NBER WORKING PAPER SERIES

BORROWING REQUIREMENTS, CREDIT ACCESS, AND ADVERSE SELECTION: EVIDENCE FROM KENYA

William Jack
Michael Kremer
Joost de Laat
Tavneet Suri

Working Paper 22686
http://www.nber.org/papers/w22686

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2016, Revised September 2019

The authors would like to thank Egor Abramov, William Glennerster, Matthew Goodkin-Gold, Kamran Jamil, Benjamin Marx, Adam Ray, Itzchak Raz, Indrani Saran and Kevin Xie for exceptional research assistance. Our gratitude also goes out to Suleiman Asman, Antony Wainaina and Nadir Shams for excellent management, field supervision and data collection. We are grateful to Joshua Angrist, Michael Boozer, Esther Duflo, Rachel Glennerster and to seminar audiences at the CEGA East Africa Evidence Summit, Nairobi; Georgetown University; Harvard University; the IGC Trade and Development Conference at Stanford University; the IPA Microfinance Conference; the MIT Development Lunch; Northwestern; Notre Dame; University of California, San Diego; Tinbergen Institute, Amsterdam; and the World Bank for comments. We thank the Gates Foundation, Google and the Agricultural Technology Adoption Initiative for funding. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Borrowing Requirements, Credit Access, and Adverse Selection: Evidence from Kenya
William Jack, Michael Kremer, Joost de Laat, and Tavneet Suri
NBER Working Paper No. 22686
September 2016, Revised September 2019
JEL No. O13,O16

ABSTRACT

Do the stringent formal sector borrowing requirements common in many developing countries restrict credit access, technology adoption, and welfare? When a Kenyan dairy's savings and credit cooperative randomly offered some farmers the opportunity to replace loans with high down payments and stringent guarantor requirements with loans collateralized by the asset itself — a large water tank — loan take-up increased from 2.4% to 41.9%. (In contrast, substituting joint liability requirements for deposit requirements did not affect loan take up.) There were no repossessions among farmers allowed to collateralize 75%of their loans, and there was only a 0.7% repossession rate among those offered 96% asset collateralization. A Karlan-Zinman test based on waiving borrowing requirements ex post finds evidence of adverse selection with lowered deposit requirements, but not of moral hazard. A simple model and rough calibration suggests that adverse selection may deter lenders from making welfare-improving loans with lower deposit requirements, even after introducing asset collateralization. We estimate that 2/3 of marginal loans led to increased water storage investment. Real effects of loosening borrowing requirements include increased household water access, reductions in child time spent on water-related tasks, and greater school enrollment for girls.

William Jack
Georgetown University
Billy.Jack@georgetown.edu

Michael Kremer
Harvard University
Department of Economics
Littauer Center M20
Cambridge, MA 02138
and NBER
mkremer@fas.harvard.edu

Joost de Laat
University of Utrecht
joostdelaat@gmail.com

Tavneet Suri
MIT Sloan School of Management
100 Main Street, E62-517
Cambridge, MA 02142
and NBER
tavneet@mit.edu

1 Introduction

Formal-sector lenders in developing countries often impose very tight borrowing requirements, such as high deposit requirements or guarantor requirements. To the extent that these requirements restrict credit access, investment, technology adoption, and welfare, there may be a strong case for steps to encourage lenders to loosen these borrowing requirements, for example by loosening regulatory caps on interest rates, strengthening legal and contract enforcement institutions to expand the scope for collateralization of debt, or even subsidizing lenders. While the evidence summarized in Banerjee et al. (2015) suggests both limited take up and limited impact of expanding credit access through standard microfinance contracts, it is possible that moving from the very restrictive borrowing requirements in many developing contracts to borrowing requirements more typical of developed countries would have a bigger impact.

We examine the impact of replacing loans with high down payments and stringent guarantor requirements with asset- collateralized loans, similar to the mortgages and car loans that are common in developed countries. In particular, we studied a Kenyan dairy's saving and credit cooperative which randomly offered different borrowing conditions to different members. Its standard borrowing conditions required that one third of loans be secured with deposits by the borrower, and that the remaining two thirds be secured with cash or shares from guarantors. Allowing borrowers to collateralize loans for water tanks using assets purchased with the loans dramatically increased borrowing. Only 2.4% of farmers borrowed under the savings cooperative's standard borrowing conditions. The loan take up rate increased to 23.9% under 25% deposit or guarantor requirements and 75% tank-collateralization. The take-up rate further increased to 41.9% when all but 4% of the loan could be collateralized with the tank. Thus more than 90% of those who wished to borrow at the available interest rate were credit-constrained. Results were similar in a separate out-of-sample test.

However, we find no evidence that joint liability expands credit access. There was no statistically significant difference in loan take up between farmers offered loans with a 25 percent deposit requirement and those offered the opportunity to substitute guarantors for all but 4 percent of the loan value.

Defaults did not increase with moderate deposit requirements and asset collateralization. In particular, there were no tank repossessions when 75% of the loan could be collateralized with the tank itself and 25% was collateralized with deposits from the borrower and/or guarantors. Reducing the deposit requirement to 4% with 96% asset-collateralization induced a 0.7% repossession rate overall, corresponding to a 1.63% repossession rate among the marginal farmers induced to borrow by the lower borrowing requirements. The hypothesis of equal rates of tank repossession under a 4% deposit requirement and under a 25% deposit or guarantor requirement is rejected at the 5.25% level using a Fisher exact test. Karlan-Zinman tests based on *ex post* waivers or borrowing requirements suggest that this difference is entirely due to adverse selection, rather than the treatment effects associated with moral hazard.

A simple model suggests that under adverse selection, a lender with market power facing interest rate caps, such as the savings and credit cooperative we study, will set deposit requirements above the socially optimal level even with asset collateralization. To see this, note that at the margin, raising deposit requirements selects out unprofitable borrowers but imposes a cost on credit-constrained inframarginal borrowers, and a profit-maximizing lender will not internalize these costs to inframarginal borrowers. A rough calibration suggests that the cooperative could increase profits by moving to 75% but not 96% asset collateralization. Consistent with the results of the calibration, after learning the results of the program, the lender changed its policy to allow 75% collateralization with the tank, but not to allow 96% collateralization.

With regards to investments, we find that those offered the opportunity to collateralize loans with the tanks were more likely to have purchased tanks and had more water storage capacity overall. These results also suggest that improving credit access can influence technology adoption (Zeller et al., 1998). Consistent with Devoto et al. (2013), our results suggest that credit provision can contribute to increased access to clean water in the developing world. Children of households offered less restrictive credit terms spent somewhat less time collecting water and tending to livestock and difference-in-difference estimates find that fewer girls in these households were out of school. We find no impact on milk production.

The primary contributions of this paper are twofold. First, we extend the literature on asset-collateralized loans in developing countries. Existing literature on transition and developed economies (Aretz, Campello, and Marchica 2016, Calomiris et al. 2016) provides evidence that when institutional reforms at the national level expand collateralization options, borrowing increases at both extensive (higher loan takeup) and intensive (more leverage) margins. One such expansion of collateralization options is the enhancement of the ability to collateralize loans with the assets that they are used to purchase (Assuncao et al. 2014).[1] Our context allows identification from randomization at the level of individual loans. The result is a novel estimate of the direct impact on loan uptake of replacing a high-deposit loan with an asset-collateralized, low-deposit loan. Secondly, we measure how repossession rates vary under different loan contracts, and use a Karlan-Zinman test to decompose the effect of lower deposit requirements on repossession into moral hazard and adverse selection effects.[2] Our model builds on the results of the Karlan-Zinman test to suggest that even after asset-collateralization is allowed, lenders will set deposit requirements which are too high from a social welfare standpoint.

We also provide results that contribute to the literature on credit access in the developing world. A large literature in development economics examines the potential for microfinance to expand access to credit, often through joint liability lending (Morduch, 1999; Hermes and Lensink, 2007). We find very large effects of asset collateralization on credit uptake consistent with Feder et al. (1988).

The rest of the paper is organized as follows: Section two provides background on smallholder dairy farming in the region we study. Section three presents a model with which we interpret the data. Section four explains the program design. Section five explains the data and our empirical specifications. Section six discusses the impact of borrowing requirements on loan

[1] Skrastins (2016) also considers asset collateralization, examining how institutional design can facilitate easier collection of debt and collateral.

[2] For a similar decomposition of deposit requirement changes into moral hazard and adverse selection effects in the developed context, see Adams, Einav and Levin (2009).

take up and on borrower characteristics. Section seven discusses the treatment, selection, and overall impacts of relaxing borrowing conditions on loan recovery and tank repossession, and calibrates the model to the data. Section eight discusses the impacts on real outcomes. Section nine concludes by discussing potential policy implications and directions for further research.

2 Background

WHO and UNICEF estimate that approximately 900 million people lack access to water at their homes (2010), with substantial consequences for global health and human development. We examine the potential of asset-collateralized credit to expand access to large rainwater harvesting tanks among a population of dairy farmers in an area straddling Kenya's Central and Rift Valley provinces. Because installation of water supply at the household level requires substantial fixed costs, there has been increasing interest in whether extension of credit can help improve access to water (Devoto et al 2011).[3]

Collection of water from distant sources limits water use, including for hand washing and cleaning, with potential negative health consequences (Wang and Hunter, 2010; Esrey 1996). It also imposes a substantial time burden, particularly for women and girls, with potentially negative consequences for schooling.[4]

Dairy farmers in particular benefit from reliable access to water because dairy cattle require a regular water supply (Nicholson (1987), Peden et al. (2007), and Staal et al (2001)). Without easy access to water, the most common means of watering cattle is to take them to a source every two or three days, which is time consuming and can expose cattle to disease (Kristjanson et al. 1999).[5]

Rainwater harvesting tanks provide convenient access to water, reducing the need to travel

[3] See also http://www.waterforpeople.org/.

[4] In our baseline survey, women report spending 21 minutes per day fetching water, three times as much as men, and our enumerators reported that women were typically more eager than their husbands to purchase tanks.

[5] During the baseline survey, it was reported that farmers spent on average ten hours per week taking their cows to the water sources.

to collect water and then carry it home. Moreover, rainwater is not subject to contamination by disease-bearing fecal matter. In the area we examine, approximately 30% of farmers are connected to piped water systems, but these systems provide water only intermittently, typically three days per week. 70% of farmers do not have any connection to a water system. Historically, many farmers in the area used stone or metal tanks to harvest rainwater or store piped water for days when piped water is not available. Approximately one-quarter of comparison group farmers had a water storage tank of more than 2,500-liter capacity at baseline. However, stone tanks are susceptible to cracking, and metal tanks are susceptible to rusting, so neither approach is particularly durable. Lightweight, durable plastic rainwater harvesting tanks were introduced about 10 years prior to the start of the study. These plastic rainwater harvesting tanks are displayed prominently at agricultural supply dealers in the area and are the dominant choice for farmers obtaining new tanks. Almost all farmers are thus familiar with the product, but since they cost about $320 or 20% of annual household consumption, very few farmers tend to own them.

Like most of Kenya's approximately one million smallholder dairy farmers, the farmers in our study sell milk to a dairy cooperative, the Nyala dairy cooperative (although not all are members of the cooperative). The Nyala dairy cooperative performs basic quality tests, cools the milk, and then sells it to a large-scale milk producer for pasteurization and sale to the national market. It keeps track of milk deliveries and pays farmers monthly. During the time period we study, selling to the Nyala dairy was more lucrative for farmers than selling on the local market or to another dairy, which would have involved higher transport costs.[6]

The Nyala dairy cooperative has an associated savings and credit association (SACCO). These are widespread in Kenya, with total membership of almost five percent of the population.[7] SAC-

[6]Casaburi and Macchiavello (2014) examine a different Kenyan context in which farmers sell to dairies even though the dairy pays a lower price than the local market, arguing that farmers value the savings opportunity generated by the monthly, rather than daily, payments provided by dairies.

[7]Until 2012, many dairy cooperatives ran SACCOs as a service to their members, with the dairy cooperative's management also overseeing the SACCO. The 2012 SACCO act made cooperatives separate farming and banking activities. SACCOs previously run by a dairy cooperative became a separate legal entity but have tended to retain strong links with the dairy cooperative.

COs are typically limited to a 12% annual interest rate, but in some cases they can charge 14% annually (SASRA, 2013). In practice, this is interpreted as 1% monthly interest and 1.2% monthly interest. As a result, SACCOs are typically conservative in their lending, imposing stringent borrowing requirements.

In the SACCO we examine, the borrower must have savings deposited in the SACCO worth 1/3 of the total amount of the loan and must find up to three guarantors willing to collateralize the remaining 2/3 of the loan with savings and/or shares in the cooperative. Borrowers and guarantors are paid the same standard 3% quarterly interest on funds deposited in the SACCO as are other depositors. These terms are fairly typical. The Nyala SACCO offers loans for a variety of purposes, mostly school fees and emergency loans in the case of illness and agricultural loans in kind (advances on feed). In the year prior to the study, it made just 292 cash loans to members, averaging KSh 25,000 ($315).

In order to examine how potential borrowers respond to different potential loan contracts, we focus on an environment in which lending is feasible. Several features of the institutional environment are favorable to lending. First, farmers who borrow agree to let the SACCO deduct loan repayments from the dairy's payments to the farmer for milk. This provides a very easy mechanism for collecting debt that not only has low administrative cost for the lender but also effectively makes repayment the default option for borrowers, instead of requiring them to actively take steps to repay debt. Second, the dairy paid a higher price for milk than alternative buyers, providing farmers with an incentive to maintain their relationship with the dairy. Finally, the SACCO may have more legitimacy in collecting debt than would an outside for-profit lender.

The physical characteristics of rainwater harvesting tanks also make them well-suited as collateral. The tanks are bulky and have to be installed next to the user's house, so a lender seeking to repossess a tank can find them easily. Moreover, tanks have no moving parts and are durable, so they preserve much of their value through the repossession and resale process. Finally, while tanks are too large to be easily transported by hand for more than a short distance, a lender seeking to repossess them can easily load them onto a truck.

3 Model

With full information there would be no need for collateral, deposits, or guarantors, and borrowers with a tank valuation up to a certain amount would get loans. However, in the presence of asymmetric information about valuations on the one hand, and outcome realizations on the other, adverse selection and moral hazard preclude attainment of the first best. In order to help motivate the empirical work in subsequent sections, we build a simple model in which a lender can respond to such imperfections by introducing non-price rationing mechanisms into credit contracts, but in doing so fails to achieve the information-constrained social optimum. .

In Section 3.1 we lay out the assumptions . We allow risk-averse potential borrowers to vary in their valuation of tanks, and in initial wealth. Given their wealth and tank valuations as well as the deposit required by the lender, potential borrowers choose whether to borrow to buy a tank, in which case they must use some of their wealth for the deposit, constraining their first-period consumption. Remaining wealth can be used for first-period consumption or additional savings for period 2. Borrowers then receive stochastic income and choose whether to repay the loan or allow the lender to repossess the tank.

In section 3.2, we first consider the problem of a borrower deciding whether to repay given the borrower's first period savings (defined to include the deposit), tank valuation, and income realization. We then solve backwards to the problem of a potential borrower deciding whether to take out a loan given their initial wealth, their tank valuation, and the required deposit. We show that if potential borrowers are credit constrained, high deposit requirements will have a selection effect on repayment in which they screen out low-valuation or low-wealth borrowers who are relatively unlikely to repay. High deposit requirements will also have a treatment effect on repayment conditional on borrowing, lowering the threshold tank valuation above which borrowers choose to repay the loan for each possible period-two income realization.

In section 3.3, we work back further to the problem of the lender choosing the size of the required deposit. To reflect our institutional context, we consider a monopoly lender with exogenously fixed interest rates. We show that, since in the presence of adverse selection, a lender fails to internalize the cost to credit-constrained inframarginal borrowers due to a high deposit requirement, stricter deposit requirements than would be socially optimal are chosen.

3.1 Assumptions

Below we describe key assumptions of the model in addition to the basic framework. These key assumptions are designed to ensure that the support of first-period wealth, second-period income, and tank valuation generate, for any deposit requirement, some marginal borrowers and some inframarginal credit-constrained borrowers. We also make some assumptions to assure that we focus on interesting/relevant cases. For example, we assume that the distribution of shocks is sufficiently wide that some borrowers will default in some states of the world. We also make some technical assumptions to ensure the profit function is well-behaved and continuous.

Borrower i's valuation of the tank is denoted θ_i. θ_i is private information encompassing utility benefits of the tank, time savings, and any dairy farming productivity and risk-reduction benefits. (These are likely to vary among farmers, for example, due to distance from other water sources, availability of household labor, and taste for clean water.) There is a continuum of potential borrowers, with water tank valuation continuously distributed over the interval $[\underline{\theta}, \overline{\theta}]$ according to some cumulative distribution function $F(\theta)$ with a probability mass function that is continuous on its support. Potential borrowers value consumption of a composite good c as well as water tanks, with preferences for potential borrower i represented by a utility function $U(\theta_i, c) = u(c_1) + u(c_2) + \theta_i I_2(T)$, where u is at least three-times continuously differentiable, $u' > 0, u'' < 0, \lim_{c \to 0} u' = \infty$ and $\lim_{c \to \infty} u' = 0$ and $I_2(T)$ is an indicator for owning a tank at period $t = 2$. c_1 and c_2 represent non-tank consumption in each of the two periods, and we impose the constraint $c_1, c_2 \geq 0$.[8] For simplicity, discounting and net present discounted value weightings are set aside, and we assume utility does not depend on tank ownership in period 1, $I_1(T)$.

Potential borrower i has an initial wealth w_i at period $t = 1$, drawn from the interval $\left[\underline{W}, \overline{W}\right]$ according to the distribution $F_w(\cdot)$ which is continuously differentiable. The realized value of w is private information, known only to the borrower. Income at period $t = 2$ is denoted y_i, and drawn stochastically from the interval $\left[\underline{Y}, \overline{Y}\right]$. In order to ensure differentiabilty of the profit function, we assume that y_i is drawn from a uniform distribution and that $\overline{Y}$ is large enough that a borrower with second-period income $\overline{Y}$ has higher wealth after repayment than a borrower with second period income $\underline{Y}$ has after repossession. Formally, $\overline{Y} > \underline{Y} + R_T P$. The final assumption we invoke to ensure differentiability is assumption A, described in the appendix. [9] The realized value of y is also private information, known only to the borrower. The distributions of initial wealth, water tank valuation and income are independent, have positive densities throughout their supports.

Potential borrowers can purchase tanks at price P in period $t = 1$ through a contract with the lender in which they must repay $R_T P$ at $t = 2$, where R_T is the gross interest rate. If they purchase a tank, then in period $t = 2$ they choose whether to repay the loan or allow the tank to be repossessed. We assume that the support of θ is wide enough that some potential borrowers are not willing to purchase tanks at full cost, but every potential borrower would purchase a tank if it were free. In particular, assume that $0 < \underline{\theta}$, and that the potential borrower with lowest endowment $\underline{W}$ and valuation $\underline{\theta}$ prefers consumption to the tank, and thus when y_i is unknown

[8]Because borrowers weigh utility from non-tank consumption against the constant utility of tank consumption, our assumptions on the marginal utility of non-tank consumption are insufficient to ensure that this constraint binds. We could ensure this, however, by assuming $\lim_{c \to 0} u(c) = -\infty$.

[9]Assumption A rules out a particular pathological behavior of the optimal savings and default cutoff functions. The uniformity and wide support of y ensures that utility is single-peaked in savings. Were this condition to fail, it is conceivable optimal savings would move discontinuously. Were it not for the possibility of this discontinuity, the results would hold for any distribution with continuous pdf and finite support. Note also that while we use the example of a uniform distribution, single-peakedness is ensured for a wider class of distributions. One sufficient condition is wide support ($Y > Y + R_T P$) and relative flatness. This condition is satisfied for truncated normal distributions with variance sufficiently large relative to their support, β distributions with small parameters, and certain triangular and trapezoidal distributions.

will not purchase the tank even if somehow assured of receiving the best possible income draw in the next period, $\overline{Y}$.[10]

If farmers borrow to buy a tank, they must make a deposit of at least the lender's requirement $D \in [0, P]$, which earns a gross interest rate R_D. The lender chooses the required deposit, but borrowers take it as a parameter. Potential borrowers may also allocate wealth to savings and they earn gross interest R_D on any saving. Gross savings, including the value of the tank deposit, are denoted S, so for those who borrow to purchase a tank, overall savings $S \geq D$, while those who do not purchase a tank are not subject to this constraint.

To ensure that the model reflects a market with credit-constrained borrowers and allows for the possibility of adverse selection effects on equilibrium outcomes, we make two assumptions. The first is that, for any deposit requirement D, there exist marginal borrowers. Specifically, we assume that the support of W and θ are wide enough that a farmer with period-1 wealth $\underline{W}$ and tank valuation $\underline{\theta}$ will prefer not to borrow even when D=0, and a farmer with period-1 wealth $\overline{W}$ and tank valuation $\overline{\theta}$ will prefer to purchase a tank even when D=P. The second assumption is that at least some borrowers are credit constrained for any deposit requirement D. Specifically, we assume the deposit requirement causes some potential borrowers to be credit constrained if they undertake the tank investment, in the sense of constraining their first period consumption below the level that would be optimal were the deposit not mandated. Since marginal utility is decreasing in consumption and consumption is always higher under default than repayment, a sufficient assumption for there to exist a positive measure of agents who are credit constrained is $u'(\underline{W}) > R_D\mathbb{E}(u'(y_i - R_T P))$. We call borrowers who satisfy $u''(w) > R_D\mathbb{E}(u''(y_i - R_T P))$ "definitely credit-constrained."

To ensure that a nonzero mass of credit-constrained farmers will choose to borrow, we assume that for any D, there is some w_i such that $u'(w_i - D) > R_D\mathbb{E}(u'(y_i + R_D D - R_T P))$, and an agent with initial wealth w_i and tank valuation $\overline{\theta} - \epsilon$ for some $\epsilon > 0$, will choose to borrow a tank. Liquidity constraints make holding wealth in the SACCO costly and are thus consistent with our

[10]This condition is assumed to hold for any reasonable deposit requirement, i.e. any D between 0 and P.

empirical result that greater deposit requirements reduce loan take up dramatically. However, the model also admits individuals who are not credit constrained, and for sufficiently high w_i these individuals will optimally choose $S > D$ (such that higher c_1 could have been chosen). We make final assumptions that $\underline{W}$ and $\underline{Y}$ are large enough so that repayment of loan principal and interest is always feasible ex ante, $\underline{W}R_D + \underline{Y} > R_T P$, and initial payment of the deposit is always feasible $\underline{W} > P$.[11] This assumption is more accurately thought of as a simplification: in the case that wealth levels are such that some farmers may find themselves unable to pay off the tank, our assumptions on u are such that those farmers will never borrow, regardless of the level of D, and thus we can ignore them for the purpose of the model and restrict our attention to those farmers for whom repayment is always feasible ex ante.

There is a limited liability constraint so that if the borrower fails to repay, the only assets which the lender can seize are the pledged deposit D and the tank. If the tank is repossessed, it is sold for δP[12] and the lender is repaid the principal and interest, as well as a repossession fee, K_B. We assume K_B is small enough that the borrower has higher wealth under repossession than under repayment. Leftover proceeds from the sale of the tank, if they exist, are returned to the borrower. We let D_F denote the deposit level at which the principal, interest, and repossession fees are exactly covered by the deposit and tank sale proceeds. We also allow for the possibility that default creates an additional utility cost $M \geq 0$ for borrowers, because it may negatively affect their relationship with the cooperative, which pays a premium price for milk, and which is owned by fellow farmers.

The lender is a monopolist with cost of capital R_D.[13] The lender chooses a required deposit

[11] Farmers also own land, and while land markets are thin and transaction costs for formal sales are high, some sales and rental transactions do take place. (For a discussion of land tenure, see Place and Migot-Adholla, 1998; Barrows and Roth 1990).

[12] The assumption that $\delta \leq 1$ is natural in the case of a scaled-up permanent program, but because tanks were made available at the wholesale price under the program we examine, and because the program was available to only some farmers, the resale value of a repossessed tank could potentially be somewhat greater than P in our context, and indeed one repossessed tank sold for more than the wholesale price. We assume, however, that δ is not so large that potential borrowers can profit by borrowing and allowing repossession ($\delta \leq R_T$ is one sufficient condition for this).

[13] The SACCO may have a small amount of capital available at very low cost from its earnings from transaction fees on payments to farmers, but we will treat its cost of capital at the margin as the 3% per quarter it pays to depositors.

value D^* to maximize expected profits. Reflecting the regulatory cap on interest rates faced by SACCOs, the gross interest rate that the lender charges to borrowers is fixed at R_T. (Empirically, the net interest rate corresponding to R_T is the 1% per month interest rate charged by the SACCO.) We assume that the lender can only offer a single variety of contracts. As we discuss in detail in section 3.4, there are several reasons to believe that a model in which the lender offered a menu of contracts would not reflect empirical reality.

Denote the total cost of repossession to the lender as K.[14] (In the program we examine, farmers were charged a KSh 4,000 repossession fee, but we estimate the full cost of repossession for the lender at KSh 8,500, even excluding intangible costs like the costs of bad publicity and the risk of vandalism, so the empirical case corresponds to $K = 8,500$ and $K_B = 4,000$.) We assume $K_B < K$ as this would reasonably be expected as a property of the optimal contract, since because farmers are risk averse, it will generally not be optimal for borrowers to fully bear the risk associated with negative income shocks that lead to tank repossession. [15]

Below, we first solve potential borrowers' problems of whether to repay conditional on having borrowed and whether to borrow given the D chosen by the lender. We then solve for the profit maximizing D^* for the lender, given borrower behavior.

3.2 The Borrowers' Problem

We first consider the problem of a borrower deciding whether to repay a loan given the deposit D, their tank valuation θ_i,, gross savings S, and second period income y_i. We then solve backwards to the first-period problem of a potential borrower deciding whether to purchase a tank given their wealth and tank valuation.

Proposition 1. *Under the conditions on the distribution of tank valuation assumed earlier, a marginal level of income exists, denoted by $y^R(\theta_i, S, D)$, at which a borrower with valuation θ_i is indifferent*

[14]For example, rental costs for a truck to move the tank, the time of staff members and the security guard who is present at repossessions, management time, the risk of negative publicity or vandalism by a disgruntled borrower.

[15]Moreover, one could imagine that if the contract imposed severe penalties on borrowers during periods when they had negative income shocks and had to allow tank repossession, some borrowers might react in ways that would create large costs for the SACCO, for example vandalizing tanks prior to repossession.

between forgoing consumption in order to make the repayment and allowing the tank to be repossessed. y_i^R is continuously differentiable with respect to all of its arguments, strictly decreasing in θ_i and S, and weakly decreasing in D. When D is such that all repossessions result in negative equity, y_i^R is strictly decreasing in D.[16]

Proof: see appendix.

When choosing whether to repay the loan, the borrower trades off utility from other consumption against utility from the tank. Since utility of consumption is concave, the cost of foregone consumption from repaying the tank loan is decreasing in second-period resources, and thus S and y. Higher θ makes repayment more attractive. y^R defines a repayment probability that is increasing in S. In general, y^R does not need to be within $[\underline{Y}, \overline{Y}]$ for every (θ_i, S, D) tuple; however our assumptions ensure that there do exist such tuples at which repayment occurs.

Corollary 2. *For definitely credit-constrained borrowers who have $S = D$, the threshold level of income for repayment y_i^R is strictly decreasing in the deposit requirement even if negative equity lending does not occur.*

This follows immediately from the fact that y_i^R is decreasing in S. Note that higher D may make the potential credit-constrained borrower worse off overall by constraining c_1, but it increases second period assets, which allows higher c_2. Diminishing marginal utility of consumption then favours repayment once the loan has been made. In the negative equity case, higher S (via D) increases c_2 under repayment,but has no effect on c_2 under repossession, so this effect is even stronger.

Having solved for repayment behavior conditional on borrowing and saving, we can now solve for borrowing and saving behavior as functions of D and w.

Proposition 3. *Potential borrowers will borrow if $\theta_i > \theta^*(D, w_i)$, where θ^* is continuously differen-*

[16]Note for this section's propositions that $\theta^R, y_i^R, \theta^*$, and u may fail to be differentiable at $D = D_F$. This is because utility in the case of repossession may not be differentiable with respect to D at this point. Thus this section's proofs all assume $D \neq D_F$. However, all of the propositions still hold at $D = D_F$ in the following sense: because all of the aforementioned functions are continuous at $D = D_F$ and continuously differentiable around $D = D_F$, if a proposition states, for example, that a function f is weakly increasing in D, we have shown that its derivative is non-positive where it exists, and thus there exists some $\epsilon > 0$ such that for all $D \in (D_F - \epsilon, D_F + \epsilon)$, $f(D) \geq f(D_F)$ if $D < D_F$ and $f(D) \leq f(D_F)$ if $D > D_F$.

tiable in D and w_i for almost all farmers. Among these farmers, θ^* is weakly increasing in D for all farmers, strictly increasing in D for some farmers, and decreasing in w_i. Hence, the repossession rate will be:

$$\rho(D) = \frac{\int_w \int_{\theta^*(D,w)}^{\bar{\theta}} F_Y(y^R(\theta, S, D)) f_\theta(\theta) f_w(w) dw d\theta}{\int_w [1 - F_\theta(\theta^*(D, w))] f_w(w) dw}. \tag{1}$$

Proof: See Appendix.

Potential borrowers compare the expected utility from borrowing to purchase the tank against the expected utility from not borrowing. The expected utility from borrowing depends on the distribution of income draws, and the subsequent optimal choice regarding whether to repay the loan and thus retain the tank. In particular, in any y realisation where borrowers subsequently choose to default on the loan, they would have been better off by not borrowing.

Borrowing to purchase the tank reduces consumption for all income realizations, and potential borrowers thus consider the gains from owning the tank against the cost of foregone consumption. Given the assumptions on the support of the cumulative distribution function $F(\theta_i)$, there will be an interval of wealth levels for which a marginal potential borrower, with valuation $\underline{\theta} < \theta^*(D, w) < \overline{\theta}$, exists. This borrower is indifferent whether to borrow. Potential borrowers with greater valuations will borrow while those with lower valuations will not. There may be some wealth levels below which even those with $\theta = \bar{\theta}$ do not borrow (and some wealth level above which everyone borrows). However, our assumptions ensure that $\theta^*(w) \in [\underline{\theta}, \overline{\theta}]$ for a nonzero mass of potential borrowers. The mass of potential borrowers who decide to borrow is given by

$$\tau(D) = 1 - \int_{\underline{w}}^{\overline{w}} F_\theta(\theta^*(D, w)) f_w(w) dw. \tag{2}$$

Proposition 4. *Potential borrowers with $\theta_i > \theta^*(D, w)$ who are definitely credit constrained will have $S = D$, and would be strictly better off with a lower required deposit. Moreover, if repossessions are negative equity, potential borrowers with a nonzero chance of default are better off with a lower deposit irrespective of whether they are credit constrained. In the case of positive equity or zero probability of default, borrowers who are not credit constrained are indifferent to marginal changes in D. Trivially,* *those with $\theta_i < \theta^*(D)$ are also indifferent to marginal changes in D since they do not borrow.*

Proof: By definition , those who are definitely credit constrained have

$$u'(w_i - D) > R_D \mathbb{E}\left(u'(y_i + R_D D - R_T P)\right). \tag{3}$$

Since $y_i + R_D S - R_T P$ is a borrower's consumption level under repayment, and borrowers have higher period 2 consumption in the case of default than in the case of repayment, $u'(y_i + R_D S - R_T P)$ represents an upper bound on a borrower's marginal period two utility. Thus definitely credit constrained borrowers have

$$u'(c_1(w_i, D)) > R_D \mathbb{E}\left(u'(c_2(w_i, D, \theta_i, S = D)\right). \tag{4}$$

The rest of the proof is immediate from Claim 4 in the proof of proposition 3 (see Appendix A).

$u'(y_i + R_D S - R_T P)$ is trivially decreasing in S for $S > 0$. Furthermore $u'(w_i - S)$ is trivially increasing in S for $S < w_i$. Thus definitely credit constrained borrowers maximize expected utility by setting S=D, and are strictly better off with a lower deposit.

To see the intuition for the impacts of changes in D on non-credit-constrained borrowers , note first that under negative-equity repossession, c_2 is decreasing in D since more wealth is seized when D increases. To see that non-credit-constrained borrowers with $\theta_i > \theta^*$ are indifferent to changes in D when default never occurs or is positive equity, note first that unconstrained borrowers who don't default ultimately recover all of $R_D D$ and thus are unaffected by changes in D. Similarly, unconstrained borrowers who *do* default also recover all of $R_D D$ when $D \geq D_F$. The third result, that those who do not borrow are indifferent to marginal changes in the required deposit, trivially follows from the fact that they do not borrow, and thus do not put down a deposit.

3.3 The Lender's Problem

Having solved the borrower's problem, we can consider a profit-maximizing lender's problem of choosing the optimal required deposit D^*.[17] Denote the lender's net profit per customer who repays a loan without a tank repossession as Π_r, equal to the interest paid by the borrower minus the cost of borrowing the capital to finance the loan, $R_D P$.

$$\Pi_r = (R_T - R_D)P \tag{5}$$

To calculate the payoff to the lender when a borrower fails to repay a loan and the tank has to be repossessed, note that the lender will seize the required deposit and the accrued interest, $R_D D$, sell the repossessed tank for δP, and incur the cost of repossession, K, in addition to the previous outlay on borrowing the capital for the loan, $R_D P$. It will have to return to the borrower any proceeds of the tank sale net of interest and repossession fees, $max\{R_D D + \delta P - R_T P - K_B, 0\}$. Hence, the lender's profit from a loan, Π_d, if the loan is defaulted upon and the tank is repossessed is

$$\Pi_d(D) = \begin{cases} K_B - K + R_T P - R_D P & \text{if positive equity default} \\ \delta P + R_D D - K - R_D P & \text{if negative equity default} \end{cases} \tag{6}$$

Define the *net loss* that the lender incurs from default as their total profit had the loan been repaid, less their profit under repossession, $L_d(D) = \Pi_r - \Pi_d(D)$ (so positive numbers indicate a relative loss).

$$L_d(D) = \begin{cases} K - K_B & \text{if positive equity default} \\ R_T P + K - \delta P - R_D D & \text{if negative equity default} \end{cases} \tag{7}$$

[17] The SACCO has major market power, so for simplicity we model it as a monopolist. While other lenders serve rural Kenya, the SACCO's unique relationship with the farmers in our sample gives it an effective monopoly on this particular type of loan for dairy farmers in the area.

Let $E(\Pi(D))$ denote expected total profits, which the lender maximizes over D. On the intensive margin, an increase in D will (weakly) reduce tank repossession risk for existing borrowers since borrowers will be less willing to allow tanks to be repossessed if they are required to make a larger deposit. Intuitively, this is because a larger deposit means that they have more resources in period $t = 2$ from which to finance consumption, reducing $u'(c_2)$. Under negative equity repossession, default also falls in D as it involves greater foregone consumption. This is the treatment effect of D. On the extensive margin, an increase in the required deposit will reduce the total number of loans and thus both the total profit from loans with no repossession and the expected loss from repossessions. This is the selection effect.

A greater deposit also directly reduces the lender's losses if borrowers fail to repay and proceeds from the tank sale are inadequate to cover the borrower's principal, interest, and tank repossession fee obligations. This never occurs in our data.

The lender's problem is thus given by

$$\max_D E(\Pi(D)) = \max_D \left\{ \int_{\underline{w}}^{\overline{w}} \int_{\theta^*(D,w)}^{\bar{\theta}} \left[\Pi_r - F(y^R(\theta, S^*(w,D), D)) L_d(D) \right] f_w(w) f_\theta(\theta) d\theta dw \right\} \quad (8)$$

where Π_r is the lender's profit per repaid loan and $\int_{\underline{w}}^{\overline{w}} \int_{\theta^*(D,w)}^{\bar{\theta}} \left[F(y^R(\theta, S^*)) \right] f_\theta(\theta) f_w(w) d\theta dw$ is the amount of tank repossessions for a given level of D.

The lender's first order condition for D^* will require equalizing the marginal cost and benefits of raising the required deposit:

$$\frac{\partial E(D)}{\partial D} = \int_{\underline{w}}^{\overline{w}} \Bigg[-\frac{\partial \theta^*}{\partial D} f_\theta(\theta^*) f_w(w) \left[\Pi_r - F(y^R(\theta, S^*, D^*)) L_d(D^*)\right] - \left(\int_{\theta^*}^{\bar{\theta}} \frac{\partial F(y^R(\theta, S^*, D))}{\partial D} f_\theta(\theta) f_w(w) d\theta \right) L_d(D^*) - \left(\int_{\theta^*}^{\bar{\theta}} F(y^R(\theta, S^*, D^*)) f_\theta f_w(w)(\theta) d\theta \right) L_d'(D^*) \Bigg] dw = 0. \quad (9)$$

A proof that this derivative exists and is continuous except at the two points mentioned below is given in appendix A. In maximising profit, the lender will not consider the welfare effects of raising the required deposit on inframarginal customers who would have borrowed in any case. Customers who are credit-constrained or have negative equity suffer a reduction in utility from an increase in the required deposit, which does not factor into the lender's choice of the required deposit rate. This creates a wedge between the private and social benefits from raising the deposit requirement that will tend to make lenders choose deposit requirements that are too high from a social point of view. As long as the lender's profits are continuously differentiable in the deposit requirement at D^* (and thus the FOC holds), reducing the deposit ratio slightly from the lender's profit maximizing level will generate a second-order reduction in profits, but a first order increase in welfare for infra-marginal borrowers.

There are two points at which profits could fail to be continuously differentiable in D. One of these points is the minimal deposit level at which all of the borrowers repay, $\tilde{D}$. Lemma 1 demonstrates that $D^* < \tilde{D}$.

Lemma 1. *The profit-maximizing deposit ratio will be such that there is some non-zero probability of repossession.*

Proof: see appendix.

Intuitively, this lemma follows from the fact that if there were zero repossessions, the lender could lower the deposit, increasing the number of borrowers with a negligible increase in the repossession rate. The other point at which profits could fail to be continuously differentiable in

D is the point, D_F, at which a borrower's net equity after the resale of a tank is zero. Specifically, D_F is the point at which the deposit plus the resale value of the tank just covers the debt on the tank plus interest and the repossession fee, K_B. Increases in D will increase loan recovery in the event of repossession only for D less than D_F. Above D_F, increases in D will affect profits only by changing the probability of tank repossession. By Lemma 1, profits are continuously differentiable with respect to D over the interval $[0, \tilde{D})$ except at D_F.

Thus for $D^* \neq D_F$, a small change in the deposit will create a second-order change in profits for the lender, but a first-order loss in welfare for infra-marginal borrowers. This generates our main result that in the presence of adverse selection generated by heterogeneous tank valuation, the lender chooses deposit requirements that are too stringent from a social point of view.[18]

Proposition 5. *If the profit-maximizing D^* is not D_F, (i.e., if $R_D D^* + \delta P - K_B - R_T P \neq 0$) or 0, then reducing the deposit requirement from the profit maximising level D^* increases social welfare.*

Proof. Social welfare is the sum of borrowers' utilities and lender's profit:

$$E(\Pi(D)) + \mathbb{U}_{total}(D),$$

where $\mathbb{U}_{total}(D)$ is the total expected utility of all the borrowers, given deposit requirement D.

If $R_D D + \delta P - R_T P - K_B \neq 0$ (i.e., $D \neq D_F$) and $D^* \neq 0$, then D^* is characterized by the lender's FOC, since lemma 1 implies $D^* < P$. This implies $\frac{\partial E(\Pi(D))}{\partial D} = 0$. As we showed before, definitely credit-constrained inframarginal borrowers strictly prefer lower deposits, and other

[18] From the standpoint of an unconstrained social planner who seeks to maximize social welfare, the first best would be to allocate tanks to every farmer who has a sufficiently high valuation. Repossessions consume resources, so would never take place. This could be implemented by setting required deposits to zero, and only allowing high valuation farmers to borrow. Further, on account of risk aversion through concave $u(c)$ it is optimal for farmers to be fully insured against income shocks. Consumption utility then becomes deterministic.

One could also consider a mechanism design problem for a planner constrained by lack of information on individual specific tank valuations and income realizations. Such a constrained planner would face the problem of designing a mechanism in which potential borrowers would reveal their tank valuations and income shocks. We will not attempt to solve this mechanism design problem, but the result that a small reduction in the deposit from the profit maximizing level will improve social welfare demonstrates that even a constrained social planner could generate higher welfare than a monopolist.

inframarginal borrowers weakly prefer lower deposits: $\frac{\partial \mathbb{U}_{total}(D)}{\partial D} < 0$. Given the assumptions on the support of w and θ, there will be a nonzero-measure group of inframarginal borrowers who are definitely credit constrained. Potential borrowers who do not borrow will be indifferent to changes in D. Hence the derivative of social welfare with respect to D is negative:

$$\frac{\partial E(D)}{\partial D} + \frac{\partial \mathbb{U}_{total}(D)}{\partial D} = \frac{\partial \mathbb{U}_{total}(D)}{\partial D} < 0.$$

Thus, a social planner that takes borrower welfare into account will set a strictly lower D than would a profit-maximizing lender. □

Since the deposit is greater than socially optimal, the equilibrium fails to achieve the information-constrained social optimum. A social planner without information on borrowers' types could still increase welfare by lowering the deposit.

Note that the lender's first order condition simplifies considerably in the empirically relevant special case where the deposit plus the resale value of the tank is great enough that the borrower has positive equity. Hence, in this case L_d is not a function of D, thus $L_d'(D) = 0$ and the FOC simplifies and can be written as:

$$\frac{\int_{\underline{w}}^{\overline{w}} \frac{\partial \theta^*}{\partial D} f_\theta(\theta^*) f_w(w) dw}{\int_{\underline{w}}^{\overline{w}} \left[\frac{\partial \theta^*}{\partial D} F(y^R(\theta^*, S^*)) f_\theta(\theta^*) - \int_{\theta^*}^{\bar{\theta}} \frac{\partial F(y^R(\theta, S^*))}{\partial D} f_\theta(\theta) d\theta \right] f_w(w) dw} = \frac{L_d(D^*)}{\Pi_r} = \frac{K - K_B}{(R_T - R_D)P}. \tag{10}$$

Here, the left hand side is the ratio of marginal borrowers to marginal tank repossessions. The marginal tank repossession term consists of two components; marginal borrowers having positive default probability, and inframarginal borrowers having increased default probability. In the empirical section we will measure this ratio. At the optimal deposit set by the lender, this ratio equals the ratio of the net costs of a tank repossession to the profits from a successful loan. $L_d > \Pi_r$ and thus this ratio must exceed one, since otherwise even loans that are defaulted upon are profitable overall.

3.4 Discussion

The model could be extended in various ways. One extension which may seem natural is to allow the lender to offer a menu of contracts, with varying interest rate/deposit requirement pairings. We have several reasons to believe that a model with a menu of contracts would not, in fact, be realistic. First, both before and after the experiment, the SACCO only offered a single set of terms for loan contracts. Additionally, the low cap on interest rates drastically limits the scope for variation in contract terms. As discussed below, the 10% inflation rate meant that SACCOs could charge no more than 2% real annual interest. The 3% quarterly nominal rate paid to depositors in the SACCO further limits the range of contracts that would have been profitable–even with no defaults–to a .5 percentage point window. In an equilibrium in which borrowers choose different deposit-interest rate pairs, all borrowers with positive deposits would still experience distortions.

Additionally, we have treated the distribution of income as independent across potential borrowers, but it is also worth considering the case in which $y_i = y_c + y_{ii}$ where y_c is a common shock, for example, due to weather or milk prices, and y_{ii} is an idiosyncratic borrower-specific shock and the common shock is observable, but idiosyncratic shocks are private information for borrowers. In this case, requiring all borrowers to be insured against aggregate risk would reduce repossessions by addressing the moral hazard that arises if borrowers allow tank repossession during periods of negative shocks, even when this is socially inefficient, because they do not face the full costs of repossession. Borrowing decisions will also be improved because borrowers will face more of the full costs of borrowing, including the cost of the risk of default. Hence this will be part of optimal contract design. The optimal response to a common shock is thus insurance, rather than a greater deposit requirement.

The model could also be extended to include guarantor requirements in addition to deposit requirements. Depending on the assumptions, substituting guarantor contracts for deposit requirements might or might not increase access to credit.The assumptions of the model ensure that there are farmers with low enough tank valuations that they choose not to borrow but enough initial wealth that they would not be credit constrained if they did borrow. They also ensure that there are farmers with too little initial wealth to borrow, but high enough tank valuation that they would borrow if they were not credit constrained. Imagine farmers could perfectly contract with each other in the sense of being able to observe each other's initial wealth, tank valuations, and income, and fully enforce all contracts. Then regardless of whether the lender offers a formal guarantor contract, high-wealth, low-valuation farmers would act as guarantors to low-wealth, high-valuation farmers. Even if the lender does not offer a guarantor contract, de facto guarantors could lend low-wealth borrowers money to pay down their deposit. Thus under this assumption, replacing a deposit requirement with a guarantor contract from the lender will not affect loan uptake. Similarly, if farmers cannot contract with each other independent of the existence of a formal guarantor contract, then loan uptake will be the same with or without such a contract, since no one will be willing to extend a guarantee.

On the other hand, if the existence of a formal guarantor contract improves farmers' ability to contract with each other, then such an arrangement will affect outcomes. Formal guarantor agreements could improve farmers' ability to contract with each other if, for example, informal borrowers had the option to default on informal lenders by choosing to use their loan funds for something other than purchasing the tank (i.e, further increasing first-period consumption), and if lenders were then unable to extract repayment in the second period. One scenario in which this would be the case is one in which would-be guarantors were concerned that borrowers might ask for "loans" only to abscond with their borrowed funds and move out of town. This option would be rendered impossible by the existence of a formal guarantor contract which would ensure that the informal borrower actually puts the guarantor's money into buying the tank. Thus formal contracts would incentivize repayment (and mitigate adverse selection of informal borrowers with no intention of repaying) by introducing the cost of a lost tank for those who default.

However, while formal guarantor contracts impact *individual* outcomes in this intermediate case, they need not necessarily increase total demand for loans in general equilibrium. High-wealth, low-valuation farmers who are near-indifferent toward borrowing but do borrow in the case of no guarantor contracts may choose not to borrow if it is possible for them to act as guarantors. Such farmers may prefer to act as guarantors for high-valuation low-wealth borrowers, and in doing so may lose enough period-one wealth to render borrowing no longer worthwhile. The net effect could be that all borrowers who enter the market when guarantor contracts are introduced are offset by guarantors leaving the market, or even that more guarantors leave the market than borrowers enter.

Thus it is an empirical question whether guarantor contracts impact outcomes, as theory would predict different outcomes depending on the nature of contracting in a given empirical context.

4 Project Design and Implementation

This section first discusses features of the loan contracts that were common across treatment arms and then discusses differences across treatment arms that were used to estimate the impact of borrowing requirements on loan take up and on tank repossession and to separately measure moral hazard and adverse selection. (We focus on the main sample and describe some slight differences in the out-of-sample group at the end of the section.)

4.1 Common Loan Features Across Treatment Arms

All farmers in the project were offered a loan to purchase a 5,000-liter water tank. As a bulk purchaser of the tank, the SACCO was able to purchase tanks at the wholesale price and get free delivery to the borrowers' farm. In the main sample, the wholesale price was KSh 4,000 (about $53) below the retail price and the SACCO passed these savings on to borrowers.[19] The price of the tank to the farmers, denoted P in the model, was KSh 24,000 (about $320), or roughly

[19] In this paper we use the dollar to Kenyan Shilling exchange rate at the time of the study which was approximately $1:KSh 75.

20 percent of annual household consumption. Borrowers also incurred installation costs for guttering systems and base construction that averaged about KSh 3,400, or 14% of the cost of the tank.

All farmers received a hand-delivered letter with the loan offer, and were given 45 days to decide whether to take up the loan. All loans were for KSh 24,000 and required an up-front deposit of at least KSh 1,000. The interest rate was 1% per month, charged on a declining balance.[20]

Since the inflation rate is about 10% per annum, the real interest rate was very low. The 1% monthly interest rate is standard for SACCOs but is below the commercial rate. All treatment arms were charged a 1% late fee per month. The interest rate on a late balance was in the ballpark of the market range, but since processing late payments was labor intensive and costly for the lender, the lender was better off when borrowers paid on time. The amount due each month was automatically deducted from the payment owed to the farmer for milk sales. If milk payments fell short of the scheduled loan payment, the farmer was required to pay the balance in cash. Debt service represented 8.4% of average household expenditures and 11.4% of median expenditures at the beginning of the loan term.

Collection procedures for late loans were as follows. When a farmer fell two full months of principal (i.e. KSh 2,000) behind, the SACCO sent a letter warning of pending default and provided two months to pay off the late amount and fees. The letter was hand-delivered to the farmer and followed up with monthly phone reminders. If the late payment was still outstanding after a further 60 days, the SACCO applied any deposits by the borrower or guarantors to the balance.

In arms other than the 100% secured joint liability arm (described below), it is possible that a balance would remain due after this. If a balance still remained, the SACCO gave the farmer

[20]Charging interest on a declining balance is common in Kenya. Borrowers repaid a fixed proportion of the principal each month plus interest on the remaining principal. Borrowers were scheduled to repay KSh 1,000 of their principal back each month for 24 months. In the first month, when farmers had not repaid any of the KSh 24,000 principal, borrowers were scheduled to repay KSh 1240. In the second month, farmers were scheduled to repay KSh 1230; in the third month they were scheduled to repay KSh 1220; and in the final month farmers were scheduled to repay the final KSh 1,000 of their principal and KSh 10 in interest.

an additional 15 days to clear it and waited to see if the next month's milk deliveries would be enough to cover the balance. If not, the SACCO would repossess the tank, charging a KSh 4,000 fee for administrative costs to the borrower from the proceeds of any tank sale. K_B was thus KSh 4,000. The full administrative costs associated with repossessing the tank, including the cost of hiring a truck, staff time, and security, was approximately KSh 8,500, so K should be considered to be at least KSh 8,500 and likely larger, since the lender also risked negative publicity or vandalism from repossession.

The SACCO was the residual claimant on all loan repayments and was responsible for administering the loan. To finance the loans to farmers, Innovations for Poverty Action (IPA) purchased tanks from the tank manufacturer, which then delivered tanks to farmers. The SACCO arm of the cooperative then deducted loan repayments from farmer's savings accounts and remitted these payments to IPA, holding back an agreed administrative fee, structured so as to ensure the SACCO was the residual claimant on loan repayments. IPA financed the loan with a grant from the Bill and Melinda Gates Foundation. To ensure that the cooperative repaid IPA, the cooperative and IPA signed an agreement with tBrookside Dairy Ltd., he milk processing plant, the dairy's customer and the largest private milk producer and processor in the country. The agreement authorized Brookside to make loan repayments directly to IPA out of the milk payments to the cooperative.

4.2 Treatment Arms

As shown in Table 1, farmers were randomly assigned to one of four experimental loan groups, two of which were randomly divided into subgroups after uptake of the loans. One group was offered loans with the standard 100% cash collateral eligibility conditions typically offered by the cooperative (and by most other formal lenders in Kenya, including SACCOs and banks). Specifically, the borrower was required to make a deposit equal to one-third of the loan amount (KSh 8,000) and to have up to three guarantors deposit the other two-thirds of the loan (KSh 16,000) with the SACCO as financial collateral. Guarantors could either be those who already had savings or shares in the cooperative or those willing to make deposits. This group will be denoted Group C (for Cash collateralization).

A second group was offered the opportunity to put down a 25% (KSh 6,000) deposit, and to collateralize the remaining 75% of the loan with the tank itself. This group is denoted Group D (for deposit).

In a third group, the borrower only had to put down 4% of the loan value (KSh 1,000) in a deposit and could find a guarantor to pledge the remaining 21% (5,000 KSh), bringing the total cash pledged against default to 25% of the loan amount. Like the deposit group, 75% of the loan could be collateralized with the tank itself. This group is denoted Group G (for guarantor). Comparing this guarantor group with the 25% deposit group isolates the impact of replacing individual with joint liability.

In a final group, denoted Group A (for Asset collateralization), 96% of the value of the loan was collateralized with the tank itself and only a 4% deposit was required.

In order to distinguish treatment and selection effects of deposit requirements, the set of farmers who took up the 25% deposit loans was randomly divided into two sub-groups. In one, all loan terms were maintained, while in the other, KSh 5,000 of deposits were waived one month after the deposit was made, leaving borrowers with a deposit of KSh 1,000, the same as borrowers in the 4% deposit group, A. The deposit (maintained) and deposit (waived) subgroups are denoted (D^M) and (D^W) respectively.

Similarly, within the guarantor group, in one subgroup loan terms were maintained and in the other, the guarantors had their pledged cash returned and were released from liability in the case of default. Borrowers were informed of this. These guarantor-maintained and guarantor-waived subgroups are denoted (G^M) and (group G^W), respectively.[21]

The selection effect of the deposit requirement on an outcome variable is the difference in the

[21] To avoid deception, at the time the loans were first offered, potential borrowers were told that they would face a 50% chance of having KSh 5,000 of the deposit requirement waived or of having the guarantor requirement waived, respectively.

variable between all borrowers in the 4% deposit group and the 25% deposit group (waived) subgroup. The deposit treatment effect is the difference in a variable's value between the deposit (maintained) and deposit (waived) subgroups. Selection and treatment effects of the guarantor requirement are defined analogously.

5 Data and empirical specifications

In this section we discuss the sampling frame, randomization, data collection, and the empirical approach.

5.1 Sampling, Surveys, and Randomization

A baseline survey was administered to 1,968 households chosen randomly from a sampling frame of 2,793 households regularly selling milk to the dairy. 1,804 farmers were offered loans in accordance with the treatment assignment shown in Table 1. 419 farmers were offered 100% secured joint-liability loans and 510 were offered 4% deposit loans.[22] 460 farmers took out loans.[23]

Midline surveys were administered to all households in the sample, in part to check that tanks had been installed and were in use, but also to collect data on real impacts, including school participation and indicators of time use, based on asking what every household member did in the 24 hours prior to the survey. Subsequently a number of shorter phone surveys were administered, each of which focused on the three months prior to the survey. Time use information was collected from households in all groups,[24] while detailed production data was elicited from

[22]The groups with the least and most restrictive loan forms were the largest because this maximized power in picking up real effects of the loans. Loans were offered in three waves, since it was unknown *ex ante* how many farmers would borrow and the total capital available for purchasing tanks was limited.

[23]Loans were given in three phases, with contractual repayment periods running from March 2010 - February 2012; May 2010 - April 2012; and September 2010 - September 2012. (As discussed below, another set of loans in an out-of-sample group began in February 2012. The total number of loan offers that were prepared was 2616, but 19 of these offers could not be delivered, so the total number of loan offers that were delivered to farmers was 2597. When a household entered into a loan agreement, a water tank was delivered within a period of three months.

[24]Specifically, 1,699 households were interviewed in September 2011: 1,710 in February 2012; and 1,660 in May 2012.

households in the 4% deposit group and the 100% secured joint-liability group.[25] Finally, administrative data from the dairy cooperative was used to construct indicators of loan recovery, repossession, late payment collection actions[26], and early repayment.

Table 2 reports F-tests for baseline balance checks across all treatment groups. Of the 26 indicators presented, one exhibits significant differences across groups at the 5-percent level, and two do so at the 10-percent level. This is in line with what would be expected when the assignment is indeed random.

In part, using the proceeds from the first set of loans, approximately 2600 additional farmers were offered loans between February and April 2012 (following a baseline survey in December 2011), providing an out-of-sample test. These loan offers were for KSh 26,000, due to an increase in the wholesale price of tanks. The monthly interest rate on these loans was 1.2% rather than one percent. We report data from this "out of sample" group on take up rates, loan recovery, and tank repossession outcomes.

These farmers were randomly assigned to receive loan offers requiring only a KSh 1,000 deposit; a KSh 6,000 deposit; or KSh 5,000 from a guarantor plus a KSh 1,000 deposit. These deposits were the same value required in the first set of loan offers but, because the loan offer was for KSh 26,000 rather than KSh 24,000, they were slightly lower as a percentage of the loan amount: i.e. 4% deposit loans; 25% deposit loans; or 21% guarantor, 4% deposit loans. No farmers received the standard Nyala 100% secured joint liability loan offer in this out-of-sample group.

5.2 Empirical Approach

Empirical specifications typically take the form:

$$y_i = \alpha + \beta_A A_i + \beta_D^M D_i + \beta_D^W D_i^W + \beta_G^M G_i + \beta_G^W G_i^W + \varepsilon_i \tag{11}$$

[25]Data was collected from 901 respondents in 2011, and from 863 respondents in February 2012.

[26]E.g. receipt of a letter warning of pending default or reclamation of security deposit

where y_i is the outcome of interest, A_i, D_i^M and G_i^M are dummy variables equal to one if farmer i was randomized to Group A, D, or G, respectively, and D_i^W and G_i^W are equal to one for those members of the deposit and guarantor groups who had their obligations waived *ex post*. The base group in this specification is therefore Group C, the 100% deposit group. For some specifications, we add a vector of individual covariates, X_i.

The overall average impact of moving from a 4% deposit requirement to a 25% deposit or guarantor requirement on take up or tank repossession or any other dependent variable is that given by the differences $\beta_D^M - \beta_A$ and $\beta_G^M - \beta_A$, respectively. The *ex post* randomized removal of deposit and guarantor requirements in groups D^W and G^W allows estimation of the selection and treatment effects of deposits and guarantors. In particular, the selection effects of being assigned to either the deposit or guarantor group are identified by $\beta_D^W - \beta_A$ and $\beta_G^W - \beta_A$, and reflect the extent to which greater deposit requirements or guarantor requirements select borrowers who behave differently than those who take up loans in the 4% deposit group due to differential selection. Under the model, this corresponds to selection of farmers with different tank valuations.

Note that in the notation of the model, the loan take up rate corresponds to $\tau(D) = 1 - \int_{\underline{w}}^{\overline{w}} F(\theta^*(D,w)) f_w(w) dw$ and the repossession rate corresponds to

$$\rho(D) = \frac{\int_w \int_{\theta^*(D,w)}^{\bar{\theta}} F_Y(y^R) f_\theta(\theta) f_w(w) dw d\theta}{\int_w [1 - F_\theta(\theta^*)] f_w(w) dw}. \tag{12}$$

Effects of changing the required deposit D, which we empirically estimate, correspond to changes in the relevant cutoff values. The selection effect corresponds to changes in θ^* while the treatment effect corresponds to changes in y^R.The repayment propensity of marginal farmers who are induced to borrow by being offered a 4% deposit requirement rather than a 25% deposit requirement is equal to the difference in repayment between the 4% and 25% deposit (waived) group, divided by the fraction of borrowers in the 4% group who would only borrow if in that group, e.g., the difference in loan take up rates between the 4% and 25% groups, divided by the take up rate in the 4% group. This corresponds to

$$\frac{\rho(6,000) - \rho(1,000)}{\frac{\tau(1{,}000) - \tau(6{,}000)}{\tau(1{,}000)}} \tag{13}$$

in the model.

The treatment effects of borrowing requirements are identified by comparing loan repayment outcomes for borrowers who have the borrowing requirements maintained with outcomes borrowers who have borrowing requirements waived *ex post*. That is, any treatment effect of the deposit requirement would show up in a difference between β_D^M and β_D^W, while a treatment effect of the guarantors would be observed if β_G^M and β_G^W differed. The treatment effects of the deposit requirement would encompass the incentive effects of borrowing requirements in the model. Specifically, as the required deposit D decreases, the cutoff value $y^R(D, \theta, S)$ rises for some borrowers and is unchanged for others. The effect of moving from $D = KSh\,6,000$ to $D = KSh\,1,000$ corresponds to $\rho(6,000) - \rho(1,000)$ in the model.

6 Loan Take up Rates

Subsection 6.1 discusses the impact of borrowing requirements on loan take up and subsection 6.2 discusses the impact of borrowing requirements on observable borrower characteristics.

6.1 Impact of Borrowing Requirements on Loan Take Up

Allowing farmers to collateralize loans with the assets purchased with the loan greatly expands access to credit. In the original sample, 2.4% of farmers borrow under the standard SACCO contract with 100% cash collateralization (Group C); 27.6% - more than ten times as many - borrow when the deposit is 25% and the rest of the loan can be collateralized with the tank (Group D); and 44.3% borrow when 96% of the loan can be collateralized and only a 4% deposit is required (Group A) (See table 4). This implies that more than 40% of all targeted farmers would like to borrow at the prevailing interest rate and use this technology, but are not doing it because of borrowing requirements. To put this slightly differently, at least $(44.3 - 2.4)/44.3 = 95\%$ of potential tank purchasers would have been prevented from purchasing tankes due to credit constraints under the standard SACCO contract.

Take up rates in the out-of-sample group are broadly comparable to those in the original experiment (Table 4), so in the combined sample, we estimate that 94% of those willing to borrow with a low deposit would be unwilling to borrow under the SACCO's original loan terms. This not only serves as a useful confirmation of the broad patterns in the data, but since farmers in the out-of-sample group had had a chance to see the original lending program in operation, it also provides some reassurance that the original results were not due to misconceptions regarding the water tanks or the loans, or to some unusual period-specific circumstances.[27]

Our second finding is that joint liability does not increase credit access relative to the deposit requirement with individual liability. In the original sample, 27.6% of farmers borrow when they have to put up a 25% deposit themselves (Group D), but only 23.5% borrow when they can ask a friend or relative to put up all but 4% of the value of the loan (Group G) (Table 4). In the out-of-sample group, the point estimates of take up rates is higher in the 21% guarantor, 4% deposit group than in the 25% deposit group, but the difference is still not significant, and in the combined sample, there is almost no difference in take up (as seen in Table 4, columns 2 and 3).

The high elasticity of loan take up with respect to asset collateralization and the lack of response to joint liability points to a potential limitation of traditional joint-liability based microfinance and suggests that addressing barriers to asset collateralization may play an important role in addressing credit constraints.

[27]Point estimates suggest that, averaging across treatment arms, approximately 2.7% fewer members of "out-of-sample "group purchased tanks through the program. The difference is not statistically significant at the 5% level, but it is at the 10% level. One might expect some decline in tank purchases due to the increase in the price of the tank and the increased interest rate.

6.2 Impact of Borrowing Requirements on Observable Borrower Characteristics

Under the model, the lender may use deposit requirements to screen out borrowers with low valuation, who are more likely to default, and it is assumed that the lender cannot directly observe borrowers' tank valuations. This raises the question of whether the borrowers under different arms differ in observables. As shown in Table 3, we find some evidence that borrowers in the 4% arm are not as well off, but overall we find remarkably small differences in observable borrower characteristics among borrowers across arms. Columns (2)-(5) report borrower characteristics by arm. In column (1) these characteristics are reported for the whole sample, including borrowers and non-borrowers in all experimental arms.

Of the 84 possible pair-wise comparisons,[28] we observe statistically significant differences at the 5% level in just four, almost exactly what would be expected under the null hypothesis of no differential selection on observables across treatment arms. Under the model, this suggests that the farmers with tank valuations intermediate between various levels of θ^* associated with different borrowing requirements are not that different on observables, suggesting that it would not be easy to screen borrowers on observables. That said, the variables in which there were significant differences mostly make sense in terms of the model. Borrowers in the 4% deposit group had lower log household assets than those in the 25% collateralized group and had lower log expenditures than those in both the deposit and guarantor groups. It is reasonable to think that poorer households might place less monetary value on a water tank than richer households, and thus might be disproportionately represented among those willing to borrow with a 4% deposit, but not under stricter borrowing requirements.

The starkest difference between the (few) farmers in the 100% secured joint-liability group who chose to borrow and farmers in other arms who chose to borrow is that the former typically chose to borrow only if they already owned a tank. 80% of borrowers already owned a tank, whereas only 43% of borrowers in the full sample owned tanks at baseline. Under the model, this could be interpreted as indicating that those who already owned tanks placed the highest

[28] $3! = 6$ pairs for each of 14 variables.

value on them. Relaxing borrowing requirements induced non-tank owners to buy tanks.

Relative to those who did not accept loan offers, borrowers tended to have more assets, higher per capita expenditure, more milk-producing cows, and more years of education, all of which might plausibly be associated with greater tank valuations under the model.[29] Under the model, differences between borrowers and non-borrowers would be starker than differences among borrowers across arms, if those with very low tank-valuation/initial wealth level pairs, who would not buy even with a low deposit, differ on observables from those with high valuations/wealth levels, but those in an intermediate range of valuation are more similar on observables.

7 Impact of Borrowing Requirements on Loan Repayment

Subsection 7.1 discusses loan recovery and tank repossession, assessing evidence for selection and treatment effects of borrowing requirements. Subsection 7.2 provides a rough calibration of the model, and subsection 7.3 discusses late payment.

7.1 Loan Recovery and Tank Repossession

No tanks were repossessed with 75% asset collateralization under either the 25% deposit (Group D) or the 21% guarantor, 4% deposit condition (Group G) (Table 5). We also observe no tank repossessions when a 25% borrowing requirement was initially imposed and all but 4% of the deposit was later waived. Rates of tank repossession were 0.7% in the 4% deposit, 96% asset collateralized group (Group A). In particular, one tank was repossessed in the original sample and two more were repossessed in the out-of-sample group. In one out of those three cases the borrower paid off arrears and reclaimed the tank after the tank had been repossessed

[29] There were few statistically significant differences between borrowers and non-borrowers in the 100% collateralized group, but there is little power to detect such differences in this group due to the small number of borrowers (see column [2]).

but before it had been resold.[30] Note that in all cases, proceeds from the tank sale were sufficient to fully pay off the principal and interest on the loan. The two tanks that were repossessed and then sold were purchased at KSh 29,000 and KSh 22,000).[31] There were thus no cases of loan non-recovery, defined as a failure to collect principal, interest, and late fee.

Aside from the small 100% secured joint-liability group (Group C), confidence intervals on loan non-recovery rates and on tank repossession rates are fairly tight, so we can reject even very low underlying probabilities of tank repossession. It is clearly impossible to use asymptotics based on the normal distribution when we observe zero or close to zero tank repossessions, but we can create exact confidence intervals based on the underlying binomial distribution. For example, in the combined 4% deposit group, all 431 loans were fully recovered (Table 5). We can therefore reject the hypothesis that the underlying loan non-recovery rate during the period of the loans was more than 0.69 percent. To see this, note that if the true rate was 0.69 percent, then the probability of observing at least one case of loan non-recovery in 431 loans would be $(1 - 0.0069)^{431} = 0.05$. Using a similar approach with three tank repossessions, we can reject the hypothesis that the underlying tank repossession rate during the period was more than 2.02 percent or less than 0.14 percent.

Table 5 displays Clopper-Pearson exact confidence intervals for the rate of tank repossessions and loan non-recovery under the point estimates for each loan type, calculated based on the combined sample, including loans from both the original sample and out-of-sample groups. (Clopper and Pearson, 1934).[32]

[30]We classify this case as a repossession since the costs of repossession were incurred.

[31]The high price relative to the loan value likely reflects the low depreciation rate on tanks as well as the fact that loans were based on the wholesale value of the tank.

[32] A two-sided confidence interval can be calculated for cases with a nonzero number of events. Letting p denote the underlying true probability of an event (tank repossession or loan non-recovery), n the number of loans, and E the number of events, the probability of observing E or fewer events is given by $\sum_{i=0}^{E} \binom{n}{i}(1-p)^{n-i}(p)^i$. The upper limit of the confidence interval is calculated by solving for p in $\sum_{i=0}^{E} \binom{n}{i}(1-p)^{n-i}(p)^i = \frac{\alpha}{2}$, where α is the significance level.

Likewise, the probability of observing E or more events is given by $\sum_{i=E}^{N} \binom{n}{i}(1-p)^{n-i}(p)^i$. The lower limit of the

While 25% borrowing requirements do not seem to select borrowers prone to tank repossession, borrowers selected by 4% requirements are more likely to have tanks repossessed. In particular, we can reject the hypothesis that the repossession rate is the same in the 4% deposit group as among a group combining both forms of 25% cash collateralization (e.g., combining the 25% deposit group and the 21% guarantor, 4% deposit group) at the 5.25% level. (Since the normal approximation is not a good approximation when the probability of an event is close to zero, we used Fisher's exact test to test for a difference between the repossession probabilities.) (As discussed below, after the end of the program, the SACCO began offering 75% asset-collateralized loans on its own, and there have been no tank repossessions. If one treated these observations as part of the sample, the p-value would be below 5%, but since these observations were not randomized and took place in a different time period, it is hard to quantify how much this should increase confidence that underlying tank repossession rates differ between samples with 75% and 96% asset-collateralized loans.) The sample size is inadequate to have this level of confidence for differences between the 96% asset-collateralized group and either the 25% deposit or guarantor group on its own.

There is no evidence of treatment effects of stricter borrowing requirements on tank repossession, since tank repossession rates did not budge off zero when deposit or guarantor requirements were waived *ex post*. We also do not find differences in repossession between individual and joint liability.[33]

7.2 Change in SACCO Policy Following the Program

We can try to assess welfare based on both the observed behavior of the lender following the trial and based on calibrating the model using the data. Starting with the simplest comparison,

confidence interval is calculated by solving for p in $\sum_{i=E}^{N} \binom{n}{i}(1-p)^{n-i}(p)^i = \frac{\alpha}{2}$.

If there are zero events, the lower limit of the confidence interval is zero. In this case, we use a one-sided confidence interval with $\alpha = 0.05$ for the upper bound. In this event, the upper bound can be calculated by solving for p in $(1-p)^n = \alpha$

[33] See Carpena et al. (2013), Karlan and Giné (2014), and Giné et al. (2011) for other work on this issue.

our data suggests that moving from the status quo policy of 100% cash collateralization to loans 75% collateralized with the asset and 25% collateralized with cash could increase loan demand without increasing repossession. This suggests that under the model it would increase both lender and borrower welfare. After the end of the program, once the SACCO had learned about demand for loans and repayment rates under various conditions, it began using its own funds to offer 75% asset-collateralized loans to farmers. (One caveat is that the model abstracts from administration costs of loans, and given the tiny gap between borrowing and lending rates, these are significant. Perhaps in response, the SACCO introduced an appraisal fee on all its loans. For the tank loan, this is equal to KSh 700.)

It seems reasonable to conjecture that the SACCO felt that with the addition of the KSh 700 fee, it was either profitable in expectation to lower the deposit requirement to 25%, or that the costs were low enough that the SACCO could afford to take this step as a way of improving members' welfare. It is not clear whether it would have been profitable to lower the borrowing requirement to 25% without the KSh 700 fee, since the SACCO's margins on lending are very small, and the SACCO most likely incurred additional administrative costs, including costs associated with late payments, by reducing borrowing requirements.

Based on knowledge of salaries in the SACCO and rough estimates of staff time allocation, we estimate that the cost of administering the additional loans would be at least covered by the KSh 700 fee plus the margin the SACCO earns on the difference between the interest rate it pays its depositors and what it charges to borrowers.

Our point estimates suggest that since allowing 75% asset collateralization did not lead to any additional tank repossessions, moving from requiring 100% cash collateralization to 75% asset collateralization would have been profitable during the period we examined. Of course while we observe no extra risk of tank repossession, we cannot reject the hypothesis of an underlying increase in tank repossession of up to 0.32 percent with 75% asset collateralization.

However, since our results raise the question of why the lender did not lower the deposit prior to the experiment, one natural hypothesis is that it did not know how borrowers would respond and feared the downside risk. Given that the SACCO did not choose to offer 96%-asset-collateralization loans, it is not clear from revealed preference alone whether doing so would have been socially optimal. While it is not clear how one should model the objective function of the SACCO, since it is a cooperative, the fact that the cooperative did not lower the borrowing requirement to 4% after learning the results of the experiment suggests that reducing the borrowing requirement was not seen as profit maximizing. If it were profit maximizing, it would have been in the interest of all cooperative members, both borrowers and non-borrowers, to lower the deposit to 4%. While reducing the borrowing requirement to 4% might have benefited borrowers, it would have reduced overall profits and thus harmed non-borrowers, which would include the median voter in the SACCO.

While the model is stylized, and not meant to capture all features of the setting we examine, a rough calibration of the model suggests conclusions similar to those drawn from the revealed preference analysis. Given that moving from 100% cash collateralization to a 25% deposit requirement induced no defaults, the model–abstracting away from administrative costs–directly suggests that this change would increase profits (see the proof of lemma 1). The model also suggests that this change would increase borrower welfare, and would thus be socially optimal. While the model suggests that lowering the deposit requirement below 25% would be socially optimal, it isn't clear what the optimal magnitude would be for this decrease. Given the data, a rough calibration based on the results above and the first order condition for profit maximization suggests that moving all the way down to a 4% deposit requirement would not have been profitable for the SACCO.

As the model's FOC for lenders makes clear, the profit-maximizing deposit level depends not on the average rate of loan recovery and tank repossession, but on the ratio of the marginal additional tank repossessions associated with a change in D to the marginal increase in total loans. To calculate the marginal repossession rate in the combined sample when moving from 25% loans to 4% loans, i.e., D decreasing from KSh$6,000$ to KSh$1,000$, note that the average repossession rate is 0.7% for 4% deposit loans, hence $\rho(1,000) = 0.007\%$, and zero for 25% loans (Table 5, column 2), hence $\rho(6,000) = 0\%$. The take up rate for 4% deposit loans is 41.89%. For 25% deposit loans, the combined sample take up is 23.93%. Thus $\frac{\tau(6{,}000)-\tau(1{,}000))dw}{\tau(6{,}000)} = (41.89-23.93)/41.89 = 42.9\%$. In other words, 42.9% of those who borrow with a 4% deposit are marginal in the sense that they would not borrow with a 25% deposit. Thus our point estimate of the marginal repossession rate is $0.007/.429 = 0.0163$, implying that 1.63% or 1 in 62 of the marginal loans made under a 4% borrowing requirement would lead to a repossession.

Whether a lender would prefer the low deposit depends on whether the marginal profit for an extra loan is more than 1/62nd as much as the repossession costs that the lender bears, $K - K_B$, which we estimate to be at least KSh 4,500. In our context, the additional profits to the lender from a successful loan are likely to be extremely small. In particular, the difference between the interest rate of 3% per quarter that the SACCO pays on deposits and the interest rate of 1% per month that it charges borrowers amounts to only KSh 53 over two years on KSh 18,000 (the amount of the loan, less the 25% deposit, since the borrower earns interest on the deposit). Since interest is paid only on the declining balance, the SACCO makes even less than this on each successful loan. This is less than the expected loss from additional unreimbursed tank repossession costs, which are KSh 4,500/62 = KSh 73. Taking into account the costs to the SACCO of processing loans would further reinforce the conclusion that moving to a 4% deposit would not have been profitable. However, the low expected loss to the lender from additional loans suggest that it is reasonably likely that moving from a 25% deposit requirement to a 4% requirement would be socially desirable, with benefits to borrowers outweighing the small costs to the lender

7.3 Late Payment

Table 6 presents late payment results for the 456 borrowers in the original sample for whom we have complete repayment data[34] Columns (1) to (3) report late payment outcomes during the loan cycle and columns (4) to (6) show payments that were late at the end of the two-year loan

[34]Data on the time of repayment are missing for four borrowers.

cycle. The notes below the table show the p-values on the existence of the selection effect that will drive wedges between private and social optima, as well as on the treatment effects. We first discuss overall effects and then selection and treatment effects.

There is evidence of 'overall 'effects of different treatments. Those offered 100% secured joint-liability loans are much less likely to be ever late than those in any other group, with point estimates of the difference ranging from 43 to 59 percentage points. Moving from a 100% secured joint-liability loan to a 96% asset-collateralized, 4% deposit loan also increases issuance of pending default letters, and it increases late balances at the end of the loan cycle by KSh 222, or about \$3. None of the ten 100% collateralized loans were late at the end of loan. This is a significantly smaller proportion than in the 4% deposit arm, but not than in the 25% deposit or guarantor arms. The extent to which loans were late, however, is tiny, as shown in Column (5) of Table 6, which reports the outstanding late balance at the end of the contractual loan period. Point estimates of the average late balance varied from 46 to 297 KSh, or less than one percent of the loan value. Mean months late in the other groups varied from 0.08 to 0.22 months, or 2-7 days.

There is some suggestive evidence, significant at the 10% level, that stricter deposit and guarantor requirements select borrowers who are less likely to be ever late (Table 6, column 1). The 25% deposit requirements selects borrowers who are 11 ($57 - 46$) percentage points less likely to be late at least once than the 4% deposit loan. Similarly, imposing a guarantor requirement leads to borrowers who are 14 ($57 - 43$) percentage points less likely to be late ever. We find no significant treatment effect of either the deposit or guarantor requirements on being ever late.

For other repayment outcomes, shown in other columns, there is little evidence of a selection effect. Column (2) reports whether a borrower received a pending default letter at some point in the loan cycle (which was typically sent when a farmer was at least two months in arrears). There is no evidence of treatment and selection effects for the deposit group. There is only a borderline significant negative treatment effect of requiring a guarantor ($p = 0.10$). According to column (3), 11 percent of borrowers had security deposits reclaimed, with no significant differences between the treatment arms and the 4% deposit groups. We cannot reject the hypotheses of no treatment effect and of no selection effect.

The model has only three periods, whereas the actual program took place over 24 months. In the last four months of the program, many farmers paid off their loans using their deposits, potentially creating a 'mechanical'effect through which larger deposits reduce late repayment that is not present in the model.[35] For outcomes at the end of the cycle, which may be influenced by the mechanical effect, we see evidence of treatment effects in columns (4)-(6), but not much evidence of selection effects.

Repaid late is a dummy variable equal to 1 if at the contractual maturity date the borrower has an outstanding balance left to pay. Column (6) in Table 6 shows the number of months by which full repayment of the loan was late (any farmers who paid early are counted as being zero months late.). There are significant treatment effects from the 25% deposit on “repaid late''and “months late.''Waiving the deposit increases the chance that borrowers are late at the end of the loan cycle by about 10 percentage points and increases the time by which loans miss the two-year end of the loan cycle by 11% of a month, or just over 3 days. This seems likely to be a mechanical effect. However, since the magnitudes are small, with the difference in the late balance less than 2 USD, these late balances themselves are unlikely to have a major impact on the profitability of lending. There is no evidence for treatment effects of guarantors on late payment outcomes.

Overall, our data does not indicate a consistent pattern in late repayment differences between the 4% and 25% groups. In three of the six measures of lateness, the point estimates indicate that there was greater late repayment in the 25% deposit group and in the other three cases the point estimates indicate there was greater late payment in the 4% loan group.

It is difficult to quantify the extra administrative costs for the SACCO caused by higher rates

[35] Although the existence of such a 'mechanical 'effect would make it difficult to decompose the treatment effect into incentive and mechanical effects, it would not interfere with distinguishing these treatment effects from the selection effects which generate a wedge between profit-maximizing and social welfare maximizing borrowing requirements.

of late payment due to reducing borrowing requirements. The SACCO made very few loans initially and handled much of the bookkeeping manually, in a way that avoided high fixed costs for software and for training staff, but that involved fairly high marginal costs for processing late payments. When payments were late, the SACCO had to manually calculate how late the payments were and send letters. In principle it would be fairly easy to build a software system that would automate this process and send out notices by text message. If a paper copy was needed, it this could be sent with milk transporters who visit farmers every day to collect milk which is delivered to the dairy daily.

One way to get a sense of the cost of late payment is to examine the extent to which the SACCO increased fees when it began making tank loans with a 25% down payment. As noted, the SACCO now applies a KSh 700 initial fee, just under three percent of the value of the loan. This suggests that KSh 700 was enough to cover both any perceived extra expected costs of tank repossession and any extra administrative cost of more frequent late payments caused by moving from the original SACCO contract to a 25% deposit contract.

Another other striking feature of the data is that early repayment was common. It is surprising that so many farmers would forego a close to zero interest loan, since 95 percent of those who bought a tank under the 4% arm were sufficiently credit constrained that they would not purchase a tank under strict borrowing requirements.

Under the standard savings and credit cooperative contract, 90% of people in the 100% secured joint-liability group repaid their loan early. On average, they were 15 months early on a 24 month contract. Even setting aside the eight months of principal in their deposit, they forewent seven months of low interest loan. Of course it is possible that some of these early payers took out new loans through the SACCO's ordinary lending program once their existing loans were paid off. However, since ordinary loans must be fully collateralized through own and guarantors'shares and deposits, paying off a loan early is still giving up access to capital. When 21% of the 25% deposit loan is waived (KSh 5,000 of a KSh 6,000 deposit), many households apply the waived funds almost fully to pay down the principal. They effectively stuck with the status quo of the contract that they signed, thus giving up KSh 5,000 of low-interest loan for more than one year.

8 Real Impact of Changing Borrowing Requirements

While micro-finance organizations often portray their loans as being for investment, there has been debate about the extent to which they are actually used for investment as opposed to for financing consumption (Banerjee et al, 2015). Asset-collateralized loans are potentially more likely to flow towards investment, since lenders making these loans presumably have stronger incentives to ensure that borrowers actually obtain the assets than lenders making un-collateralized loans.

In this section, we show that loosening borrowing requirements for loans to purchase 5,000 liter rainwater harvesting tanks indeed led to increased investment in large tanks, although approximately one-third of the additional loans taken under the looser borrowing requirements may have been used to finance investments which would have taken place in any case. Since the rainwater harvesting tanks represent a new technology, our findings also provide evidence for the idea that access to credit may facilitate technology adoption.

Within the water literature, our findings are consistent with Devoto et al. (2011) in suggesting that expanding access to credit had real effects on access to water, and time use. Difference-in-difference estimates suggest that access to credit to purchase tanks also increased girls ' schooling. Table 8 presents ITT estimates of the impact of assignment to the 4% deposit group, as opposed to the 100% secured joint-liability group, on tank ownership, water storage capacity, cow health, and milk production. These data were collected in a series of survey rounds of farmers in the two groups. We present our results in terms of a simple difference-in-differences framework, comparing these groups before and after loan offers were made. All specifications include survey round fixed effects.

Assignment to the 4% deposit group (Group A) rather than the 100% secured joint-liability group (Group C) increased the likelihood of owning any kind of tank by 17.5 percentage points, an increase of about 35% compared with the counterfactual (note that about 45% of all households had a tank at baseline) and led to an approximately 60 percent increase in household water storage capacity. Both increases are significant at the 1 percent level (as shown in columns 1 and 2). There is a 27% increase in ownership of a tank with 2,500 liter capacity or more. Since the difference in loan take up between Group C and Group A is approximately 40%, we estimate that approximately two-thirds of the additional loans generated new tank investments, while one-third financed purchases that would have taken place in any case.

We find no significant effects on milk production (Table 8). The point estimate is that log production increases by 0.047 points, but this is insignificant, with a t-statistic just under one (column 6).[36]

There is evidence that farmers offered favorable credit terms were more likely to sell milk to the dairy to pay off their loans. Table 9 is based on monthly administrative data from the dairy on milk sales for farmers in all arms of the study. It compares the 4% deposit group (Group A) to all other groups using an ITT approach. Column 4 suggests more Group A farmers sold milk to the dairy. While assignment to the 4% deposit group does not significantly affect the quantity of sales (column 2 and 5), there is some evidence of an effect outside the top five percentiles during the period before loan maturation (although again this effect shows up only in differences, not in levels).

Devoto et al (2011) find that household water connections generated time savings. Table 10 reports estimates of the impact of treatment assignment on time use and schooling for children between the ages of 5 and 16. We present time-use results for the full sample (columns (1) and (2)), and separately for households with (columns (3) and (4)) and without (columns (5) and (6)) piped water. Odd-numbered columns measure time spent fetching water in minutes per day

[36] Table 8, column 4, suggests provision of water tanks reduced sickness among cows. Biologically, it is quite plausible that rainwater harvesting could improve cow health, because it reduced the need for cattle to travel to ponds or streams to drink and thus reduces their exposure to other cattle. However, since there were baseline differences in cow health (as reflected in the coefficient on treatment in this column), it is also possible that this simply reflects mean reversion.

per household member, and even-numbered columns measure time spent tending to livestock, again in minutes per day per household member.

Treated girls spent 3.17 fewer minutes per day fetching water (significant at the 1% level). Boys spent 9.66 fewer minutes per day tending to livestock, (significant at the 10% level) with smaller effects for girls that are not statistically significant (Columns 1 and 2, respectively). The greater access to credit for the purchase of tanks allows females in treatment households to make up nearly all of the gender differential (point estimate -2.22 minutes per day per female, column 1, row 1) in time spent fetching water, significant at the 10 % level. Access to credit to purchase water tanks reduces time spent by girls tending to livestock by 12 min/day in households with piped water. In households without piped water, it reduces time spent by boys tending to livestock by 15 min/day.

Difference-in-difference estimates suggest that greater access to credit also reduced school drop-out rates for girls (Table 11). Observations in each regression are at the individual child level, with standard errors clustered at the household level. Enrollment rates in general were very high at baseline, at about 98%for both boys and girls. Over time, some students dropped out, so these rates were 3-5 percentage points lower in the survey following the loan offers. While access to credit had no impact on boys' enrollment, girls in households assigned to the treatment group were less likely to drop out - the implied treatment effect on girls is 4 percentage points. The effect of treatment on girls' school enrollment, while significant in a difference-in-differences specification, is not significant in levels.

9 Conclusion

In high-income countries, households can often borrow to purchase assets with a relatively small down payment. In contrast, formal-sector lenders in low-income countries typically impose very stringent borrowing requirements. Among a population of Kenyan dairy farmers, we find credit access is greatly constrained by strict borrowing requirements. 42% of farmers borrowed to purchase a water tank when they could primarily collateralize the loan with the tank and only had to make a deposit of 4% of the loan value, but a small fraction (2.4%) borrowed under the lender's standard contract, which required that loans had to be 100% collateralized with pre-existing financial assets of the borrower and guarantors.

Lower borrowing requirements are associated not only with increased borrowing, but with increased investment in the new technology. With regards to repayments, we find that when 75% of the loan could be collateralized with the tanks, all borrowers repaid in full. However, reducing required deposits to 4% of the loan value selected marginal borrowers with a 1.63% rate of failing to pay and having their tanks repossessed (although we see no moral hazard effect). Finally, we find no evidence that substituting guarantors for deposit requirements expands credit access, casting doubt on the extent to which joint liability can serve as a substitute for the type of asset-collateralization common in developed countries.

A simple adverse selection model suggests that since tight borrowing requirements select safer borrowers, profit-maximizing lenders will have socially excessive incentives to choose tight deposit requirements. One policy implication is that legal and institutional barriers to using assets to collateralize debt could potentially have large effects on credit access, investment, and technology adoption. In general, weak property rights or contract enforcement could inhibit collateralization of loans with assets purchased with the loan. In our context, the lender experienced no problems repossessing collateral, and the key barrier to reducing borrowing requirements may have been financial repression in the form of regulatory limits on the interest rate SACCOs can charge customers. Adverse selection implies borrowing limits are too stringent, so regulatory limits on interest rates push in the wrong direction.[37]

A back of the envelope calculation suggests that only a small increase in the interest rate would be needed to offset the cost of the higher tank repossession rate among those who borrow

[37]Note that this conclusion is robust to the possibility that shocks to income might be correlated across borrowers, and that repossession rates might have been higher in bad states of the world. Lenders will have private incentives to consider any such correlations in setting deposit requirements. Moreover, since aggregate shocks are observable, they are better addressed through insurance than through high deposit requirements.

with a 4% down payment.[38]

Financial repression can alternatively be relaxed through upfront fees. After seeing the results of the program, the SACCO introduced the financial innovation of imposing a KSh 700 initial fee and of reducing its deposit requirement to 25%. The fee provides an upper bound on the relaxation in financial repression needed to enable expanded credit access in our setting.

Note also that the SACCO could have easily have covered the administrative costs of the program by retaining some portion of the approximately $50 gap between the wholesale price the SACCO paid for the tanks and the price at which tanks were sold to the farmer. In the program we examined, the tanks were sold to the farmer at the wholesale price, but if the SACCO charged farmers even 20% of the retail price markup, it could have raised this KSh 700 to cover administrative costs. [39]

Increasing the fee for tank repossession could also increase the lender's incentives to reduce borrowing requirements. However, increasing the tank repossession fee would have undesirable risk-sharing properties since farmers will only experience tank repossession if hit by negative income shocks. Limited liability constraints might make it difficult to collect large repossession fees from defaulting borrowers.

The model does not, however, simply suggest removing barriers to asset collateralized loans. Since strict borrowing requirements select more profitable borrowers, the model suggests that profit-maximizing lenders will face socially-excessive incentives for tight borrowing requirements. The market failure identified in the paper creates a potential case for policymakers to encourage less restrictive borrowing requirements by subsidizing such loans - the opposite of

[38] In particular, since one out of 62 marginal borrowers has a tank repossession, and since the extra cost incurred by the SACCO from a tank repossession is approximately KSh 4,500, an increase in profits per loan of KSh 4,500/62 = KSh 72.58 would have been enough to make this worthwhile for the lender in this particular season. This corresponds to an increase in the annual interest rate of approximately three tenths of one percent. In reality, a bigger increase might be needed, since lenders would also have to consider the cost of any additional late payments associated with moving to a 4% deposit ratio.

[39] Indeed, we estimate that 30% of the wholesale-retail markup would be sufficient to cover not only the SACCO's administrative costs of lending to farmers, but also the administrative costs of a larger entity lending to SACCOs. The fairly similar take up rates in the original sample and the out-of-sample group suggest that tank demand is not terribly price elastic, so it seems likely that there would be substantial tank demand even with somewhat higher prices.

existing regulatory policy. Of course, while we have argued that adverse selection will create market failures that lead to excessive borrowing requirements, there is also the danger of a government failure, with large-scale government subsidies to allow lower borrowing requirements turning into favors for the politically connected and possibly triggering bailouts or costly SACCO failures if borrowing requirements dropped too low. Still, it may be possible to isolate particular types of subsidies that would be useful and that would limit the downside risk to the government.

Most SACCOs are small and handle transactions manually, making administrative costs fairly high, and thus discouraging lending. Differences in loan administration efficiency and in administrative costs relative to loan value may partially account for differences in borrowing requirements between low and high-income countries. The development of better ICT technology for the sector could potentially radically lower the cost of handling late payments. Since it seems unlikely that the developer of better software for SACCOs could fully extract the social value of such software, subsidizing the creation of better software for managing SACCO accounts might be welfare improving.

Studies that would shed light on the impact of relaxing borrowing requirements in contexts beyond the context of rainwater harvesting tanks and the dairy industry examined here would constitute public goods to the extent that their results might inform multiple lenders. As noted, a second out-of-sample test in Kenya after the initial study generated similar results to those presented above. A similar pilot program was implemented by the J-PAL Africa policy team in Rwanda. In the first phase, 43 out of about 160 farmers took up the loan, with only one default. Since the second Nyala test, the lender has extended the program, using its own resources, and has also experienced high repayment rates. Thirteen SACCOs have chosen to implement similar programs without subsidies. Additionally, following the results of this study, a major commercial bank in Kenya (Equity Bank) has started a program with another tank manufacturer in which it is making loans to finance tank purchases.

More ambitiously, policymakers could offer to insure borrowers and/or lenders against observable negative shocks to the state of the world, such as droughts or price declines, potentially just offering bridging loans that would allow lenders to defer payment during such periods, with the loans still incurring interest.

One area we hope to explore in future work is whether prospect theoretic preferences could help explain why demand for loans is so responsive to the possibility of collateralizing loans using assets purchased with the loan and why repayment rates are so high. Under prospect theory (Kahneman and Tversky, 1979), people value gains relative to a reference point less than they disvalue losses relative to that reference point. Prospect theoretic agents may be averse to pledging an existing asset as collateral to obtain a new asset like a water tank, so they would have low take up rates when high deposits are required. However, prospect theoretic agents would be more likely to take up loans if they can use assets purchased with the loan as collateral, because this limits risk to existing assets. Once the tank is purchased, their reference point will shift, creating a strong incentive for prospect-theoretic farmers to retain possession. This could account for the very high repayment rates.

Prospect theory can also potentially explain the finding that the largest difference in observable characteristics between those borrowing in the 100% secured joint-liability group and those borrowing in the other arms is that 80% of borrowers in the 100% secured joint-liability loan arm already owned tanks. This is surprising from a diminishing returns perspective, but it is consistent with loss aversion, since most of the existing tanks are stone or metal and thus susceptible to loss from cracking or rust. Prospect theory might also help explain why farmers who made 25% deposits and later had them waived often simply applied the waived deposit toward paying down the loan early.

Adams, William, Liran Einav, and Jonathan Levin. 2009. "Liquidity Constraints and Imperfect Information in Subprime Lending."*American Economic Review*, 99 (1), 49-84.

Anderson, Michael (2008). "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Presschool, and Early Training Projects," *Journal of the American Statistical Association*, 103(484), pp. 1481-1495.

Attanasio, Orazio, Britta Augsburg, Ralph De Haas, Emla Fitzsimons and Heike Harmgart (2015). "The Impacts of Microfinance: Evidence from Joint-Liability Lending in Mongolia," *American Economic Journal: Applied Economics*, 7(1), pp. 90-122.

Banerjee, Abhijit, Esther (2005)."Growth Theory through the Lens of Development Economics," *Handbook of Economic Growth*, pp. 473-552.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster and Cynthia Kinnan (2010)."The miracle of microfinance? Evidence from a randomized valuation," Working paper, MIT.

Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman (2015). "Six Randomized Evaluations of Microcredit: Introduction and Further Steps," *American Economic Journal: Applied Economics*, 7(1), pp. 1-21.

Barrows, Richard, and Michael Roth. (1990). "Land tenure and investment in African agriculture: Theory and evidence," *The Journal of Modern African Studies*, 28(02), pp. 265-297.

Beck, Thorsten, and Asli Demirguc-Kunt (2006). "Small and medium-size enterprises: Access to finance as a growth constraint, " *Journal of Banking and Finance*, 30(11), pp. 2931-2943.

Besley, Timothy J. (1994). "How do market failures justify interventions in rural credit markets? , " *The World Bank Research Observer*, 9(1), pp. 27-47.

Besley, Timothy J. and Stephen Coate (1995). "Group Lending, Repayment Incentives and Social Collateral, " *Journal of Development Economics*, 46(1), pp. 1-18.

Billingsley, P. (1995) Probability and Measure. Wiley, p. 212.

Carpena, Fenella, Shawn Cole, Jeremy Shapiro and Bilal Zia (2013). "Liability Structure in Small-scale Finance. Evidence from a Natural Experiment, " *World Bank Economic Review*, 27(3), pp. 437-69.

Casaburi, Lorenzo, and Rocco Macchiavello (2014). "Reputation, Saving Constraints, and Interlinked Transactions: Evidence from the Kenya Dairy Industry," Working paper, Warwick

Casey, Katherine, Rachel Glennerster, and Edward Miguel (2012). "Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan," *Quarterly Journal of Economics*, forthcoming.

Clingingsmith, David, Asim Khwaja, and Michael Kremer (2009)."The Impact of the Hajj: Religion and Tolerance in Islam's Global Gathering," *Quarterly Journal of Economics*, 124(3), pp. 1133-1170.

Clopper, C.J and Egon S. Pearson (1934). "The use of confidence or fiducial limits illustrated in the case of the binomial," *Biometrika*, (26), pp. 404-413.

Crépon, Bruno, Florencia Devoto, Esther Duflo and William Parienté (2011). "Impact of microcredit in rural areas of Morocco: evidence from a randomized evaluation," Working paper, MIT.

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

de Mel, Suresh, David McKenzie and Christopher Woodruff (2009). "Are women more credit constrained? Experimental evidence on gender and microenterprise returns," *American Eco-*

nomic Journal: Applied Economics, 1(3), pp. 1-32.
Devoto, Florence, Esther Duflo, Pascaline Dupas, William Parienté and Vincent Pons (2011). "Happiness on tap: piped water adoption in urban Morocco," Working paper, MIT.
Djankov, Simeon, Caralee McLiesh and Andrei Shleifer (2007). "Private credit in 129 countries," *Journal of financial Economics*, 84(2), pp. 299-329.
Duflo, Esther, Michael Kremer, Jonathan Robinson (2008). "How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya," *American Economic Journal*, 98(2), pp. 473-552.
Enterprise Survey Data (2015). "Finance," World Bank.
Esrey, Steven A. (1996). "Water, waste, and well-being: a multicountry study," *American Journal of Epidemiology*, 143(6), pp. 608-623.
Fafchamps, Marcel, David McKenzie, Simon Quinn and Christopher Woodruff (2011). "When is capital enough to get female microenterprises growing? Evidence from a randomized experiment in Ghana," NBER Working Paper Series, Working Paper 17207.
Feder, Gershon and David Feeny (1991). "Land tenure and property rights: theory and implications for development policy," *The World Bank Economic Review*, 5(1), pp. 135-153.
Feder, Gershon, Tongroj Onchan and Tejaswi Raparla (1988). "Collateral, guaranties and rural credit in developing countries: evidence from Asia," *Agricultural Economics*, 2(3), pp.231-245.
Feigenberg, Ben, Erica Field and Rohini Pande (2012). "The economic returns to social interaction: experimental evidence from microfinance," Working paper, Harvard.
Field, Erica, Rohini Pande, John Papp and Natalia Rigol (2010). "Term Structure of Debt and Entrepreneurial Behavior: Experimental Evidence from Microfinance", Working paper, Harvard.
Giné, Xavier and Dean Karlan (2011). "Group versus individual liability: short and long term evidence from Philippine microcredit lending groups," Working paper, Yale.
Giné, Xavier, Karuna Krishnaswamy and Alejandro Ponce (2011). "Strategic Default in Joint Liability Groups: Evidence from a Natural Experiment in India," mimeo, World Bank.
Giné, Xavier, Jessica Goldberg, and Dean Yang (2012)."Credit Market Consequences of Improved Personal Identification: Field Experimental Evidence from Malawi", *American Economic Review*, forthcoming.
Hermes, Niels, and Robert Lensink (2007). "The empirics of microfinance: what do we know?" *The Economic Journal*, 117(517), F1-F10.
Kahneman, Daniel and Amos Tversky (1979). "Prospect theory: An analysis of decision under risk, " *Econometrica: Journal of the Econometric Society*, 47(2), pp. 263-291.
Karlan, Dean and Jonathan Zinman (2009). "Observing unobservables: identifying information asymmetries with a consumer credit field experiment," *Econometrica*, 77(6), pp. 1993-2008.
Karlan, Dean and Xavier Giné (2014). "Group versus Individual Liability: Short and Long Term Evidence from Philippine Microcredit Lending Groups," *Journal of Development Economics*, 107, pp. 65-83.
Kling, Jeffrey, Jeffrey Liebman, and Lawrence Katz (2007). "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75(1), pp. 83-119.
Kremer, Michael, Jean Lee, Jonathan Robinson and Olga Rostapshova (2011). "The return to capital for small retailers in Kenya: evidence from inventories," Working paper, World Bank.
Kristjanson, Patricia M., Brent Murray Swallow, Gareth Rowlands, R.L. Kruska and P.N. De Leeuw (1999). "Measuring the costs of African animal trypanosomosis, the potential benefits

of control and returns to research," *Agricultural systems*, 59(1), pp. 79-98.

La Porta, Rafael, Florencio Lopez-De-Silanes, Andrei Shleifer and Robert W. Vishny (1997). "Legal Determinants of External Finance," /emphThe Journal of Finance, 52(3), pp. 1131-1150.

Luoto, Jill, Craig McIntosh, Bruce Wydick (2007). "Credit Information Systems in Less Developed Countries: A Test with Microfinance in Guatemala," *Economic Development and Cultural Change*, 55(2), pp. 313-334.

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

Microcredit Summit Campaign (2014). Data reported to the campaign in 2013,

Morduch, Jonathan (1999). "The microfinance promise, " /emphJournal of Economic Literature, 37(4), pp. 1569-1614.

Nicholson, Mark (1987). "The effect of drinkingfrequency on some aspects of the productivity of Zebu cattle," *Journal of Agricultural Science*, 108(1), pp. 119-128.

Peden, Don, Faisal Ahmed, Abiye Astatke, Wagnew Ayalneh, Mario Herrero, Gabriel Kiwuwa, Tesfaye Kumsa, Bancy Mati, A. Misra, Denis Mpairwe, Girma Tadesse, Tom Wassenaar and Asfaw Yimegnuhal (2007). "Water and livestock for human development," in D. Molden (ed.), *Comprehensive assessment of water management in agriculture*, Oxford University Press, Oxford, pp. 485-514.

Place, Frank and Shem E. Migot-Adholla. (1998). "The economic effects of land registration on smallholder farms in Kenya: evidence from Nyeri and Kakamega districts," *Land Economics*, Vol. 73, No. 3., pp. 360-373.

Rajan and Zingales. (1998) "Financial Dependence and Growth," *The American Economic Review*, Vol. 88, No. 3. (Jun., 1998), pp. 559-586.

The SACCO Societies Regulatory Authority. (2013) "Sacco Supervision Annual Report 2013 (Deposit Taking Saccos)," `http://www.sasra.go.ke/index.php/resources/publications#.VqfjDvkrI1k`.

Staal, S.J., I. Baltenweck, O. Bwana, G. Gichungu, M. Kenyanjui, B. Lukuyu, K. Muriuki, H. Muriuki, F. Musembi, L. Njoroge, D. Njubi, A. Omore, M. Owango and W. Thorpe (2001). "Dairy systems characterisation of the greater Nairobi milk shed," Smallholder Dairy Project Research Report.

Stiglitz, Joseph and Andrew Weiss (1981). "Credit rationing in markets with imperfect information," *American Economic Review*, 79(1), pp. 159-209.

Wang, Xia, and Paul Hunter (2010). "A systematic review and meta-analysis of the association between self-reported diarrheal disease and distance from home to water source," *The American Journal of Tropical Medicine and Hygeine*, 83(3), pp. 582-584.

Water for People (No date), About. `http://www.waterforpeople.org/about/`.

White, Gilbert F., David J. Bradley, and Anne U. White (1972). Drawers of water. Chicago: University of Chicago Press.

WHO and UNICEF (2010), Progress on Sanitation and Drinking Water: 2010 update. World Health Organization (WHO) and

UNICEF: Joint Monitoring Programme for Water Supply and Sanitation.

World Bank (2012), World development indicators onlinedatabase, `http://data.worldbank.org/data-catalog/world-development-indicators`.

World Bank (2014), Global financial inclusion database, `http://datatopics.worldbank.org/financialinclusion/`.

Zeller, Manfred, Aliou Diagne, and Charles Mataya (1998). "Market access by smallholder farmers in Malawi: Implications for technology adoption, agricultural productivity and crop income," *Agricultural Economics*, 19(1), pp. 219-229.

A Proofs for the Model Section

Proposition 1.

Under the conditions on the distribution of tank valuation assumed earlier, a marginal level of income exists, denoted by $y^R(\theta_i, S, D)$, at which a borrower with valuation θ_i is indifferent between forgoing consumption in order to make the repayment and allowing the tank to be repossessed. y_i^R is continuously differentiable with respect to all of its arguments, strictly decreasing in θ_i and S, and weakly decreasing in D. When D is such that all repossessions result in negative equity, y_i^R is strictly decreasing in D.

Proof. If the borrower repays the lender, her second-period utility is

$$U_{2,r}(y_i, S; \theta_i) = \theta_i + u(y_i + R_D S - R_T P), \tag{14}$$

that is, the benefit of the tank, θ_i, plus the consumption utility from resources remaining once the loan principal and interest $R_T P$ are repaid. Consumption is financed from the remainder of the gross returns from savings and the income draw. To derive the utility of a borrower who does not repay the loan and allows the tank to be repossessed, first consider the net proceeds the borrower receives from the sale of the tank. In the event of repossession, a borrower will receive their net equity in the tank (from the lender's point of view) if it is positive and will lose the required deposit if their net equity is negative. The net equity of the borrower is equal to the total value of the tank and the required deposit, $R_D D + \delta P$, minus the total claims of the lender in the event of default, $R_T P + K_B$. Hence, in the event of default, the borrower faces a financial cost from default of $\min\{R_T P + K_B, R_D D + \delta P\}$. Since the borrower's assets before repossession have value $R_D S + \delta P$, a defaulting borrower receives net proceeds from the first period of $\max\{R_D S - (R_T - \delta)P - K_B, R_D(S - D)\}$, and has total second-period utility of

$$U_{2,d}(y_i, S, D; \theta_i) = u(\max\{y_i + R_D S + \delta P - R_T P - K_B, y_i + R_D(S - D)\}) - M \tag{15}$$

where the final term captures the disutility from harming their relationship with the SACCO M. Consumption is financed by the period two endowment y_i, any net proceeds from the sale of the tank, and any non-deposit savings. Loan defaults only occur when low income is realized, since high-income borrowers will have a reduced marginal utility of consumption and thus prefer to repay the loan, and potential borrowers will not borrow if they know that they will allow the tank to be repossessed for all income realizations.[40] Note also that whether any default would be positive or negative equity is determined prior to and independently of the period two income draw, depending only on whether $\delta P + R_D D \geq R_T P + K_B$. Comparing the utilities from repayment and default yields the condition for repossession, conditional on borrowing at $t = 1$. A borrower will only default upon the loan and allow the tank to be repossessed if she earns low enough period-two income that the utility from defaulting exceeds the utility from repayment:

$$U_{2,repossession}(y_i, S; \theta_i) > U_{2,repay}(y_i, S; \theta_i). \tag{16}$$

[40]Recall that the the borrower receives no utility benefit from the tank if it is repossessed, but still incurs the repossession fee.

Under the conditions on the distribution of tank valuation assumed earlier, a marginal level of income exists, denoted by $y^R(\theta_i, S, D)$, at which a borrower with valuation θ_i is indifferent between repaying the loan and allowing the tank to be repossessed. Since $u'(c)$ is decreasing, and default gives higher consumption, repayment is preferred at any higher y_i. First consider the case where D is such that any loan default involves positive equity. In this case y^R is defined by:

$$\theta_i + u(y^R + R_D S - R_T P) = u(y^R + R_D S + \delta P - R_T P - K_B) - M. \tag{17}$$

Since

$$\theta_i + u(y^R + R_D S - R_T P) - u(y^R + R_D S + \delta P - R_T P - K_B) + M \tag{18}$$

is continuously differentiable, and has nonzero derivative with respect to y^R (this follows from the fact that $y^R + R_D S - R_T P < y^R + R_D S + \delta P - R_T P - K_B$), the continuous differentiability of y^R follows from the implicit function theorem.

Clearly, higher θ_i allows a higher consumption differential between default and repayment at the point of indifference. This translates to a lower y^R. Letting $c_{2,r}$ denote second period consumption in the case of repayment and $c_{2,d}$ in the case of default, total differentiation gives:

$$d\theta_i + \left(u'(c_{2,r}) - u'(c_{2,d})\right)(dy^R + R_D dS) = 0 \tag{19}$$

$$\Rightarrow \frac{\partial y^R}{\partial \theta_i} = -\frac{1}{u'(c_{2,r}) - u'(c_{2,d})} < 0 \tag{20}$$

$$\Rightarrow \frac{\partial y^R}{\partial S} = -R_D < 0 \tag{21}$$

Separately, in the case where negative equity repossession can occur, y^R is defined by:

$$\theta_i + u(y^R + R_D S - R_T P) = u(y^R + R_D(S - D)) - M \tag{22}$$

Again, continuous differentiability of y^R is direct from the implicit function theorem. By total differentiation:

$$d\theta_i + u'(c_{2,r})(dy^R + R_D dS) - u'(c_{2,d})(dy^R + R_D(dS - dD)) = 0 \tag{23}$$

$$\Rightarrow \frac{\partial y^R}{\partial \theta_i} = -\frac{1}{u'(c_{2,r}) - u'(c_{2,d})} < 0 \tag{24}$$

$$\Rightarrow \frac{dy^R}{dS} = -R_D < 0 \tag{25}$$

$$\Rightarrow \frac{dy^R}{dD} = -\frac{u'(c_{2,d})}{u'(c_{2,r}) - u'(c_{2,d})} R_D < 0 \tag{26}$$

These results reflects that, for a borrower with given θ_i who has positive equity, the decision to repay only depends on their wealth, and thus higher S reduces y^R. In the negative equity case,the direct effect of D (holding S constant) is to decrease c_2 under default, again reducing y^R. Higher θ_i increases the benefits of repayment, and thus justifies incurring the greater foregone consumption utility associated with lower y_i. □

Proposition 3. *Potential borrowers will borrow if $\theta_i > \theta^*(D, w_i)$, where θ^* is continuously differentiable in D and w_i for almost all farmers. Among these farmers, θ^* is weakly increasing in D for all farmers, strictly increasing in D for some farmers, and decreasing in w_i. Hence, the repossession rate will be:*

$$\frac{\int_w \int_{\theta^*(D,w)}^{\bar{\theta}} F_Y(y^R(\theta, S, D)) f_\theta(\theta) f_w(w) d\theta dw}{\int_w [1 - F_\theta(\theta^*(D))] f_w(w) dw}. \tag{27}$$

Proof. At period $t = 1$, potential borrowers i will borrow if expected utility from not borrowing is lower than expected utility from borrowing. The utility potential borrowers receive if they do not borrow, denoted as $\bar{U}$, is equal to their consumption utility across the two periods $u(c_1^0) + u(c_2^0)$ where second-period consumption is $c_2^0 = (w - c_1^0)R_D + y_i$. This is evaluated at the consumption profile that maximises expected utility, characterised by the Euler equation $u'(c_1^0) = R_D \mathbb{E}(u'(c_2^0))$. Borrowers, knowing their θ_i, will allow their tanks to be repossessed if they have a low income realization, $y_i \leq y^R(\theta_i, D)$. Then, the borrower's expected utility from borrowing will be equal to the expectation over all possible income outcomes that include income realizations that lead to default, $U_d(y_i, D; \theta_i)$, and that lead to keeping the tank, $U_r(y_i, D; \theta_i)$. This will exceed the expected utility from not borrowing, and thus the individual will choose a savings amountt S (and thus a c_1) and borrow, if

$$U^*(D, w_i, \theta_i) = \max_{S \geq D} \left(\int_{\underline{Y}}^{y_i^R} U_d(y_i, S, D; \theta_i, w_i) f_Y(y_i) dy_i + \int_{y_i^R}^{\bar{Y}} U_r(y_i, S, D; \theta_i, w_i) f_Y(y_i) dy_i \right) \geq \overline{U(w_i)}. \tag{28}$$

Note that the value $U_d(y_i, S, D; \theta_i, w_i)$ depends on whether D is sufficiently large to preclude negative equity repossession. Since we consider only borrowers who can always repay the tank, the utility cost of repayment for a borrower of a given wealth level with a given deposit requirement is finite. Thus for any borrower we consider, there is some $\theta_{repay} \in [0, \infty)$ for which she repays the loan with nonzero probability. As is shown below, utility from borrowing is continuous, increasing, and weakly convex in θ whenever there is a nonzero probability of repayment (that is, whenever $\theta > \theta_{repay}$). Furthermore, borrowers who do not value tank ownership are strictly worse off borrowing. Thus, for all $w \in [\underline{W}, \overline{W}]$, there exists a marginal tank valuation, denoted by $\theta^*(D, w) \in [0, \infty)$, where a potential borrower with wealth w would be indifferent regarding whether to borrow. $\theta^*(D, w)$ need not be within the support of θ for all w, but under our assumptions, for every $D \in [0, P]$ there is a range of w for which $\theta^*(D, w) \in [\underline{\theta}, \overline{\theta}]$. Higher valued potential borrowers will borrow while lower valued potential borrowers will not. Thus, the mass of potential borrowers with a fixed w who borrow is given by $1 - F_\theta(\theta^*(D, w))$, with the mass of defaults given by $\int_{\theta^*(D,w)}^{\bar{\theta}} F_Y(y^R(\theta, S) f_\theta(\theta) d\theta$. Integrating over the distribution of w gives the population borrowing and default rates. To show the proposition's claims about the derivatives of θ^*, we proceed in five steps. First, we show that overall utility given S, D, w and θ is continuously differentiable in all of its arguments. Next we use that fact to demonstrate that $S^*(D, w, \theta)$, the optimal amount of savings, is continuously differentiable in its arguments for almost all farmers. From there, we show that overall utility from borrowing and optimizing savings, $U^*(D, w, \theta)$ is continuously differentiable in all of its arguments almost everywhere. Having shown this, we prove proposition 4, that U^* is weakly decreasing in D for all farmers and strictly decreasing in D for some farmers even in the case of positive equity loans. Lastly, we use the last two facts to prove the remaining parts of proposition 3.

Claim 1: *Overall utility from borrowing $U_{overall}(\theta, w, S, D)$, given a savings level S, is continuously differentiable in each of its arguments.*

Proof. Overall utility is given by

$$U_{overall} = u(w_i - S) + \int_{\underline{Y}}^{y^R(S,D,\theta)} [u(c_{2,default}(S, D, y)) - M] f_y(y) dy + \int_{y^R(S,D,\theta)}^{\overline{Y}} [u(y + R_D S - R_T P) + \theta] f_y(y) dy. \quad (29)$$

The proofs of claims one and two assume that $y^R \neq \overline{Y}$ and $y^R \neq \underline{Y}$. We will show at the end of the proof of claim two that these cases occur for only a zero-measure set of farmers.

The right hand side of equation 28 is trivially differentiable in w_i, with derivative $u'(w_i - S)$, which is continuous. By proposition 1, y^R is continuously differentiable in all of its arguments. Lastly, u is continuously differentiable in c_2, and in cases of both repayment and repossession, c_2 is continuously differentiable with respect to S and D. Thus by Leibniz' rule, the expression is differentiable with respect to S, D, and θ. Noting that the envelope theorem gives that changes in y^R are second-order, we have

$$\frac{\partial}{\partial \theta} U_{overall} = \int_{y^R(S,D,\theta)}^{\overline{Y}} f_y(y) dy = 1 - F(y^R). \quad (30)$$

$$\frac{\partial}{\partial S} U_{overall} = -u'(w_i - S) + R_D \bigg(\int_{\underline{Y}}^{y^R(S,D,\theta)} u'(c_{2,default}(S, D, y)) f_y(y) dy \quad (31)$$

$$+ \int_{y^R(S,D,\theta)}^{\overline{Y}} u'(y + R_D S - R_T P) f_y(y) dy \bigg). \quad (32)$$

$$\frac{\partial}{\partial D} U_{overall} = \frac{\partial c_{2,default}}{\partial D} \int_{\underline{Y}}^{y^R(S,D,\theta)} u'(c_{2,default}(S, D, y)) f_y(y) dy. \quad (33)$$

The continuity of each of these expressions is immediate from the fact that u' is continuous and the fundamental theorem of calculus.[41] □

Claim 2: *Optimal savings $S^*(D, w, \theta)$ is continuously differentiable in all of its arguments for almost all farmers.*

[41] Attentive readers might notice that $\frac{\partial c_{2,default}}{\partial D}$ is not continuous at $D = D_F$. Recall, however, that for the purpose of these propositions, we assume $D \neq D_F$.

Proof. We have

$$\frac{\partial^2}{\partial S^2}U_{overall} = u''(w_i - S) + R_D\Bigg(R_D\int_{\underline{Y}}^{y^R(S,D,\theta)} u''(c_{2,default}(S,D,y))f_y(y)dy \tag{34}$$

$$+\frac{\partial y^R}{\partial S}u'(c_{2,default}(S,D,y^R))f_y(y^R) + R_D\int_{y^R(S,D,\theta)}^{\overline{Y}} u''(y + R_D S - R_T P)f_y(y)dy \tag{35}$$

$$-\frac{\partial y^R}{\partial S}u'(y^R + R_D S - R_T P)f_y(y^R)\Bigg). \tag{36}$$

Recall from proposition 1 that $\frac{\partial y^R}{\partial S} = -R_D$. Furthermore, since $Y \sim Unif[\underline{Y}, \overline{Y}]$, $f_y(y) = (\overline{Y} - \underline{Y})^{-1}$ for all $y \in [\underline{Y}, \overline{Y}]$, and zero otherwise. Combining these facts with the continuity of u″ and the fundamental theorem of calculus, we derive, for $y^R \in [\underline{Y}, \overline{Y}]$,

$$\frac{\partial^2}{\partial S^2}U_{overall} = u''(w_i - S) + R_D^2 f_y(y^R)\Bigg(u'(\overline{Y} + R_D S - R_T P) - u'(c_{2,default}(S,D,\underline{Y}))\Bigg). \tag{37}$$

Note that this expression is continuous in S, D and y^R. By the assumption that $\overline{Y} + R_D S - R_T P > c_{2,default}$, the concavity of u yields that both terms in this expression are negative. For $y \notin [\underline{Y}, \overline{Y}]$, the right hand side of equation 33 is

$$u''(w_i - S) + R_D^2\Bigg(\int_{\underline{Y}}^{y^R(S,D,\theta)} u''(c_{2,default}(S,D,y))f_y(y)dy + \int_{y^R(S,D,\theta)}^{\overline{Y}} u''(y + R_D S - R_T P)f_y(y)dy\Bigg). \tag{38}$$

This expression is also continuous, and trivially negative. Thus,

$$\frac{\partial^2}{\partial S^2}U_{overall} < 0. \tag{39}$$

The concavity of $U_{overall}$ with respect to S, along with the assumptions that $\lim_{c\to 0} u'(c) = \infty$ and $\lim_{c\to\infty} u'(c) = 0$ and the continuity of $\frac{\partial U_{overall}}{\partial S}$ ensure that there is some unique (possibly negative) $S_{max} \in \mathbb{R}$ such that

$$\frac{\partial U_{overall}}{\partial S}(S_{max}) = 0. \tag{40}$$

We have from equation 30 and the fact that $c_{2,default}$ is continuously differentiable with respect to D when $D \neq D_F$ that $\frac{\partial U_{overall}}{\partial S}$ is differentiable in D and

$$\frac{\partial^2 U_{overall}}{\partial S \partial D} = R_D \bigg(\frac{\partial c_{2,default}}{\partial D} \int_{\underline{Y}}^{y^R} u''(c_{2,default}) f_y(y) dy + \frac{\partial y^R}{\partial D} u'(c_{2,default}(S, D, y^R)) f_y(y^R) - \frac{\partial y^R}{\partial D} u'(y^R + R_D S - R_T P)) f_y(y^R) \bigg). \quad (41)$$

This expression is continuous. We also have

$$\frac{\partial^2 U_{overall}}{\partial S \partial \theta} = R_D \bigg(\frac{\partial y^R}{\partial \theta} u'(c_{2,default}(S, D, y^R)) f_y(y^R) - \frac{\partial y^R}{\partial \theta} u'(y^R + R_D S - R_T P)) f_y(y^R) \bigg), \quad (42)$$

which is also continuous.

It is also immediate from equation 30 that $\frac{\partial U_{overall}}{\partial S}$ is continuously differentiable with respect to w. Using all of these facts, and the fact that

$$\frac{\partial^2}{\partial S^2} U_{overall} < 0 \quad (43)$$

for all S, we can apply the implicit function theorem to derive that S_{max} is continuously differentiable with respect to D, w, and θ.

If $S_{max} > D$, $S^* = S_{max}$, and so we have that S^* is continuously differentiable with respect to D, w, and θ. If $S_{max} < D$, $S^* = D$. Since marginal changes in D, w, and θ still leave $S_{max} < D$, S^* has constant derivative 0 with respect to w and θ and one with respect to D whenever $S_{max} < D$. S^* may fail to be continuously differentiable when $S_{max} = D$. However, note that $\frac{\partial S_{max}}{\partial w} > 0$ where it exists. This follows from the fact that $U_{overall}$ is concave in S and (as can be seen in equation 28), the marginal utility of S is increasing in w. Furtheremore, at the points where S_{max} is not differentiable with respect to w (in particular, the w values for which y^R is equal to $\underline{Y}$ or $\overline{Y}$), it is both left and right-differentiable, with negative semi-derivatives. Thus, given θ, $S_{max} = D$ holds for at most one value of w, and thus for a zero measure of borrowers.

Similarly, $\frac{\partial y^R}{\partial \theta}$ is negative where it exists. At both $\underline{Y}$ and $\overline{Y}$, y^R is both left and right differentiable with respect to θ with negative semi-derivatives. Since changes in w don't affect y^R directly, this implies that in the case of constrained savings ($S_{max} < D$,) $y^R = \underline{Y}$ or $y^R = \overline{Y}$ for *any w* for only a zero measure (two-element) set of θ. Furthermore, in the unconstrained case, changes in w affect y^R only through changes in S_{max}. Since S_{max} is increasing in w everywhere, $\frac{\partial y^R(S^*)}{\partial \theta}$ is negative where it exists. Similarly at both $\underline{Y}$ and $\overline{Y}$, y^R is both left and right differentiable with respect to w with negative semi-derivatives. Thus in the unconstrained case, y^R is equal to one of its endpoints for only a zero-measure set of w given any θ. Thus, given any D, there is are at most two values of θ for which y^R is equal to one of its endpoints for more than a zero-measure set of w. Thus the claim is proven. □

Claim 3:*Let $U^*(D, w, \theta)$ denote total utility from borrowing with optimized savings. U^* is continuously differentiable in all of its arguments whenever $S_{max} \neq D$, $y^R \neq \underline{Y}$, and $y^R \neq \overline{Y}$.*

Proof. Note that

$$U^*(D,w,\theta) = U_{overall}(D, S^*(D,w,\theta), w, \theta). \tag{44}$$

Thus differentiability is immediate from claims one and two, and

$$\frac{\partial}{\partial w}U^*(D,w,\theta) = \frac{\partial U_{overall}}{\partial S^*}\frac{\partial S^*}{\partial w} + \frac{\partial U_{overall}}{\partial w}. \tag{45}$$

And analogous expressions hold for the derivatives with respect to θ and D. Recall that we either have $S^* = S_{max}$ or $S^* = D$. If $S^* = S_{max}$, then $\frac{\partial U_{overall}}{\partial S^*} = 0$, and

$$\frac{\partial}{\partial x}U^*(D,w,\theta) = \frac{\partial U_{overall}}{\partial x} \tag{46}$$

for each variable $x \in \{D, \theta, w\}$. Thus continuous differentiability follows from claim 1. If $S^* = D$, $\frac{\partial S^*}{\partial w} = \frac{\partial S^*}{\partial \theta} = 0$, and thus we can again ignore the S^* in the relevant derivative, and so continuous differentiability with respect to w and θ again follows immediately from claim 1. If $S^* = D$, $\frac{\partial S^*}{\partial D} = 1$, so

$$\frac{\partial}{\partial D}U^*(D,w,\theta) = \frac{\partial U_{overall}}{\partial S^*} + \frac{\partial U_{overall}}{\partial D}, \tag{47}$$

and continuous differentiability follows from claims 1 and 2. □

Claim 4 (Proposition 4): *Potential borrowers with $\theta_i > \theta^*(D,w)$ who are definitely credit constrained will have $S = D$, and they would be strictly better off with a lower required deposit. Moreover, if repossessions are negative equity, potential borrowers with a nonzero chance of default are also better off with a lower deposit irrespective of whether they are credit constrained. In the case of positive equity or zero probability of default, borrowers who are not credit constrained are indifferent to marginal changes in D. Trivially, those with $\theta_i < \theta^*(D)$ are also indifferent to marginal changes in D since they do not borrow.*

Proof. Recall from the proof of claim 3 that for non-credit-constrained borrowers (those who set $S^* > D$,)

$$\frac{\partial U^*}{\partial D} = \frac{\partial U_{total}}{\partial D}. \tag{48}$$

It is thus immediate from equation 32 that U^* is unchanging in D in the positive equity case and decreasing in D in the negative equity case. For credit-constrained borrowers (those who set $S^* = D$), we have

$$\frac{\partial U^*}{\partial D} = \frac{\partial U_{total}}{\partial D} + \frac{\partial U_{overall}}{\partial S^*}. \tag{49}$$

The first term in this expression is zero in the positive equity case and negative in the negative equity case. To sign the second term, recall that borrowers are credit-constrained if and only if

$$S_{max} < D, \tag{50}$$

where S_{max} is the unique point at which $\frac{\partial U_{total}}{\partial S} = 0$. But since U_{total} is concave in S, this means that $S^* = D > S_{max}$ implies $\frac{\partial U_{overall}}{\partial S^*} < 0$. Thus the expression is strictly negative in both the positive and negative equity cases. □

Proof of Proposition 3

Proof. We have that

$$\frac{\partial U^*}{\partial \theta} = 1 - F(y^R) \tag{51}$$

for all levels of θ. Since borrowers are strictly worse off borrowing if they have a repayment probability of zero, $\theta = \theta^*$ implies that $F(y^R) < 1$. This fact, along with claim 3, allows us to apply the implicit function theorem, giving that θ^* is continuously differentiable in D and w whenever $S_{max} \neq D$, $y^R \neq \underline{Y}$ and $y^R \neq \overline{Y}$. **It is at this point that we invoke assumption A, which states that $S_{max} = D$ or $y^R = \underline{Y}$ at θ^* for at most a zero-measure set of w.** (Note that we can never have $y^R = \overline{Y}$ at θ^*, since borrowers who will always default are strictly worse off borrowing). Thus continuous differentiability in D and w holds for all but a zero-measure set of w. Since U^* is increasing in w faster than $\overline{U}$ is, θ^* is decreasing in w. [42] For those farmers for whom U^* is strictly decreasing in D, θ^* is increasing in D. For those farmers for whom U^* is unchanging in D, θ^* is unchanging in D.

For a fixed w, the repossession rate is decreasing in the deposit requirement D, because θ^* is increasing in D (adverse selection) and y^R is decreasing in D (moral hazard).

□

□

Assumption A:

$S_{max} = D$ or $y^R = \underline{Y}$ at θ^* for at most a zero-measure set of w, and at w^* for at most a zero-measure set of θ.

Although S_{max} is increasing in w, it may be increasing in θ. But θ^* is decreasing in w. It is thus possible, in principle, that $S_{max} = D$ could hold at θ^* for a nonzero-measure set of w. In such a case, the profit function could fail to be differentiable. However, this condition would require peculiar behavior: by the existence of credit-constrained borrowers, $S_{max} < D$, at $(\underline{W}, \theta^*(\underline{W}))$. Thus in order for S_{max} to be equal to D for a positive-measure set of w, one of two things would need to happen. In one case $S_{max}(\theta^*)$ would need to be increasing or decreasing in w until it hits D, at which point its derivative with respect to w would need to be exactly zero for an interval of w's. In the other case, S_{max} would need to bounce above and below D so pathologically as w increases as to be equal to D at an uncountable number of points. (Analogous behavior could yield that $S_{max} = D$ at w^* for a nonzero-measure set of θ, where w^* is as defined below.) We have no reason to think this bizarre behavior is especially probable, and thus reasonable priors are that the parameters are almost always such that assumption A holds. Exactly analogous logic holds for the $y^R = \underline{Y}$ case.

Derivative of Expected Profit

[42] That U^* is increasing in w faster than $\overline{U}$ is follows from the fact that borrowers always have lower second-period consumption than non-borrowers, and thus higher savings. The result is thus immediate from the envelope theorem.

Proof. To show that expected profit is continuously differentiable in D whenever $D \neq D_F$, it is convenient to change the order of integration to

$$E(\Pi(D)) = \left\{ \int_{\underline{\theta}}^{\bar{\theta}} \int_{w^*(D,\theta)}^{\overline{W}} \left[\Pi_r - F(y^R(\theta, S^*(w,D), D))L_d(D)\right] f_w(w) f_\theta(\theta) d\theta dw \right\}. \quad (52)$$

Note that the existence of a w^* for every θ follows from two facts. First $lim_{w \to \infty} U^* - \overline{U} = \theta$, since as w grows, repayment probability approaches one and the consumption differential between borrowers and non-borrowers approaches an infinitesimal share of consumption. Secondly, $lim_{w \to D} U^* - \overline{U} = -\infty$, since consumption is always lower in the case of borrowing.

Because optimal savings is always changing in w, but not always changing in θ, it simplifies the proof to change the order of integration and consider w^* rather than θ^*. However, we will show at the end of the proof that the resulting expression for the derivative of expected profits is equal to the one used in the body of the paper.

Consider the functions $Z : \mathbb{R}^3 \to \mathbb{R}$ and $H : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$Z(w_0, \theta, D) = \int_{w_0}^{\bar{W}} \left[\Pi_r - F(y^R(\theta, S^*(w,D), D))L_d(D)\right] f_w(w) dw \quad (53)$$

and

$$H(\theta, D) = (w^*(\theta, D), \theta, D). \quad (54)$$

Note that

$$E(\Pi(D)) = \int_{\underline{\theta}}^{\bar{\theta}} Z(H(D)) f_\theta(\theta) d\theta. \quad (55)$$

We proceed by demonstrating the continuous differentiability of various terms in Z and H using the implicit function theorem. Assume for the below (through equation 64) that y^R is not equal to either of the endpoints of its support. Consider first the case of credit-constrained borrowers, who have $S_{max} < D$ and thus set $S^* = D$. Define $F_1 : \mathbb{R}^4 \to \mathbb{R}^1$, which we will use to define y^R given a fixed w, θ and D. Set

$$F_1(y, w, \theta, D) = \theta_i + M + u(y + R_D D - R_T P) - u(c_{2,default}). \quad (56)$$

The total differential dF_1 is represented by

$$\left[u_r' - u_d' \quad 0 \quad 1 \quad R_D(u_r' - u_d') - \tfrac{\partial c_{2,default}}{\partial D} u_d' \right], \quad (57)$$

where u_r' denotes the marginal utility of consumption under repayment, $u'(y^R + R_D D - R_T P)$, and u_d' the marginal utility of consumption under default, $u'(c_{2,default})$. It can be verified that each entry in dF_1 is continuous in (y, w, θ, D)-space, and thus F_1 is continuously differentiable over $\mathbb{R}^4$. Furthermore, $u_r' - u_d' > 0$. Thus by the implicit function theorem, y^R is continuously differentiable with respect to (w, θ, D), and thus also with respect to each individual term in this vector.

In order to show continuous differentiability of w^*, we define a new function $G_1 : \mathbb{R}^4 \to \mathbb{R}^2$ which can be used to jointly determine y^R and w^* for a fixed θ and D. We define

$$G_1(y, w, \theta, D) = \begin{bmatrix} \theta_i + M + u(y + R_D D - R_T P) - u(c_{2,default}) \\ U(y, w, \theta, D) - \overline{U}(w) \end{bmatrix}. \tag{58}$$

The total differential dG_1 is given by

$$\begin{bmatrix} u'_r - u'_d & 0 & 1 & R_D(u'_r - u'_d) - \frac{\partial c_{2,default}}{\partial D} u'_d \\ \frac{\partial U}{\partial y} & \left(\frac{\partial U}{\partial w} - \frac{\partial \overline{U}}{\partial w}\right) & \frac{\partial U}{\partial \theta} & \frac{\partial U}{\partial D} \end{bmatrix}. \tag{59}$$

This is equal to

$$\begin{bmatrix} u'_r - u'_d & 0 & 1 & R_D(u'_r - u'_d) - \frac{\partial c_{2,default}}{\partial D} u'_d \\ 0 & u'_b - u'_n & 1 - F(y) & \frac{\partial U}{\partial D} \end{bmatrix} \tag{60}$$

where u'_b denotes the marginal utility of first-period wealth for borrowers, which is in this case given by $u'(w - D)$, and u'_n denotes the marginal utility of first-period wealth for non-borrowers, given by $u'(w - S_n)$, where S_n satisfies the non-borrower's euler equation. It can again be shown that each entry in dG_1 is continuous as a function of (y, w, θ, D) and and thus dG_1 is continuous. Furthermore

$$det\left(\begin{bmatrix} u'_r - u'_d & 0 \\ 0 & u'_b - u'_n) \end{bmatrix}\right) = (u'_r - u'_d)(u'_b - u'_n). \tag{61}$$

Since nonborrowers save less than borrowers with the same initial wealth level, this expression is always positive, and thus the matrix is invertible. Thus we can apply the implicit function theorem to derive that y^R and w^*, when defined jointly, are continuously differentiable with respect to (θ, D).

We can demonstrate the same results in the non-constrained case, in which $S^* = S_{max} > D$, through an analogous process. In this case, we define $F_2 : \mathbb{R}^5 \to \mathbb{R}^2$ and $G_2 : \mathbb{R}^5 \to \mathbb{R}^3$ by

$$F_2(S, y, w, \theta, D) = \begin{bmatrix} \frac{\partial}{\partial S} U \\ \theta_i + M + u(y + R_D D - R_T P) - u(c_{2,default}) \end{bmatrix}, \tag{62}$$

and

$$G_2(S, y, w, \theta, D) = \begin{bmatrix} \frac{\partial}{\partial S} U \\ \theta_i + M + u(y + R_D D - R_T P) - u(c_{2,default}) \\ U(y, w, \theta, D) - \overline{U}(w) \end{bmatrix}. \tag{63}$$

It can again be verified that dF_2 and dG_2 are continuous in $\mathbb{R}^5$. Furthermore, the relevant determinant for dF_2 is equal to

$$\frac{\partial^2 U}{\partial S^2}(u'_r - u'_d) - R_D \frac{\partial^2 U}{\partial S \partial y}.$$

We showed in the proof of claim two that this expression is always negative. [43] The relevant determinant for dG_2 is equal to

$$\left[\frac{\partial^2 U}{\partial S^2}(u_r' - u_d') - R_D \frac{\partial^2 U}{\partial S \partial y}\right](u_b' - u_d'). \tag{64}$$

This expression is also negative.

Thus in all cases such that $D \neq D_F$, $S_{max} \neq D$, $y^R \neq \underline{Y}$, and $y^R \neq \overline{Y}$, S^*, y^R, and w^* are continuously differentiable with respect to (S^*, y^R, w, θ, D). With this established, we can move to the continuous differentiability of the component functions of profit.

We return now to consideration of the functions, Z and H, that we defined above. Much of the remainder of the proof is built around an extension of Leibniz' integral rule that states that if a function $f(w, t)$ is measurable and integrable over w, and is differentiable in t for all but a zero-measure set of w's in the interval A, with derivative bounded on A in absolute value by an integrable function, then $\int_A f(w, t)$ is differentiable with derivative $\int_A f'(w, t)$. (Billingsley 1995)

We claim, given this result, that Z is continuously differentiable in D and θ for all but two possible θ values. These are the values at which $y^R = \overline{Y}$ and $y^R = \underline{Y}$ for more than a zero-measure set of w. Call them θ_U and θ_L, respectively. To see that Z is continuously differentiable for all other θ, recall that we showed above that $\left[\Pi_r - F(y^R(\theta, S^*(w, D), D))L_d(D)\right]$ is continuously differentiable with respect to (w, θ, D) whenever $S_{max} \neq D$, $y^R = \overline{Y}$ and $y^R = \underline{Y}$. Recall from claim two of the proof of proposition three that for a given θ, one of these conditions holds for at most three w (call them ω_1, ω_2, and ω_3.). By the leibniz' rule extension, we thus have differentiability of Z as long as the derivatives of

$$\left[\Pi_r - F(y^R(\theta, S^*(w, D), D))L_d(D)\right]$$

with respect to D and θ are bounded in absolute value by an integrable function over $[\underline{W}, \overline{W}] \setminus \{\omega_i | i \in \{1, 2, 3\}\}$. Note that the derivative with respect to D is

$$\left(-\frac{\partial y^R}{\partial D} f(y^R) L_d(D) - F(y^R) L_d'(D)\right).$$

Every term in this expression except for $\frac{\partial y^R}{\partial D}$ is trivially bounded. But note that $\frac{\partial y^R}{\partial D}$ can take one of two values: the value for the unconstrained case in which the borrower saves S_{max} or the value for the constrained case in which the borrower saves D. We have already shown that both of these expressions are continuous in w, and thus are bounded in absolute value on $[\underline{W}, \overline{W}]$. Thus $\frac{\partial y^R}{\partial D}$, and so the whole expression of interest, is bounded in absolute value by a constant (and therefore integrable) function.

[43] In that case we labeled this whole expression as $\frac{\partial^2 U_{overall}}{\partial S^2}$, because we were only interested in S^*, and so took y^R as a function of S^* rather than determining their derivatives jointly.

Thus Z is continuously differentiable in D whenever $\theta \neq \theta_L$ and $\theta \neq \theta_U$, and in particular,

$$\frac{\partial}{\partial D} Z = \int_{w_0}^{\overline{W}} \left(-\frac{\partial y^R}{\partial D} f(y^R) L_d(D) - F(y^R) L_d'(D) \right) f_w(w) \tag{65}$$

Note also that the differentiability of Z in w is immediate by the continuity of y^R in w, and we have

$$\frac{\partial}{\partial w_0} Z(w_0, D) = -\left[\Pi_r - F(y^R(\theta, S^*(w_0, D), D)) L_d(D) \right] f_w(w_0), \tag{66}$$

which is continuous with respect to (w_0, θ, D).[44]

From our results above, we also have that H is continuously differentiable whenever θ and D are such that $S_{max} \neq D$ at w^* and y^R is not equal to one of the endpoints of its support. Recall that assumption A ensures that w^* is not so pathological that for some D, $S_{max}(w^*) = D, y^R = \underline{Y}$ or $y^R = \overline{Y}$ for a nonzero mass of θ. By a similar argument to that which we used to show the boundedness of $\frac{\partial y^R}{\partial D}$, we have that $\frac{\partial w^*}{\partial D}$ is bounded in absolute value over the set of all $\theta \in [\underline{\theta}, \overline{\theta}]$ such that $S_{max}(w^*) \neq D$, $y^R \neq \underline{Y}$, and $y^R \neq \overline{Y}$.

Putting these together, we derive that $Z \circ H$ is continuously differentiable in $\mathbb{R}^2$ for all but a zero-measure set of θ with derivative

$$\begin{aligned} &-\frac{\partial w^*}{\partial D} \left[\Pi_r - F(y^R(\theta, S^*(w^*, D), D)) L_d(D) \right] f_w(w^*) \\ &\qquad + \int_{w^*}^{\overline{W}} \left(-\frac{\partial y^R}{\partial D} f(y^R) L_d(D) - F(y^R) L_d'(D) \right) f_w(w). \end{aligned} \tag{67}$$

Given this, since $E(\Pi(D)) = \int_{\underline{\theta}}^{\overline{\theta}} Z(H(D)) f_\theta(\theta) d\theta$, we can again invoke the Leibniz' rule extension to derive that $E(\Pi(D))$ is continuously differentiable in D with derivative

$$\begin{aligned} \int_{\underline{\theta}}^{\overline{\theta}} \Bigg[&-\frac{\partial w^*}{\partial D} \left[\Pi_r - F(y^R(\theta, S^*(w^*, D), D)) L_d(D) \right] f_w(w^*) \\ &+ \int_{w^*}^{\overline{W}} \left(-\frac{\partial y^R}{\partial D} f(y^R) L_d(D) - F(y^R) L_d'(D) \right) f_w(w) dw \Bigg] f_\theta(\theta) d\theta. \end{aligned} \tag{68}$$

That the second line of this expression (integrated over θ) is equal to the analogous expressions in the body of the paper is immediate from a change in the order of integration. To see that the first line is equal to the analogous expression in the body of the paper, consider the function

[44] Technically, Z could fail to be differentiable when w^* is equal to one of the endpoints of its support. However, w^* is strictly decreasing in θ, and so this can occur for only a zero-measure set of θ. Thus as with other zero-measure discontinuity points (we won't repeat another argument along these lines given the frequency with which they appear in this proof), we can work around this.

$\Phi : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\Phi(D, D_0) = \int_{\underline{\theta}}^{\overline{\theta}} \int_{w^*(D,\theta)}^{\bar{W}} \left[\Pi_r(D_0) - F(y^R(\theta, S^*(w, D_0), D_0))L_d(D_0)\right] f_w(w) f_\theta(\theta) d\theta dw. \quad (69)$$

That is, for a given deposit requirement D_0, Φ is a function which encompasses just the external margin effects of D: changes in D change the limits of the integral, but not the integrand. We can change the order of integration to yield

$$\Phi(D, D_0) = \int_{\underline{w}}^{\overline{w}} \int_{\theta^*(D,w)}^{\bar{\theta}} \left[\Pi_r(D_0) - F(y^R(\theta, S^*(w, D_0), D_0))L_d(D_0)\right] f_w(w) f_\theta(\theta) d\theta dw. \quad (70)$$

Assumption A assures that Φ is differentiable at $D = D_0$, and taking derivatives of both of the expressions for Φ above yields the desired result.

□

Lemma 1. *The profit-maximizing deposit ratio will be such that there is some non-zero probability of repossession.*

Proof. Assume for contradiction that D^* is such that the overall probability of repossession is zero. Let $\mathbb{P}(D, w)$ denote the probability of an individual with initial wealth level w borrowing and defaulting when the deposit requirement is D. Let Ω_0 denote the set of all w such that repossession occurs with nonzero probability for $D = D^*$. Recalling that we have assumed the probability of repossession is zero when the deposit level is D^*, we have

$$0 = \int_{\underline{w}}^{\overline{w}} \mathbb{P}(D^*, w) dw \quad (71)$$

$$= \int_{\Omega_0} \mathbb{P}(D^*, w) dF_w \quad (72)$$

By definition, for any $w \in \Omega_0$,

$$\mathbb{P}(D^*, w) > 0.$$

Thus

$$\int_{\Omega_0} \mathbb{P}(D^*, w) dF_w = 0$$

$$\implies \mu(\Omega_0) = 0$$

$$\implies \mu(\Omega_0^c) = 1.$$

Note that Ω_0^c, the complement of Ω_0, is the set of all w such that $\mathbb{P}(D^*, w) = 0$

Recall that the derivative of expected profit with respect to the deposit ratio (for $D \neq D_F$)

is

$$\frac{\partial E(\Pi(D))}{\partial D} = \int_{\underline{w}}^{\overline{w}} \Bigg[-\frac{\partial \theta^*}{\partial D} f_\theta(\theta^*) f_w(w) \left(\Pi_r - F(y^R(\theta, S^*(w,D), D)) L_d(D^*)\right)$$
$$- \left(\int_{\theta^*}^{\bar\theta} \frac{\partial F(y^R(\theta, S^*, D))}{\partial D} f_\theta(\theta) f_w(w) d\theta \right) L_d(D^*)$$
$$- \left(\int_{\theta^*}^{\bar\theta} F(y^R(\theta, S^*, D)) f_\theta f_w(w)(\theta) d\theta \right) L'_d(D^*) \Bigg] dw \quad (73)$$

By the fact that Ω_0 has measure zero, this is equal to

$$\int_{\Omega_0^c} \Bigg[-\frac{\partial \theta^*}{\partial D} f_\theta(\theta^*) \left(\Pi_r - F(y^R(\theta, S^*(w,D), D)) L_d(D^*)\right)$$
$$- \left(\int_{\theta^*}^{\bar\theta} \frac{\partial F(y^R(\theta, S^*, D))}{\partial D} f_\theta(\theta) d\theta \right) L_d(D^*)$$
$$- \left(\int_{\theta^*}^{\bar\theta} F(y^R(\theta, S^*, D)) f_\theta(\theta) d\theta \right) L'_d(D^*) \Bigg] dF_w \quad (74)$$

When $\mathbb{P}(D^*, w) = 0$, by definition $F(y^R(\theta, S^*, D) = 0$ for all $\theta > \theta^*(D^*)$. Since y^R is weakly decreasing in D, this implies that $\frac{\partial F(y^R(\theta, S^*, D))}{\partial D} = 0$.[45] Thus

$$\int_{\Omega_0^c} - \left(\int_{\theta^*}^{\bar\theta} \frac{\partial F(y^R(\theta, S^*, D))}{\partial D} f_\theta(\theta) d\theta \right) L_d(D^*) dF_w \quad (75)$$
$$= \int_{\Omega_0^c} - \left(\int_{\theta^*}^{\bar\theta} F(y^R(\theta, S^*, D)) f_\theta(\theta) d\theta \right) L'_d(D^*) dF_w \quad (76)$$
$$= 0. \quad (77)$$

So

$$\frac{\partial E(D)}{\partial D} = \int_{\Omega_0^c} -\frac{\partial \theta^*}{\partial D} f_\theta(\theta^*) \left(\Pi_r - F(y^R(\theta, S^*(w,D), D)) L_d(D^*)\right) dF_w \quad (78)$$
$$= \int_{\Omega_0^c} -\frac{\partial \theta^*}{\partial D} f_\theta(\theta^*) \Pi_r dF_w \quad (79)$$

By assumption, there exists a range of w for which $\theta^* \in [\underline{\theta}, \overline{\theta}]$, and for w in this range, $\frac{\partial \theta^*}{\partial D} > 0$. Since Ω_0^c has measure one, its intersection with this range has nonzero measure, and thus

$$\frac{\partial E(D^*)}{\partial D} = \int_{\Omega_0^c} -\frac{\partial \theta^*}{\partial D} f_\theta(\theta^*) \Pi_r dF_w < 0,$$

[45] Over the measure one set on which it exists.

and profit is not maximized. □

Table 1: Program design

Treatment (loan) description	Deposit amount (KSh)	Guarantor amount (KSh)	Collateralization with tank (KSh)	Offers
4% deposit (A)	1,000	0	23,000	510
100% cash collateralized loan (C)	8,000	16,000	0	419
25% deposit loan, maintained (D^M)	6,000	0	18,000	225
25% deposit loan, waived (D^W)	6,000 $\rightarrow$ 1,000	0	18,000	225
21% guarantor loan, 4% deposit, maintained (G^M)	1,000	5,000	18,000	225
21% guarantor loan, 4% deposit, waived (G^W)	1,000	5,000 $\rightarrow$ 0	18,000	200

Note: Loan amount is KSh 24,000 for all treatment groups.
All amounts in KSh (roughly KSh 75=$1)

Table 2: Baseline randomization checks

	Mean	F-test stat	P-value
Milk production (Aug 2009 - Jan 2010)			
(1) Average monthly milk production	207.4	1.229	0.297
(2) Monthly milk per cow	133.2	0.523	0.719
(3) Monthly cows calved down	0.103	2.691**	0.030
Milk sales (Aug 2009 - Jan 2010)			
(4) Monthly sales to dairy	69.01	1.175	0.320
(5) Sold milk to dairy dummy	0.480	2.129*	0.075
Livestock (Aug 2009 - Jan 2010)			
(6) At least one cow died	0.318	0.539	0.707
(7) At least one cow got sick	0.516	2.091*	0.080
(8) Zerograzing dummy	0.177	0.265	0.901
(9) Zero or semi-zerograzing dummy	0.749	1.899	0.108
Assets			
(10) Household assets (ln KSh)	12.27	0.976	0.420
(11) Value of livestock (ln Ksh)	11.29	1.038	0.386
(12) Monthly cows producing milk	1.660	1.858	0.115
(13) Baseline piped water	0.315	0.726	0.574
(14) Own water tank	0.428	0.256	0.906
(15) Own water tank > 2500 liters	0.241	0.444	0.777
Schooling			
(16) Kids (5-16) enrolled in school	0.975	0.302	0.877
(17) Girls (5-16) enrolled in school	0.980	0.554	0.696
(18) Boys (5-16) enrolled in school	0.970	0.261	0.903
Household characteristics			
(19) Household head education (years)	8.459	1.193	0.312
(20) Female household head	0.201	0.603	0.660
Time use (minutes per day)			
(21) Farming	87.0	1.298	0.269
(22) Livestock	77.2	0.665	0.616
(23) Fetching water	14.3	1.556	0.184
(24) Working	38.8	0.172	0.953
(25) School (Girls 5-16)	330.5	0.647	0.629
(26) School (Boys 5-16)	336.3	1.033	0.390

Note: Milk volumes in liters per month. Reported means are across all six loan groups. The F-stat tests for equality of means across all six loan groups. Certain time use variables are omitted due to space constraints. One excluded time use variable (socializing with neighbors) has a significant F-test statistic. Including the ten omitted time use variables, we conduct baseline checks on 39 variables. Standard errors are clustered at the household level when necessary.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Borrower characteristics across arms

	(1) Full sample incl. non-borrowers	(2) 100% collateralized borrowers	(3) 25% deposit borrowers	(4) 4% deposit 21% guarantor borrowers	(5) 4% deposit borrowers
(1) Log household assets	12.28	12.30	12.60	12.68	12.44
	[0.02]	[0.25]	[0.10]	[0.10]	[0.06]
(2) Log per capita expenditure	10.37	10.36	10.56	10.64	10.41
	[0.02]	[0.10]	[0.07]	[0.07]	[0.04]
(3) Avg cows producing milk	1.67	1.80	1.94	2.04	1.93
	[0.03]	[0.18]	[0.17]	[0.17]	[0.08]
(4) Milk per cow (liters)	142.7	142.7	163.9	143.6	148.4
	[2.27]	[23.57]	[10.34]	[10.34]	[5.91]
(5) Monthly sales to dairy (liters)	78.2	86.3	106.1	89.3	115.1
	[4.14]	[32.96]	[13.44]	[13.44]	[22.99]
(6) Education (years) of HH head	8.46	10.30	9.78	9.08	9.14
	[0.11]	[1.54]	[0.36]	[0.36]	[0.30]
(7) Female HH head	0.20	0.20	0.18	0.24	0.15
	[0.01]	[0.13]	[0.03]	[0.03]	[0.02]
(8) Girls as % of HH	0.13	0.05	0.13	0.11	0.10
	[0.00]	[0.04]	[0.01]	[0.01]	[0.01]
(9) Piped water access	0.32	0.40	0.27	0.30	0.34
	[0.01]	[0.16]	[0.04]	[0.04]	[0.03]
(10) Own tank	0.43	0.80	0.49	0.46	0.49
	[0.01]	[0.13]	[0.05]	[0.05]	[0.03]
(11) Own big tank (> 2500 L)	0.24	0.40	0.30	0.33	0.24
	[0.01]	[0.16]	[0.04]	[0.04]	[0.03]
(12) Number of big tanks	0.32	0.40	0.41	0.43	0.30
	[0.02]	[0.16]	[0.07]	[0.07]	[0.04]
(13) Practice zero grazing	0.18	0.20	0.18	0.19	0.23
	[0.01]	[0.13]	[0.03]	[0.03]	[0.03]
(14) Practice zero/semi zerograzing	0.75	1.00	0.81	0.77	0.80
	[0.01]	[0.00]	[0.04]	[0.04]	[0.03]

Note: Standard errors in brackets.
All data is pre-treatment. Log per capita expenditure is measured in log Kenya shillings per year.
There are significant differences between borrowers and non-borrowers at the 5% level in the first three rows, columns (3)-(5); row 5, columns (4) and (5); row 6, column (5); row 10, column (2); row 11, column (4); and row 14, column (3).

Table 4: Loan take-up rates and standard errors

	Original sample		Out of sample loans		Combined data		
	Loans taken up/offers	Rate (percent)	Loans taken up/offers	Rate (percent)	Total loans taken up/offers	Overall Rate (percent)	P-value of difference (percent)
100% cash collateralized loan (C)	10/419	2.39			10/419	2.39	
		[0.75]				[0.75]	
25% deposit loan (D)	124/450	27.55	233/1042	22.36	357/1492	23.93	0.031
		[2.11]		[1.29]		[1.10]	
21% guarantor, 4% deposit loan (G)	100/425	23.53	261/1036	25.19	361/1461	24.71	0.50
		[2.06]		[1.35]		[1.13]	
4% deposit (A)	226/510	44.31	205/519	39.50	431/1029	41.89	0.12
		[2.20]		[2.15]		[1.54]	

Note: The original sample loans were offered during March 2010, May 2010, and June 2010. The out of sample loans were offered Feb to April 2012. Standard errors shown in brackets. Standard errors calculated as $SE = \sqrt[2]{p(1-p)/n}$, where p is the percentage of loan take-up and n is the number of offers.

Table 5: Tank repossession and loan non-recovery rates: combined sample

Group	Tank repossession		Loan non-recovery	
	Count	Rate (percent)	Count	Rate (percent)
4% deposit (A)	3/431	0.7	0/431	0
		(0.14, 2.02)		(0, 0.85)
25% deposit (D)	0/357	0	0/357	0
		(0, 0.83)		(0, 0.83)
21% guarantor, 4% deposit (G)	0/361	0	0/361	0
		(0, 0.83)		(0, 0.83)
100% cash collateralized (C)	0/10	0	0/10	0
		(0, 25.89)		(0, 25.89)
Treatment effect on repossession p value	0.0525			

4% deposit = 25% deposit or guarantor

Note: Tank repossession and loan non-recovery data include loans from the original sample and out of sample groups. Of the three tank repossessions in the 4% group, one repossession was in the original sample while two were in the out-of-sample group. 25% deposit or guarantor refers to the aggregate of the 25% deposit and 21% guarantor, 4% deposit groups. 95% Clopper-Pearson exact confidence intervals are displayed in parentheses under the point estimates for each of the rates. One-sided tests were conducted for cases with zero repossessions. Treatment effect on repossession is obtained by conducting Fishers Exact Test for the difference between rates of 4% deposit and 25% deposit or guarantor groups. Note that including the additional 152 loans the Nyala cooperative has offered independently, the p-value is 0.0362.

Table 6: Late repayment

	During loan cycle			Late at end loan		
	(1)	(2)	(3)	(4)	(5)	(6)
	Late ever	Rec'd pending default letter	Security deposit reclaimed	Repaid late	Late balance (KSh)	Months late
4% deposit loan	0.57***	0.29***	0.09***	0.12***	221.79***	0.13***
	[0.11]	[0.03]	[0.02]	[0.02]	[50.02]	[0.03]
25% deposit loan, maintained	0.59***	0.33***	0.16***	0.02	45.67	0.02
	[0.12]	[0.06]	[0.05]	[0.02]	[33.04]	[0.02]
25% deposit loan, waived	0.46***	0.28***	0.08**	0.12***	161.90**	0.13***
	[0.12]	[0.06]	[0.04]	[0.04]	[66.76]	[0.05]
21% guarantor loan, 4% deposit, maintained	0.51***	0.18***	0.10**	0.06*	101.91	0.08*
	[0.13]	[0.05]	[0.04]	[0.03]	[63.43]	[0.05]
21% guarantor loan, 4% deposit, waived	0.43***	0.32***	0.14***	0.14***	297.52***	0.22**
	[0.13]	[0.07]	[0.05]	[0.05]	[111.67]	[0.09]
Constant(100% secured joint-liability loan)	0.11	0.00	0.00	0.00	0.00	0.00
	[0.11]	[0.00]	[.]	[.]	[0.00]	[.]
Deposit Selection Effect P-value	0.10	0.97	0.80	0.99	0.47	0.99
25% dep loan waived = 4% dep loan						
Guarantor Selection Effect P-value	0.07	0.64	0.38	0.66	0.54	0.34
25% guar loan waived = 4% dep loan						
Deposit Treatment Effect P-value	0.13	0.55	0.2	0.02	0.12	0.03
25% dep loan maintained = 25% dep loan waived						
Guarantor Treatment Effect P-value	0.42	0.10	0.54	0.18	0.13	0.16
25% guar loan maintained = 25% guar loan waived						
Mean of dependent variable	0.64	0.28	0.11	0.10	180.36	0.12
Observations	456	456	456	456	456	456

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Heteroskedasticity-robust standard errors in brackets.

Table 7: Early repayment

	(1) Repaid early	(2) Months early	(3) Months of principal in deposit	(4) Foregone months of low interest loan	(5) Months of repayment freed by waiver
100% cash collateralized loan (C)	0.900	15.000***	8	7.000***	–
	[0.100]	[2.431]		[2.431]	
25% deposit loan, maintained (D^M)	0.594	5.500***	6	−0.500	–
	[0.062]	[0.835]		[0.835]	
25% deposit loan, waived (D^W)	0.383	4.957***	1	3.957***	5
	[0.063]	[1.113]		[1.113]	
4% deposit, 21% guarantor loan, maintained (G^M)	0.560	3.804***	1	2.804***	–
	[0.071]	[0.810]		[0.810]	
4% deposit, 21% guarantor loan, waived (G^W)	0.320	5.214***	1	4.214***	–
	[0.067]	[1.281]		[1.281]	
4% deposit loan (A)	0.239	1.875***	1	0.875***	–
	[0.028]	[0.322]		[0.322]	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Real impacts on water access, cow health, and milk production: 4% deposit arm versus 100% cash collateralized arm

	(1)	(2)	(3)	(4)	(5)	(6)
	Own tank	Log total capacity	Own large tank	Any cow was sick	Production	Log production
Treat*Post	0.175***	0.609***	0.265***	−0.133***	0.831	0.047
	[0.023]	[0.083]	[0.030]	[0.036]	[12.979]	[0.048]
Treatment	−0.051	−0.174	−0.046*	0.102***	12.473	−0.033
	[0.033]	[0.109]	[0.028]	[0.033]	[12.566]	[0.052]
Constant	0.445***	6.932***	0.253***	0.449***	221.331***	5.207***
	[0.027]	[0.095]	[0.024]	[0.025]	[8.419]	[0.037]
Dep Var Mean	0.518	7.114	0.334	0.409	311.554	5.532
Round FE	Yes	Yes	Yes	Yes		
HH Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2649	1830	1830	5099	5151	4960

Note: All household survey data is collapsed by survey round (Nov 2011, Feb 2012, May 2012, and Sept 2012). All endline household survey data was collected only in the 100% cash collateralized and the 4% deposit treatment groups.
In column (3), owning a large tank refers to owning a tank that can hold at least 2500 liters of water.
Milk production is reported in liters.
Standard errors clustered at the household level are reported in brackets.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Milk sales

	(1)	(2)	(3)	(4)	(5)	(6)
	Sold milk	Milk sales	Milk sales, 5% trim	Sold milk	Milk sales	Milk sales, 5% trim
Treat*Post	0.034*	1.851	8.942*	0.037**	7.379	10.246**
	[0.018]	[13.269]	[4.898]	[0.017]	[6.070]	[4.703]
Treat*Post loan maturation				−0.010	−0.330	−3.854
				[0.019]	[6.913]	[5.476]
Treatment	−0.021	−2.428	−6.623	−0.021	−4.216	−6.623
	[0.017]	[10.708]	[5.124]	[0.017]	[6.541]	[5.125]
Constant	0.484***	44.517***	45.222***	0.484***	45.893***	45.222***
	[0.018]	[8.310]	[4.299]	[0.018]	[5.259]	[4.299]
TreatPost + TreatPost Maturation				0.028	7.049	6.393
SE				0.025	8.675	6.893
Dep Var Mean	0.690	186.474	159.187	0.690	159.187	131.890
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
HH Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	78476	78476	74556	78476	77693	74556

Note: All data is from administrative sources and covers all treatment groups.
Data is for each household for each month from July 2009 to May 2013.
Milk sales are reported in liters.
Treatment is defined as being offered a 4% deposit loan.
In column (3) and (6), sales are trimmed by excluding the top five percent of sales.
All specifications include month fixed effects.
Standard errors clustered at household level are reported in brackets.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Time use impacts on children 5-16 (minutes per day)

	Full sample		Piped water		No piped water	
	(1) Fetching water	(2) Tending livestock	(3) Fetching water	(4) Tending livestock	(5) Fetching water	(6) Tending livestock
Treatment*Female	−2.21*	5.57	−2.35	−16.56*	−1.98	13.61*
	[1.32]	[6.15]	[2.24]	[9.81]	[1.61]	[7.57]
Treatment	−0.96	−9.66*	0.45	5.01	−1.55	−14.84**
	[1.03]	[5.72]	[1.53]	[8.73]	[1.27]	[7.13]
Female	3.30***	−28.05***	2.94*	−18.47**	3.33**	−31.64***
	[1.09]	[5.27]	[1.74]	[7.31]	[1.34]	[6.67]
Constant	8.11***	30.59***	6.30**	25.11***	8.86***	32.81***
	[1.14]	[4.57]	[1.89]	[6.01]	[1.38]	[5.91]
Effect for Girls	−3.171***	−4.085	−1.902	−11.554**	−3.525**	−1.232
SE	[1.182]	[3.748]	[1.693]	[4.879]	[1.458]	[4.748]
Dep Var Mean	5.515	28.356	3.438	25.539	6.246	29.346
Observations	4109	4109	1069	1069	3040	3040

Note: All time use variables are in minutes per day per child. Analysis includes data from the early 2011 follow-up, Sept 2011, Feb 2012, May 2012, and Sept 2012 surveys. All specifications include time (survey round) fixed effects. Standard errors clustered at the household level are reported in brackets.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11: School enrollment impacts of tanks (children, 5-16)

	(1) Enrolled girl (5-16) dummy	(2) Enrolled boy (5-16) dummy
Treatment*Post	0.040**	−0.009
	[0.019]	[0.020]
Treatment	-0.012	0.001
	[0.012]	[0.011]
Post	−0.047***	−0.034**
	[0.016]	[0.016]
Constant	0.984***	0.983***
	[0.008]	[0.009]
Observations	1088	1080

Note: Enrollment variable equals one if the child is enrolled in school.

Panel observations only, so observations are excluded if the child was younger than five at baseline. Aging of the sample thus likely accounts for downward trend in enrollment captured by the coefficient on Post.

Standard errors clustered at the household level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$