108*
	[1.376]	[0.070]	[0.070]	[0.362]	[0.280]	[1.489]	[0.023]	[0.030]	[0.057]	[0.056]
D 3 (dummy for Flexible 2)	-0.467	-0.058	-0.028	-0.189	0.431*	-2.144	-0.044*	-0.077**	0.007	0.050
	[1.259]	[0.078]	[0.048]	[0.307]	[0.247]	[1.326]	[0.024]	[0.029]	[0.048]	[0.046]
D 4 (dummy for non-borrower)	-0.583	-0.022	-0.039	-0.198	0.078	0.246	-0.011	-0.016	0.019	0.001
	[1.206]	[0.063]	[0.041]	[0.273]	[0.211]	[1.333]	[0.023]	[0.030]	[0.044]	[0.041]
R^2	0.000	0.007	0.001	0.000	0.012	0.013	0.005	0.015	0.000	0.010
F -stat. for zero slopes of all dummies	0.06	1.54	0.44	0.14	2.59**	3.63***	1.42	4.95***	0.11	2.24*
F -stat. for zero slopes of *D* 1, *D* 2, and *D* 3	0.06	1.31	0.16	0.16	1.12	1.13	1.27	2.85**	0.10	2.86**
Number of observations	1,440	1,440	1,440	1,440	1,440	1,437	1,440	1,440	1,252	1,428
	C: Dependent variable: Landholdings					D: Dependent variable: Liquid asset				
	Dummy for owning the house	Dummy for owning farm land	Size of operational farmland for *aus*	Size of operational farmland for *aman*		Total value of household assets (BDT)	Dummy for owning livestock animals	Number of cows and bulls owned	Number of goats and sheep owned	Number of chickens and ducks owned
Intercept	0.306***	0.056**	0.567**	2.167***	2.339***	2827***	0.656***	0.378***	0.464***	2.961***
	[0.093]	[0.022]	[0.276]	[0.737]	[0.792]	[333]	[0.076]	[0.097]	[0.128]	[0.717]
D 1 (dummy for Flexible 1)	0.150	-0.033	0.956	0.242	1.294	425	0.039	0.117	0.042	0.250
	[0.113]	[0.023]	[0.804]	[1.160]	[1.254]	[527]	[0.091]	[0.139]	[0.161]	[0.837]
D 2 (dummy for Flex. 1+IGA)	0.072	-0.022	0.728	0.322	-1.022	613	-0.011	0.272	0.153	-0.561
	[0.129]	[0.029]	[0.728]	[1.518]	[1.036]	[473]	[0.093]	[0.230]	[0.182]	[0.838]
D 3 (dummy for Flexible 2)	0.125	-0.017	-0.147	-1.222	-0.861	411	-0.025	0.003	0.092	-0.192
	[0.121]	[0.029]	[0.318]	[0.834]	[0.917]	[395]	[0.090]	[0.121]	[0.158]	[0.812]
D 4 (dummy for non-borrower)	0.083	-0.014	0.753	0.186	0.553	381	-0.042	0.069	0.011	-0.689
	[0.090]	[0.023]	[0.621]	[1.162]	[1.041]	[319]	[0.077]	[0.110]	[0.135]	[0.676]
R^2	0.009	0.003	0.004	0.004	0.008	0.002	0.004	0.009	0.002	0.005
F -stat. for zero slopes of all dummies	0.67	0.88	1.01	1.17	1.82	0.50	1.19	0.61	0.43	1.81
F -stat. for zero slopes of *D* 1, *D* 2, and *D* 3	0.65	0.85	1.17	1.46	1.68	0.60	0.31	0.76	0.28	0.61
Number of observations	1,440	1,440	1,440	1,440	1,440	1,440	1,440	1,440	1,440	1,440

Notes: Robust standard errors clustered at the village level are shown in squared brackets. Significant at the 10% (*), 5% (**), and 1% (***) levels.

Source: Estimated by the authors, using the microdata described in the text.