Identification Strategy:
A Field Experiment on Dynamic Incentives in Rural Credit Markets*

Xavier Giné
Development Economics Research Group, World Bank
and Bureau for Research and Economic Analysis of Development (BREAD)

Jessica Goldberg
Ford School of Public Policy and Department of Economics, University of Michigan

Dean Yang
Ford School of Public Policy and Department of Economics, University of Michigan;
Bureau for Research and Economic Analysis of Development (BREAD);
and National Bureau of Economic Research (NBER)

May 2009

Abstract

How do borrowers respond to improvements in a lender's ability to punish defaulters? We implemented a randomized field experiment in Malawi examining the impact of fingerprinting of borrowers, which improves the lender's ability to withhold future loans from individuals who have previously defaulted while rewarding good borrowers with increased loan sizes. Study participants were smallholder farmers applying for input loans for growing a cash crop, paprika. Farmers were randomly allocated to either: 1) a control group, or 2) a treatment group that was fingerprinted as part of the loan application. Both treatment and control groups were given a presentation on the importance of credit history in ensuring future access to credit. For the subgroup of farmers with the highest ex ante default risk, fingerprinting led to substantially higher repayment rates. By contrast, fingerprinting had no impact on repayment for farmers with low ex ante default risk. Additional evidence indicates that, in the high-default-risk subgroup, fingerprinting resulted in higher repayment due to reductions in adverse selection (smaller loan sizes) and lower moral hazard (e.g., less diversion of fertilizer from the paprika crop).

Keywords: credit, microfinance, adverse selection, moral hazard, enforcement

---

* Gine: xgine@worldbank.org. Goldberg: jegoldbe@umich.edu. Yang: deanyang@umich.edu. Santhosh Srinivasan deserves the highest accolades for top-notch field work management and data collection. Lutamyo Mwamlima and Ehren Foss were key contributors to the success of the field work, and Niall Keleher helped organize wrap-up data entry. We appreciate the consistent and active support of Michael Carter, Charles Chikopa, Sander Donker, Lena Heron, David Rohrbach, and Mark Visocky. This project was funded by the World Bank Research Committee, USAID's BASIS AMA CRSP research facility, and the USAID Malawi country office.

1. Introduction

A central problem in many developing countries is the absence of a national system that allows individuals to be uniquely identified. Under these circumstances, loan defaulters can often avoid sanction by simply applying for new loans under different identities. Lenders respond by limiting the supply of credit, particularly in rural areas due to the inability to sanction unreliable borrowers and, conversely, to reward reliable borrowers with expanded credit.

As a result, many smallholder farmers are severely constrained by the inability to finance crucial inputs such as fertilizer and improved seeds, particularly for export crops. The following quote from 1973 by Robert McNamara when he was the World Bank president exemplifies this view: "The miracle of the Green Revolution may have arrived, but for the most part, the poor farmer has not been able to participate in it. He simply cannot afford to pay for the irrigation, the pesticide, the fertilizer… For the small holder operating with virtually no capital, access to capital is crucial."

The idea that credit markets may be mired in problems of asymmetric information and imperfect enforcement is not new. The theoretical literature has characterized debtors as individuals who may not reveal their borrowing type truthfully (adverse selection), may fail to exert effort to maximize the profits of their productive enterprise (ex-ante moral hazard), may misreport their production output (ex-post moral hazard), or may default on the loan even when output was sufficient (opportunistic default).

Textbook treatments include Laffont and Martimort (2003), Macho-Stadler and Perez-Castillo (2001), and Salanie (1997). An emerging empirical literature has examined various aspects of the limited commitment problem in credit markets. Chiappori (forthcoming) surveys the literature related to developed countries. Gine and Klonner (2005) document that asymmetric information on Tamilnadu fishermen's ability affects their access to credit for technological innovations. Karlan and Zinman (forthcoming) conduct a field experiment that shows the significant role of adverse selection and moral hazard in loan defaults in South Africa. Ligon, Thomas and Worall (1999) and Paulson, Townsend, and Karaivanov (2006) provide empirical evidence on opportunistic default in India and Thailand respectively. Visaria (forthcoming) documents the positive impact of expedited legal proceedings on loan repayment among large Indian firms.

In this paper we implement a randomized field experiment to estimate the impact of biometric identification (fingerprinting) in a context—rural Malawi—where credit supply has been limited due to difficulties in enforcing sanctions against defaulters. Fingerprinting makes the threat of future credit denial credible because it makes it easier for financial institutions to withhold new loans from past defaulters, and to reward responsible past borrowers with increased credit. This potentially reduces the various types of limited commitment problems outlined above and therefore raises repayment. Adverse selection problems should decline, because in an environment of asymmetric information and credible ex post sanction, riskier individuals should take out smaller loans or avoid borrowing altogether. Borrowers should have greater incentives to ensure that production is successful, either by exerting more effort or choosing less risky projects (lower moral hazard), and—whenever production could cover the loan repayment—should be less likely to default intentionally or opportunistically.

In this study, smallholder farmers organized in groups of 15-20 members applied for agricultural input loans to grow paprika and were randomly allocated to either: 1) a control group, or 2) a treatment group where each member had a fingerprint collected as part of the loan application. Both treatment and control groups were given a training session on the importance of credit history in ensuring future access to credit.

For the subgroup of farmers with the highest ex ante default risk, fingerprinting led to substantially higher repayment rates. By contrast, fingerprinting had no impact on repayment for farmers with low ex ante default risk. Additional evidence indicates that higher repayment for the high-default-risk subgroup is due to reductions in adverse selection (smaller loan sizes) and lower moral hazard (more land devoted to paprika, and less diversion of fertilizer from the paprika crop).

The remainder of this paper is organized as follows. Section 2 describes the experimental design and survey data, Section 3 describes the regression specifications, and Section 4 presents the empirical results. Section 5 provides additional discussion and robustness checks, and Section 6 concludes.

2. Experimental design and survey data

The experiment was carried out as part of the Biometric and Financial Innovations in Rural Malawi (BFIRM) project, a cooperative effort among Cheetah Paprika Limited (CP) and the Malawi Rural Finance Corporation (MRFC). CP is a privately owned agri-business company established in 1995 that offers extension services and high-quality inputs to smallholder farmers via an out-grower paprika scheme.[1] The farmer receives extension services and a package of seeds, pesticides and fungicides at subsidized rates in exchange for the commitment to sell the paprika crop to CP at harvest time. CP is by far the largest paprika purchaser in the country.[2] CP has a staff of six extension officers and 15 field assistants in the locations chosen for the study. The staff maintained a database of all current and past paprika growers and handled the logistics of supplying farmers with the package of inputs as well as the purchase and payment of the crop.

MRFC is a government-owned microfinance institution that provided financing for the in-kind loan package for 1/2 to 1 acre of paprika designed by CP. The loan did not include any cash to buy the inputs. Instead, borrowers had to take an authorization form from MRFC to a preferred agricultural input supplier who in turn delivered the goods and billed MRFC at a later date. The loan amount was roughly 17,000 Malawi Kwacha (approximately $120). Sixty percent of the loan went towards fertilizer (one 50 Kg bag of D-compound fertilizer and two 50 Kg bags of CAN fertilizer); the rest went toward the CP input package: thirty-three percent covered the cost of nine bags of pesticides and fungicides (2 Funguran, 2 Dithane, 2 Benomyl, 1 Cypermethrin, 1 Acephate and 1 Malathion) and the remaining seven percent for the purchase of 0.4 Kg of seeds. While all farmers that took the loan were given the CP package, farmers had the option to borrow only one of the two available CAN bags. Expected yield for farmers using this package on one acre of land was between 400 and 600 kg, compared to 200 kg with no inputs.[3]

[1] Extension services consist of preliminary meetings to market paprika seed to farmers and teach them about the growing process, additional group trainings about farming techniques, individual support for growers provided by the field assistants, and information about grading and marketing the crop.
[2] In 2007, CP purchased approximately eighty-five percent of the one thousand tons of paprika produced annually in Malawi.
[3] Yield is computed under the conservative assumption that farmers will divert one 50 Kg bag of CAN fertilizer towards maize cultivation. While larger quantities of inputs would result in higher output for experienced paprika-growers, the package described here was designed by extension experts to maximize expected profits for novice, small-holder growers.

In keeping with standard MRFC practices, farmers were expected to raise a 20 percent deposit, and were charged interest of 33 percent per year (or 30 percent for repeat borrowers). Within a group, take-up of the loan was an individual decision, but the subset of farmers who took up the loan was told that they were jointly liable for each others' loans. In practice, however, joint liability schemes in Malawi are seldom enforced.[4]

The timeline of the experiment is presented in Figure 1. In July 2007, CP told farmers in the study areas to re-organize themselves into clubs of 15 to 20 members to accommodate MRFC's group lending rules.[5] Many of these clubs were already in existence, primarily to ease delivery of Cheetah extension services and collection of the crop. Our study sample consists of 249 clubs and approximately 3,500 members in Dedza, Mchinji, Dowa and Kasungu districts. Figure 2 illustrates the locations of the study areas on a map of Malawi.

Farmer clubs in the study were randomly assigned to be fingerprinted (treatment) or not (control), with an equal probability of being in either the treatment or control group. During the months of August and September of 2007, CP staff provided the list of all paprika growing clubs in each area to be visited that week, and the BFIRM team carried out the assignment of treatment status. Randomization was carried out after stratifying by geographical area (EPA) and week of club visit.

A club visit always included a survey and a training session. Both treatment and control groups were given a presentation on the importance of credit history in ensuring future access to credit. In addition, fingerprints were collected for treatment clubs, and in treatment clubs farmers were told how their fingerprint uniquely identified them for credit reporting to all major Malawian rural lenders. The training emphasized that defaulters would face exclusion from future borrowing, while borrowers in good standing could be rewarded with larger loans in the future, and that future credit providers would be able to access the applicant's credit history simply by checking his or her fingerprint. Appendix A provides the script used during the training.

[4] See Giné and Yang (2009) for another example of limited enforcement of joint liability loans.

[5] A typical CP group has between 15 and 30 farmers and is organized around a paprika collection point. MRFC's lending groups have at most 20 farmers, so most of the CP groups participating in the study had to be split to be able to access MRFC's loans.

Loan applicants from fingerprinted clubs had their right thumb fingerprint recorded electronically during the loan application process. Members from the BFIRM team carried a laptop with a fingerprint scanner (FB80 Pro) connected to it. The laptop used the VeriFinger 5.0 software to collect fingerprints.[6] After they were collected, a demonstration program was used to show participants that the laptop was now able to identify an individual with only his or her fingerprint. One farmer was chosen at random, his or her right thumb was again placed on the scanner, and immediately his or her name along with the demographic information provided appeared on the screen. The control group was not fingerprinted, but as already mentioned, also received the same training emphasizing the importance of one's credit history and how it influences one's future credit access.

Prior to the training and the collection of fingerprints, farmers were administered a brief household socioeconomic survey.[7] The survey included questions on individual demographics (education, household size, religion), income generating activities and assets including detailed information on crop production and crop choice, livestock and other assets, risk preferences, past and current borrowing activities, and past variability of income. Summary statistics from the baseline survey are presented in Table 1, and variable definitions are provided in Appendix B.

After the completion of the survey, the names and locations of the members that applied for loans were handed over to MRFC credit officers so that they could screen and approve the clubs according to their protocols. Among other standard factors, MRFC conditions lending on the club's successful completion of 16 hours of training. MRFC approved loans for 2,063 customers in 121 clubs (out of 239). Of the customers approved for loans, some failed to raise the required down payment and others opted not

[6] We purchased the VeriFinger 5.0 Software Development Kit and had a programmer develop a data capture program that would allow the user to (i) enter basic demographic information such as the name, address, village, loan size and the unique BFIRM identifier, (ii) capture the fingerprint with the scanner and (iii) review the fingerprint alongside the demographic information. The database created was then merged with MRFC's administrative data.

[7] These survey data were collected prior to the farmers' being informed about the role of biometrics in the project and their treatment status, to ensure that farmers' survey answers were not influenced by knowledge of the nature of the experiment.

to borrow for other reasons. The final sample consists of 1,147 loan customers.[8] These customers received loan packages with an average value of MK 16913 ($121 US).

The project also implemented follow-up surveys of farmers in April 2008, a month before paprika harvest and in August 2008, once crops had been sold and income received. The formal loan maturity (payment) date was September 30, 2009. In the analysis we use April and August 2008 survey data along with administrative data from both CP and MRFC.

Balance of baseline characteristics across treatment vs. control groups

To confirm that the randomization across treatments achieved the goal of balance in terms of pre-treatment characteristics, Table 2 presents the means of several baseline variables for the control group as reported prior to treatment, alongside the difference vis-à-vis the treatment group (mean in treatment group minus mean in control group). We also report statistical significance levels of the difference in treatment-control means. These tests are presented for both the full baseline sample as well as the loan recipient sample.

Overall, there is substantial balance between the two groups in both the full baseline sample and the loan recipient sample. In the full baseline sample, the difference in means for the treatment and control groups is not significant for any of the 11 baseline variables. In the loan recipient sample, for 10 out of these 11 baseline variables, the difference in means between treatment and control groups is not statistically significantly different from zero at conventional levels, and so we cannot reject the hypothesis that the means are identical across treatment groups. For only one variable, the indicator for the study participant being male, is the difference statistically significant (at the 10% level): the fraction male in the treatment group is 6.6 percentage points lower than in the control group.[10]

---

[8] While a natural question at this point is whether selection into borrowing was affected by treatment status, treatment and control groups did not differ in their rates of MRFC loan approval or the fraction of farmers who ended up with a loan (as will be detailed in the results section below).

[10] It will turn out, however, that the regression results to come are not substantially affected by the inclusion or exclusion in the regressions of a large set of control variables (including the "male" indicator).

3. Regression Specification

Because the treatment is assigned randomly at the club level, its impact on the various outcomes of interest (say, repayment) can be estimated via the following regression equation:

(1) $Y_{ij} = \alpha + \beta B_j + \gamma X_{ij} + \varepsilon_{ij},$

where Y_{ij} = repayment outcome for individual *i* in club *j* (e.g., equal to 1 if repaying in full and on time, and 0 otherwise), B_j is biometric identification (1 if fingerprinted and 0 if not), and X_{ij} is a vector of club and individual farmer characteristics collected at baseline. ε_{ij} is a mean-zero error term. Treatment assignment at the club level creates spatial correlation among farmers within the same club, so standard errors must be clustered at the club level (Moulton 1986). Inclusion of the vector X_{ij} of baseline characteristics can reduce standard errors by absorbing residual variation. In our case, we include the baseline characteristics reported in Table 1, as well as indicators for the two stratification variables (loan officer fixed effects and week of loan offer fixed effects) and all interactions between the dummy variables for loan officer and week of loan offer.

The coefficient β on the biometric treatment status indicator is the impact of being fingerprinted on the dependent variable of interest.

We also examine the interactions between the randomized treatment and baseline characteristics. In particular we are interested in the ex-ante probability of default. For example, it may be the case that those whose ex-ante characteristics make them more likely to default on the loan see larger improvements in repayment in response to fingerprinting, compared to individuals who are very likely to repay. To test this question, the following regression equation is useful:

(2) $Y_{ij} = \alpha + \rho(B_j * D_j) + \beta B_j + \chi X_{ij} + \varepsilon_{ij},$

D_{ij} is a variable representing the individual's predicted likelihood of repayment (its main effect is included in the vector X_{ij}). The coefficient ρ on the interaction term B_j *

D_j reveals the extent to which the impact of biometric identification's on repayment varies according to the borrower's predicted repayment.

To implement equation (2) examining heterogeneity in the effect of fingerprinting, we construct an index of predicted repayment. This index is either interacted linearly with the treatment indicator, or it is converted into indicators for quintiles of the distribution of predicted repayment and then interacted with the treatment indicator. (See the Appendix for details on the construction of the predicted repayment variable.) In all regression results where the fingerprinting indicator is interacted with predicted repayment, we report bootstrapped standard errors because the predicted repayment variable is a generated regressor.[11]

4. Empirical Results: Impacts of Fingerprinting

This section presents our experimental evidence on the impacts of fingerprinting on a variety of inter-related outcomes. We examine impacts on loan approval and borrowing decisions, on repayment outcomes, and on intermediate farmer actions and outcomes that may ultimately affect repayment.

Tables 3 through 7 will present regression results from estimation of equations (1) and (2) in a similar format. In each table, each column will present regression results for a given dependent variable. Panel A will present the coefficient on treatment (fingerprint) status from estimation of equation (1).

Then, to examine heterogeneity in the effect of fingerprinting, Panels B and C will present results from estimation of versions of equation (2) where fingerprinting is interacted linearly with predicted repayment (Panel B) or with dummy variables for quintiles of predicted repayment (Panel C). In both Panels B and C the respective main effects of the predicted repayment variables are also included in the regression (but for brevity the coefficients on the predicted repayment main effects will not be presented). In Panel C, the main effect of fingerprinting is not included in the regression, to allow each

[11] We calculate standard errors for regressions in the form of equation (2) from 200 bootstrap replications. In each replication, we re-sample borrowing clubs from our original data, compute predicted repayment based on the new sample, and re-run the regression in question using the new value of predicted repayment for that replication.

of the five quintile indicators to be interacted with the indicator for fingerprinting in the regression. Therefore, in Panel C the coefficient on each fingerprint-quintile interaction should be interpreted as the impact of fingerprinting on borrowers in that quintile, compared to control group borrowers in that same quintile.

Finally, in Tables 3 through 7 the mean of the dependent variable in a given column will be given in the bottom row of the table.

A. Loan approval, take-up, and amount borrowed

The first key question to ask is whether fingerprinted farmers were more likely to have their loans approved by the lender, or were more likely to take out loans, compared to the control group. This question is important because the degree of selectivity in the borrower pool induced by fingerprinting status affects interpretation of any effects on repayment and other outcomes.

As it turns out, there is no indication that fingerprinting induced differential approval by the lender or loan-take-up by borrowers. Columns 1 and 2 of Table 3 present results from estimation of equations (1) and (2) for the full baseline sample where the dependent variables are, respectively, an indicator for the lender's approving the loan for the given farmer (mean 0.63), and an indicator for the farmer ultimately taking out the loan (mean 0.35).[12]

There is no evidence that the rate of loan approval or take-up differs substantially across the treatment and control groups on average: the coefficient on fingerprinting is not statistically different from zero in either columns 1 or 2, Panel A.

There is also no indication of selectivity in the resulting borrowing pool across subgroups of borrowers with different levels of predicted repayment. The coefficient on the interaction of fingerprinting with predicted repayment is not statistically significantly different from zero in either columns 1 or 2 of Panel B. When looking at interactions with quintiles of predicted repayment (Panel C), while the fingerprint-quintile 2 interaction is positive and significantly different from zero at the 10% level in the loan approval

[12] Not all farmers who were approved for the loan ended up taking out the loan. Anecdotal evidence indicates that a substantial fraction of non-take-up among approved borrowers resulted when borrowers failed to raise the deposit amounting to 20% of the loan amount.

regression, none of the interaction terms with fingerprinting are significantly different from zero in the loan take-up regression.

While there is no indication that the pool of ultimate borrowers was itself substantially affected by fingerprinting, it does appear that – conditional on borrowing – fingerprinted borrowers took out smaller loans. In Column 3 of Table 3, the dependent variable is the total amount borrowed in Malawi kwacha (mean 16,912.60, or roughly US$121). Panel A indicates that loans of fingerprinted borrowers were MK 697 (roughly US$5) smaller than loans in the control group on average, a difference that is significant at the 10% level.

Inspecting the coefficients on the interactions of fingerprinting with predicted repayment, it appears that this effect is confined exclusively to borrowers in the lowest quintile of expected repayment. Differences between fingerprinted and non fingerprinted borrowers are small and not significant in quintiles two through four, but in quintile one, where fingerprinted borrowers take out loans that are smaller by MK 2,722 (roughly US$19) than those in the corresponding quintile in the control group, the difference is marginally significant (the t-statistic is 1.63).

We view this result – voluntarily lower borrowing amounts on the part of fingerprinted borrowers in the lowest quintile – as evidence that fingerprinting reduces adverse selection in the credit market, albeit on a different margin than is usually thought of in the credit context. The existing literature tends to emphasize that improved enforcement should lead low-quality borrowers to be excluded from borrowing entirely – in other words, the improvement of the borrower pool operates on the *extensive* margin of borrowing. Our result here that low-quality borrowers (those in the lowest quintile of predicted repayment) voluntarily take out smaller loans leads the overall loan pool in money terms to be less weighted towards the low-quality borrowers, but in this case the improvement in the borrowing pool operates on the *intensive* margin of borrowing, rather than the extensive margin.

Interpretation of subsequent differences in the repayment rates (discussed below) should keep this result in mind. Improvements in repayment among fingerprinted borrowers (particularly among those in the lowest quintile) may in part result from their decisions to take out smaller loans at the very outset of the lending process and improve their eventual likelihood of repayment.

B. Loan repayment

How did fingerprinting affect ultimate loan repayment? Columns 1-3 of Table 4 present estimated effects of fingerprinting for the loan recipient sample on three outcomes: outstanding balance (in Malawi kwacha), fraction of loan paid , and an indicator for whether the loan is fully paid, all by September 30, 2009 (the official due date of the loan, after which the loan is officially past due).

Results for all three outcomes are similar: fingerprinting improves loan repayment, in particular for borrowers expected ex ante to have poorer repayment performance. Coefficients in Panel A indicate that fingerprinted borrowers have lower outstanding balances, higher fractions paid, and are more likely to be fully paid as of the due date (and the coefficient in the regression for fraction paid is statistically significant at the 10% level).

In Panel B, the fingerprinting/predicted repayment interaction term is statistically significantly different from zero at the 5% level for outstanding balance and the 1% level for the other two repayment outcomes. The effect of fingerprinting on repayment is larger the lower is the borrower's ex ante likelihood of repayment. In Panel C, it is evident that the effect of fingerprinting is isolated in the lowest quintile of expected repayment, with coefficients on the fingerprint-quintile 1 interaction all being statistically significantly different from zero at the 5% or 1% level and indicating beneficial effects of fingerprinting on repayment (lower outstanding balances, higher fraction paid, and higher likelihood of full repayment). Coefficients on other fingerprint-quintile interactions are all smaller in magnitude and not statistically significantly different from zero (with the exception of the negative coefficient on the fingerprint-quintile 5 interaction for fraction paid, which is odd and may simply be due to sampling variation).

The magnitudes of the repayment effect found for the lowest predicted-repayment quintile are large. The MK7,202.65 effect on outstanding balances amounts to 40% of the average loan size for borrowers in the lowest predicted-repayment quintile. The 49.9

percentage point increase in fraction paid and the 54.3 percentage point increase in the likelihood of being fully paid are also large relative to bottom quintile percentages of 62% and 52% respectively.

C. Intermediate outcomes that may affect repayment

In this section we examine decisions that farmers make throughout the planting and harvest season that may contribute to higher repayment among fingerprinted farmers. The dependent variables in the remaining results tables (Tables 5-7) are available from a smaller subset of loan recipients (N=520) that were successfully interviewed in the August 2008 follow-up survey round.[13] Columns 4-6 of Table 4 present regression results for repayment outcomes that are analogous to those in columns 1-3, but where the sample is restricted to this 520-observation sample. The results confirm that the repayment results in the 520-observation sample are very similar to those in the overall loan recipient sample, in terms of both magnitudes of effects and statistical significance levels.

Land area allocated to various crops

One of the first decisions that farmers make in any planting season (that typically starts in November and December) is the proportion of land allocated to different crops. Table 5 examines the average and heterogeneous impact of fingerprinting on land allocation; the dependent variables across columns are fraction of land used in maize (column 1), 7 cash crops (columns 2-8), and all cash crops combined (column 9).[14]

Why might land allocation to different crops respond to fingerprinting? In the absence of fingerprinting, farmers might divert land and other resources towards other crops and away from paprika, even though the loan is intended for paprika inputs. This is a form of moral hazard: diversion of paprika inputs towards maize (the primary staple

[13] Column 2 of Appendix Table 1 examines selection of loan recipients into the 520-observation August 2008 survey sample. The regressions are analogous in structure to those in the main results tables (Panels A, B, and C), and the dependent variable is a dummy variable for attrition from the baseline (September 2007) to the August 2008 survey. In no case is fingerprinting or fingerprinting interacted with predicted repayment statistically significantly associated with attrition from the survey.

[14] For each farmer, the value of the variables across columns 1-8 add up to 1.

crop) may be preferred by the farmer because it is seen as less risky, but it reduces the likelihood of loan repayment if land and other inputs generate lower economic profits when devoted to maize. Fingerprinting may then discourage such diversion of inputs and land to other crops, as farmers face increased incentives to generate cash profits that are sufficient for loan repayment.

While none of the effects of fingerprinting in Table 5 (either overall in Panel A or in interaction with predicted repayment in Panels B and C) are statistically significant at conventional levels, there is suggestive evidence that there is an impact of fingerprinting on land allocation for borrowers in the first predicted-repayment quintile. In this group, the effect of fingerprinting on land allocated to paprika (column 5, first row of Panel C) is marginally significant (with a t-statistic of 1.63) and positive, indicating that fingerprinting leads farmers to allocate 8.3 percentage points more land to paprika.

It is worth considering that the effect on land allocated to paprika may be smaller than it might be otherwise because preparing and allocating land took place earlier in the agricultural season than our treatment. If land is less easily reallocated from one crop to another, then we would anticipate smaller short run effects on land allocation than the use of inputs such as fertilizer and chemicals (to which we turn next). In the long run, when farmers incorporate the additional cost of default due to fingerprinting into their agricultural planning earlier in the season, we might find larger impacts on land allocation.

Inputs used on paprika

After allocating land to different crops, the other major farming decision made by farmers is input application. Input allocation takes place later in the agricultural cycle than land allocation, and agricultural inputs are more fungible than land. Also, inputs are added multiple times throughout the season, so farmers can incorporate new information about the costs of default into their use of inputs but cannot change land allocation after planting. Thus, we may expect use of inputs to respond more quickly to the introduction of fingerprinting than would allocation of land.

Table 6 examines the effect of fingerprinting on the use of inputs on the paprika crop.[15] The dependent variables in the first 5 columns (all denominated in Malawi kwacha) are applications of seeds, fertilizer, chemicals, man-days (hired labor), and all inputs together. Columns 6 and 7 look at, respectively, manure application (denominated in kilograms because this input is typically produced at home and not purchased) and the number of times farmers weeded the paprika plot. We view the manure and weeding dependent variables as more purely capturing labor effort exerted on the paprika crop, while the other dependent variables capture both labor effort and financial resources expended.

The results for paid inputs (columns 1-5) indicate that – particularly for farmers with lower likelihood of repayment – fingerprinting leads to higher application of inputs on the paprika crop. In Panel B, the coefficients on the fingerprint-predicted repayment interaction are all negative in sign, and the effects on the use of fertilizer and paid inputs in aggregate are statistically significantly different from zero. In Panel C, the coefficient on the fingerprint-quintile 1 interaction is positive and significantly different from zero at the 5% confidence level for spending on seeds and is marginally significant for spending on fertilizer (t-statistic 1.44) and for all paid inputs (t-statistic 1.55). The negative and significant impact on use of paid labor in the fourth quintile is puzzling and may be attributable to sampling variation.

Results for inputs not purchased in the market are either nonexistent or ambiguous. No coefficient is statistically significantly different from zero in the regressions for manure (column 6) or times weeding (column 7).

In sum: for borrowers with lower likelihoods of repayment, fingerprinting leads to increased expenditure of financial resources on market-purchased inputs used in growing paprika. While this effect is often only marginally significant for borrowers in the lowest predicted repayment quintile, the magnitudes in that quintile are substantial. For the lowest predicted-repayment subgroup, fingerprinted farmers used MK6,541 (US$47)

[15] We also examined the impact of fingerprinting on use of inputs on all crops combined. Results were very similar to Table 6's results for input use on the paprika crop only, indicating that essentially all documented increases in input application occurred on the paprika crop. (Results are available from the authors on request.)

more paid inputs in total, which is substantial compared to the mean in the lowest predicted-repayment subgroup of MK7,440 (US$53).

Farm profits

Given these effects of fingerprinting on intermediate farming decisions such as land allocation and input use, what is the effect on agricultural revenue and profits?

Columns 1-3 of Table 7 present regression results where the dependent variables are market crop sales, the value of unsold crops, and profits (market sales plus value of unsold crops minus value of inputs used), all denominated in Malawi kwacha. The magnitude of the differences between the value of sales, unsold harvest, and overall profits for fingerprinted and non fingerprinted farmers overall (Panel A), and in the bottom two quintiles (Panel C) is large and positive, the effects are imprecisely estimated and not statistically significant.

To help deal with the problem of outliers in the profit figures, column 4 presents regression results where the dependent variable is the natural log of agricultural profits.[16] The effect of fingerprinting in the bottom quintile of predicted repayment is positive but not statistically significant (t-statistic 1.11).

In sum, then, it remains possible that increased use of paid inputs led ultimately to higher revenue and profits among fingerprinted farmers in our sample, but the imprecision of the estimates prevents us from making strong statements in this regard.

5. Discussion and additional robustness checks

In sum, the results indicate that for the lowest predicted-repayment quintile, fingerprinting leads to substantially higher loan repayment. In seeking explanations for this result, we have provided evidence that for this subgroup fingerprinting leads farmers to take out smaller loans, devote more land to paprika, and apply more inputs on paprika.

[16] For seven (7) observations profits are zero or negative, and in these cases ln(profits) is replaced by 0. These observations are not driving the results, as results are essentially identical when simply excluding these 7 observations from the regression.

We view these results so far as indicating that – for the farmers with the lowest ex ante likelihood of repaying their loans – fingerprinting leads to reductions in adverse selection and ex-ante moral hazard. The reduction in adverse selection (a reduction in the riskiness of the loan pool) comes about not via the extensive margin of loan approval and take-up, but through farmers' decisions to take out smaller loans if they are fingerprinted (the intensive margin of loan take-up).

Ex-ante moral hazard is the problem that borrower behavior that is unobserved to the lender may be detrimental for repayment. We interpret changes in intermediate outcomes and behaviors – such as increased land use and input application for paprika – as reductions in ex-ante moral hazard. It is possible that in the absence of fingerprinting, borrowers in the lowest predicted-repayment subgroup were not using the paprika inputs received via the loan transaction on their farms for paprika (they could, for example, been selling them in the market or bartering them away). Then when such borrowers were fingerprinted, they became more likely to use the inputs as intended, expanding land allocated to paprika and using the inputs on that crop as the loan required.

Evidence for a reduction in ex-post moral hazard

Reductions in *ex-ante* moral hazard help encourage higher loan repayment by improving farm output so that farmers have higher incomes with which to make loan repayments. Reductions in adverse selection – reduced loan sizes for the "worst" borrowers – also help increase repayment performance. But a question that remains is whether any of the increase in repayment is due to reductions in *ex-post* moral hazard. In other words, are there reductions in strategic or opportunistic default by borrowers, holding constant loan size and farm profits?

We investigate this by running regressions for repayment outcomes analogous to regressions in columns 4-6 of Table 4, but where we include as independent variables in the regression controls for agricultural profits and the total originally borrowed.[17] The profits and total borrowed variables are flexibly specified as indicators for the borrower

[17] We limit ourselves to the 520-observation sample because of the need to control for profits, which was only observed among those in the August 2008 survey.

being in the 1st through 10th decile of the distribution of the variable (one indicator is excluded in each resulting group of 10 indicators, so there so there are 18 additional variables in each regression.)

We cannot reject the hypothesis that fingerprinting has no effect on repayment once we control for agricultural profits and original loan size. Coefficient estimates that were previously statistically significant in columns 4-6 of Table 4 are now uniformly smaller in magnitude and not statistically significantly different from zero. Indeed, the previously significant coefficients on the fingerprint * quintile 1 interaction across the columns are roughly cut in half.

There is no evidence, therefore, that a reduction in ex-post moral hazard – increases in repayment even conditional on amount borrowed and agricultural profits – is also an important contributor to the increased repayment we observe among fingerprinted farmers in the lowest predicted-repayment quintile.

Are the results due to selection?

It is also important to ask whether any of observed impacts of fingerprinting could be due to selection. Several pieces of evidence help rule out this possibility. First of all, we have shown earlier (Table 2) that fingerprinting has no effect on the composition of the treatment and control groups according to baseline variables. In addition, we find no effect of fingerprinting in the distribution of borrowers across quintiles in the treatment and control groups. Finally, we show that there is no effect of fingerprinting on the predicted repayment index among borrowers, in the treatment control groups overall as well as in quintile-by-quintile comparisons. (Results are available from authors on request.)

6. Conclusion

For all the recent empirical work on the imperfections in credit markets in developing countries, to our knowledge is the first research that directly estimates the impact of improved enforcement on loan repayment in rural areas. Such an estimate is highly valuable from a theoretical standpoint in clarifying the extent to which imperfect enforcement contributes to high default rates and thus low supply of credit.

In discussions of potential public policies that can help increase the supply of credit to rural areas, an often-cited central priority is the establishment of institutions such as credit bureaus that can effectively create public information on a borrower's past borrowing history (Conning and Udry 2005, Fafchamps 2004). These findings also provide evidence for the benefits of establishing a national credit bureau in Malawi (and elsewhere) to centralize such information that uses fingerprints as the unique identifier.

Appendix A: Biometric training script

Benefits of Good Credit
Having a record of paying back your loans can help you get bigger loans or better interest rates.

Credit history works like trust. When you know someone for a long time, and that person is honest and fair when you deal with him, then you trust him. You are more likely to help him, and he is more likely to help you. You might let him use your hoe (or something else that is important to you), because you feel sure that he will give it back to you. Banks feel the same way about customers who have been honest and careful about paying back their loans. They trust those customers, and are more willing to let them borrow money.

MRFC already gives customers who have been good borrowers a reward. It charges them a lower interest rate, 30 percent instead of 33 percent. That means that for the loan we have described today, someone who has a good credit history would only have to pay back 8855, instead of 8971.

Another way that banks might reward customers they trust is by letting them borrow bigger amounts of money. Instead of 7700 MK to grow one acre of paprika, MRFC might lend a trusted customer 15400, to grow two acres.

To earn trust with the bank, and get those rewards, you have to be able to prove to the bank that you have taken loans before and paid them back on time. You can do that by making sure that you give the bank accurate information when you fill out loan applications. But if you call yourself John Jacob Phiri one year, and Jacob John Phiri the next year, then the bank might not figure out that you are the same person, so they won't give you the rewards you have earned.

Costs of Bad Credit
But trust can be broken. If your neighbor borrows your radio and does not give it back or it gets ruined, then you probably wouldn't lend him anything else until the radio had been replaced.

Banks work the same way. If you take a loan and break the trust between yourself and the bank by not paying back the loan, then the bank won't lend to you again. This is especially true if you have a good harvest but still choose not to pay back the loan.

When you apply for a loan, one of the things that a bank does to decided whether or not to accept your application is that it looks in it's records to see if you have borrowed money before. If you have borrowed but not paid back, then you will be turned down for the new loan. This is like you asking your neighbors if someone new shows up in the village and asks you to work for him. You might first ask around to see if the person is fair to his employees and pays them on time. If you learn that the person does not pay his workers, then you won't work for him. Banks do the same thing by checking their records.

MRFC does not ever give new loans to people who still owe them money. And MRFC shares information about who owes money with other banks, so if you fail to pay back a loan from MRFC, it can stop you from getting a new loan from OIBM or another lender, also

Remainder of script is administered to fingerprinted clubs only

Biometric Technology
Fingerprints are unique, which means that no two people can ever have the same fingerprints. Even if they look similar on a piece of paper, people with special training, or special computer equipment, can always tell them apart.

Your fingerprint can never change. It will be the same next year as it is this year. Just like the spots on a goat are the same as long as the goat lives, but different goats have different spots.

Fingerprints can be collected with ink and paper, or they can be collected with special machines. This machine stores fingerprints in a computer. Once your fingerprint is stored in the computer, then the machine can recognize you, and know your name and which village you come from, just by your fingerprint! The machine will recognize you even if the person who is using it is someone you have never met before. The information from the machines is saved in many different ways, so if one machine breaks, the information is still there. Just like when Celtel's building burned, people's phone numbers did not change.

Administer the following after all fingerprints have been collected:

Demo
Now, I can figure out your name even if you don't tell me. Will someone volunteer to test me? (Have a volunteer swipe his finger, and then tell everyone who it was).

The bank will store information about your loans with your fingerprint. That means that bank officers will know not just your name, but also what loans you have taken and whether or not you have paid them back. They will be able to tell all of this just by having you put your finger on the machine.

Before, banks used your name and other information to find out about your credit history. But now they will use fingerprints to find out. This means that even if you tell the bank a different name, they will still be able to find all of your loan records. Names can change, but fingerprints cannot.

Having your fingerprint on file can make it easier to earn the rewards for good credit history that we talked about earlier. It will be easy for the bank to look up your records and see that you have paid back your loans before. It will also be easier to apply for loans, because there will be no new forms to fill out in the future!

But, having your fingerprint on file also makes the punishment for not paying back your loan much more certain. Even if you tell the bank a different name than you used before, or meet a different loan officer, or go to a different branch, the bank will just have to check your fingerprint to find out whether or not you paid your loans before. Having records of fingerprints also makes it easy for banks to share information. Banks will share information about your fingerprints and loans. If you don't pay back a loan to MRFC, OIBM will know about it!

Appendix B: Variable definitions
[TO BE ADDED]

Appendix C: Construction of predicted repayment variable

To construct the predicted repayment variable, we first limit the sample to individuals in the *control* group (N=563), and run a regression of a repayment outcome (fraction of loan repaid by September 30, 2008) on various farmer- and club-level baseline characteristics. Conceptually, the resulting index will be purged of any bias introduced by effects of fingerprinting on repayment because it is constructed using coefficients from a regression predicting repayment for only the control or non-fingerprinted farmers.

Appendix Table 2 presents results from this exercise. Statistically significant results in column 1, which only includes farmer-level (individual) variables on the right-hand-side, indicates that older farmers and those who do not self-identify as risk-takers have better repayment performance on the loan. Inclusion of a complete set of fixed effects for loan officer * week of initial loan offer interactions raises the R-squared substantially (from 0.05 in column 1 to 0.46 in column 2). The explanatory power of the regression is marginally improved further in column 3 (to an R-squared of 0.48) when the age and education are broken up into categorical variables (instead of being entered linearly).

We then take the coefficient estimates from column 3 of Appendix Table 2 and predict the fraction of loan repaid for the *entire* sample (both control and treatment groups). This variable, which we call "predicted repayment", is useful for analytical purposes because it is a single index that incorporates a wide array of baseline information (at the individual and locality/ loan officer level) correlated with repayment outcomes.

References

Ashraf, Nava, Dean Karlan, and Wesley Yin (2006), "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *Quarterly Journal of Economics*.

Banerjee, Abhijit V., "Contracting Constraints, Credit Markets, and Economic Development", in Mathias Dewatripont, Lars Peter Hansen and Stephen Turnovsky (eds.), *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Volume III*, Cambridge, U.K.: Cambridge University Press, 2003.

Banerjee, Abhijit V. and Andrew F. Newman, 1993. "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101(2), pp 274-298.

Bencivenga, Valerie R. and Bruce D. Smith. 1991. "Financial Intermediation and endogenous Growth" *Review of Economic Studies* 58, pp. 195-209.

Chiappori, Pierre Andre (forthcoming), "Econometric Models of Insurance under Asymmetric Information", in G Dionne (ed.), *Handbook of Insurance*: North Holland.

Conning, Jonathan and Christopher Udry, "Rural Financial Markets in Developing Countries," in R. E. Everson, P. Pingali, and T.P. Schultz (eds.), *The Handbook of Agricultural Economics, Vol. 3: Farmers, Farm Production, and Farm Markets*, Elsevier Science, 2005.

Duflo, Esther, Rachel Glennerster, and Michael Kremer, "Use of Randomization in Development Economics Research: A Toolkit," NBER Technical Working Paper T0333, December 2006.

Duflo, Esther, Michael Kremer, and Jonathan Robinson (2009), "Nudging Farmers to Use Fertilizer: Evidence from Kenya," working paper, UC Santa Cruz, MIT, and Harvard University.

Fafchamps, Marcel, *Market Institutions in Sub-Saharan Africa: Theory and Evidence*. MIT Press, 2004.

Giné, Xavier, and Stefan Klonner (2005), "Financing a New Technology in Small-scale Fishing: the Dynamics of a Linked Product and Credit Contract," working paper, World Bank.

Giné, Xavier and Dean Yang, "Insurance, Credit, and Technology Adoption: Field Experimental Evidence from Malawi," *Journal of Development Economics*, Vol. 89, 2009, pp 1-11.

Karlan, Dean, and Jonathan Zinman (forthcoming), "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment", *Econometrica*.

Laffont, Jean-Jacques, and David Martimort (2003), *The principal agent model: The economic theory of incentives*: Princeton University Press.

Ligon, Ethan, Jonathan P. Thomas, and Tim Worrall (2002), "Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies", *Review of Economic Studies* 69 (1):209-244.

Lloyd-Ellis, Huw & Bernhardt, Dan, 2000. "Enterprise, Inequality and Economic Development," *Review of Economic Studies,* Blackwell Publishing, vol. 67(1), pp. 147-68, January.

Macho-Stadler, and Perez-Castrillo (2001), *An Introduction to the Economics of Information: Incentives and Contracts*. 2nd ed: Oxford University Press.

Moulton, Brent, “Random Group Effects and the Precision of Regression Estimates,” *Journal of Econometrics*, 32, 3, August 1986, p. 385-397.

Paulson, Anna L., Robert M. Townsend and Alexander Karaivanov (2006), "Distinguishing Limited Commitment from Moral Hazard in Models of Growth with Inequality", Journal of Political Economy.

Salanié, Bernard (1997), *The economics of contracts: a primer*. Cambridge: MIT Press.

Stiglitz, Joseph E. (1974), "Incentives and Risk Sharing in Sharecropping", *Review of Economic Studies* 41:397-426.

Visaria, Sujata (forthcoming), “Legal Reform and Loan Repayment: The Microeconomic Impact of Debt Recovery Tribunals in India,” *American Economic Journal: Applied Economics*.

Table 1: Summary statistics

	Mean	Standard Deviation	10th Percentile	Median	90th Percentile	Observations
Baseline Characteristics						
Male	0.80	0.40	0	1	1	1147
Married	0.94	0.24	1	1	1	1147
Age	39.96	13.25	24	38	59	1147
Years of Education	5.35	3.50	0	5	10	1147
Risk Taker	0.56	0.50	0	1	1	1147
Days of Hunger Last Year	6.05	11.05	0	0	30	1147
Late Paying Previous Loan	0.13	0.33	0	0	1	1147
Income SD	27568.34	46296.41	3111.27	15556.35	57841.34	1147
Years of Experience Growing Paprika	2.22	2.36	0	2	5	1147
Previous Default	0.02	0.14	0	0	0	1147
No Previous Loans	0.74	0.44	0	1	1	1147
Take-up						
Approved	0.99	0.08	1	1	1	1147
Any Loan	1.00	0.00	1	1	1	1147
Total Borrowed (MK)	16912.60	3908.03	13782	16100	20136.07	1147
Land Use						
Fraction of Land used for Maize	0.43	0.16	0.28	0.40	0.63	520
Fraction of land used for Soya/Beans	0.15	0.16	0.00	0.11	0.38	520
Fraction of land used for Groundnuts	0.13	0.12	0.00	0.11	0.29	520
Fraction of land used for Tobacco	0.08	0.12	0.00	0.00	0.27	520
Fraction of land used for Paprika	0.19	0.13	0.00	0.18	0.36	520
Fraction of land used for Tomatoes	0.01	0.03	0.00	0.00	0.00	520
Fraction of land used for Leafy Vegetables	0.00	0.02	0.00	0.00	0.00	520
Fraction of land used for Cabbage	0.00	0.01	0.00	0.00	0.00	520
Fraction of Land used for all cash crops	0.57	0.16	0.38	0.60	0.72	520
Inputs						
Seeds (MK, Paprika)	247.06	348.47	0	0	560	520
Fertilizer (MK, Paprika)	7499.85	7730.05	0	5683	18200	520
Chemicals (MK, Paprika)	671.31	1613.13	0	0	2500	520
Man-days (MK, Paprika)	665.98	1732.99	0	0	2400	520
All Paid Inputs (MK, Paprika)	9084.19	8940.13	0	8000	19990	520
KG Manure, Paprika	90.84	313.71	0	0	250	520
Times Weeding, Paprika	1.94	1.18	0	2	3	520
Outputs						
KG Maize	1251.30	1024.36	360	1080	2160	520
KG Soya/Beans	83.14	136.86	0	40	200	520
KG Groundnuts	313.89	659.34	0	143	750	520
KG Tobacco	165.47	615.33	0	0	400	520
KG Paprika	188.14	396.82	0	100	364	520
KG Tomatoes	30.56	126.29	0	0	0	520
KG Leafy Vegetables	29.94	133.24	0	0	0	520
KG Cabbage	12.02	103.79	0	0	0	520
Revenue and Profits						
Market sales (MK)	65004.30	76718.29	9800	44000	137100	520
Profits (market sales + value of unsold crop - cost of inputs, MK)	117779.20	303100.80	33359	95135	261145	520
Value of Unsold Harvest (Regional Prices, MK)	80296.97	288102.70	24645	70300	180060	520
Repayment						
Balance, Sept. 30	2080.86	5663.98	0	0	9282	1147
Fraction Paid by Sept. 30	0.84	0.33	0	1	1	1147
Fully Paid by Sept. 30	0.74	0.44	0	1	1	1147

Table 2: **Tests of balance in baseline characteristics between treatment and control group**

Variable:	Full baseline sample: Mean in control group	Full baseline sample: Difference in treatment (fingerprinted) group	Loan recipient sample: Mean in control group	Loan recipient sample: Difference in treatment (fingerprinted) group
Male	0.81	-0.034 (0.022)	0.80	-0.066* (0.037)
Married	0.92	-0.005 (0.011)	0.94	0.003 (0.016)
Age	39.50	-0.092 (0.666)	39.96	-0.088 (1.171)
Years of education	5.27	-0.028 (0.172)	5.35	-0.124 (0.272)
Risk taker	0.57	-0.031 (0.031)	0.56	0.013 (0.051)
Days of hunger in previous season	6.41	-0.787 (0.818)	6.05	-0.292 (1.329)
Late paying previous loan	0.14	0.009 (0.023)	0.13	0.030 (0.032)
Standard deviation of past income	25110.62	1084.922 (1724.142)	27568.34	-1158.511 (2730.939)
Years of experience growing paprika	2.10	0.090 (0.139)	2.22	0.299 (0.223)
Previous default	0.03	0.001 (0.010)	0.02	0.008 (0.010)
No previous loan	0.74	-0.000 (0.026)	0.74	-0.020 (0.041)
Observations	3206		1147	

Stars indicate significance at 10% (*), 5% (**), and 1% (***) levels.
Notes: Each row presents mean of a variable in the baseline (September 2008) survey in the control group, and the difference between the treatment group mean and the control group mean of that variable (standard error in parentheses). Differences and standard errors calculated via a regression of the baseline variable on the treatment group indicator; standard errors are clustered at the club level.

Table 3: **Impact of fingerprinting on loan approval, loan take-up, and amount borrowed**

	(1)	(2)	(3)
Sample:	All Respondents	All Respondents	Loan Recipients
Dependent variable:	Approved	Any Loan	Total Borrowed (MK)
Panel A			
Fingerprint	0.038	0.051	-696.799*
	(0.053)	(0.044)	(381.963)
Panel B			
Fingerprint	0.207	0.108	-2812.766
	(.161)	(.145)	(2371.685)
Predicted repayment * fingerprint	-0.219	-0.074	2630.653
	(.197)	(.168)	(2555.167)
Panel C			
Fingerprint * Quintile 1	0.093	0.075	-2721.780
	(.115)	(.111)	(1666.068)
Fingerprint * Quintile 2	0.180*	0.102	-258.179
	(.096)	(.086)	(828.500)
Fingerprint * Quintile 3	-0.030	0.061	-458.924
	(.082)	(.073)	(596.109)
Fingerprint * Quintile 4	-0.001	-0.037	-101.028
	(.086)	(.082)	(575.968)
Fingerprint * Quintile 5	-0.017	0.039	-400.620
	(.100)	(.089)	(784.509)
Observations	3206	3206	1147
Mean of dependent variable	0.63	0.35	16912.60

Stars indicate significance at 10% (*), 5% (**), and 1% (***) levels.
Notes: Each column presents estimates from three separate regressions: main effect of fingerprinting in Panel A, linear interaction with predicted repayment in Panel B, and interactions with quintiles of predicted repayment in Panel C. All regressions include credit officer * week of initial loan offer fixed effects, baseline characteristics (male, five-year age categories, one-year education categories, and marriage), and baseline risk indicators (dummy for self-reported risk-taking, days of hunger in the previous season, late payments on previous loans, standard deviation of income, years of experience growing paprika, dummy for default on previous loan, and dummy for no previous loans). Standard errors on Panel A coefficients are clustered at the club level, while those in Panels B and C are bootstrapped with 200 replications and club-level resampling.

Table 4: Impact of fingerprinting on loan repayment

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	Loan recipients	Loan recipients	Loan recipients	Loan recipients included in August 2009 survey	Loan recipients included in August 2009 survey	Loan recipients included in August 2009 survey
Dependent variable:	Balance, Sept. 30	Fraction Paid by Sept. 30	Fully Paid by Sept. 30	Balance, Sept. 30	Fraction Paid by Sept. 30	Fully Paid by Sept. 30
Panel A						
Fingerprint	-996.430	0.073*	0.096	-875.314	0.073*	0.085
	(754.301)	(0.040)	(0.062)	(670.297)	(0.044)	(0.069)
Panel B						
Fingerprint	-9727.739**	0.716***	0.842***	-8931.946*	0.684***	0.759***
	(4199.085)	(.110)	(.178)	(5162.708)	(.196)	(.213)
Predicted repayment * fingerprint	10855.103**	-0.799***	-0.928***	10046.221*	-0.761***	-0.841***
	(4499.549)	(.121)	(.196)	(5446.717)	(.206)	(.240)
Panel C						
Fingerprint * Quintile 1	-7202.647**	0.499***	0.543***	-8016.543*	0.566***	0.599***
	(2969.045)	(.127)	(.147)	(4347.488)	(.195)	(.198)
Fingerprint * Quintile 2	-1028.696	0.066	0.163	1799.143	-0.098	-0.071
	(1871.298)	(.105)	(.160)	(1914.282)	(.111)	(.168)
Fingerprint * Quintile 3	-297.918	0.005	-0.004	-586.977	0.038	0.052
	(901.013)	(.048)	(.091)	(871.625)	(.055)	(.105)
Fingerprint * Quintile 4	775.231	-0.037	-0.045	549.532	-0.029	-0.065
	(883.076)	(.046)	(.078)	(821.086)	(.053)	(.113)
Fingerprint * Quintile 5	1404.812	-0.078*	-0.084	289.061	-0.006	0.007
	(951.535)	(.046)	(.074)	(804.733)	(.054)	(.110)
Observations	1147	1147	1147	520	520	520
Mean of dependent variable	2080.86	0.84	0.74	1439.16	0.89	0.79

Stars indicate significance at 10% (*), 5% (**), and 1% (***) levels.
Notes: Each column presents estimates from three separate regressions: main effect of fingerprinting in Panel A, linear interaction with predicted repayment in Panel B, and interactions with quintiles of predicted repayment in Panel C. All regressions include credit officer * week of initial loan offer fixed effects, baseline characteristics (male, five-year age categories, one-year education categories, and marriage), and baseline risk indicators (dummy for self-reported risk-taking, days of hunger in the previous season, late payments on previous loans, standard deviation of income, years of experience growing paprika, dummy for default on previous loan, and dummy for no previous loans). Standard errors on Panel A coefficients are clustered at the club level, while those in Panels B and C are bootstrapped with 200 replications and club-level resampling. Sample limited to individuals who took out loans in 2008 (columns 1-3), or who took out loans and were included in follow-up survey in 2009 (columns 4-6).

Table 5: **Impact of fingerprinting on land use**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: Fraction of land used for…	Maize	Soya/Beans	Groundnuts	Tobacco	Paprika	Tomatoes	Leafy Vegetables	Cabbage	All cash crops
Panel A									
Fingerprint	0.001	0.015	-0.012	-0.004	0.005	-0.001	-0.003	-0.001	-0.001
	(0.019)	(0.019)	(0.016)	(0.016)	(0.014)	(0.003)	(0.003)	(0.001)	(0.019)
Panel B									
Fingerprint	-0.009	-0.025	-0.025	-0.033	0.079	0.009	0.006	-0.003	0.009
	(.092)	(.094)	(.060)	(.062)	(.064)	(.010)	(.015)	(.004)	(.092)
Predicted repayment * fingerprint	0.013	0.049	0.016	0.036	-0.092	-0.013	-0.011	0.003	-0.013
	(.101)	(.105)	(.068)	(.066)	(.073)	(.013)	(.016)	(.005)	(.101)
Panel C									
Fingerprint * Quintile 1	-0.061	-0.013	-0.008	-0.012	0.083	0.005	0.007	-0.002	0.061
	(.066)	(.063)	(.052)	(.050)	(.051)	(.008)	(.012)	(.003)	(.066)
Fingerprint * Quintile 2	0.065	0.019	-0.014	-0.019	-0.035	-0.005	-0.010	-0.002	-0.065
	(.052)	(.042)	(.041)	(.030)	(.037)	(.008)	(.008)	(.002)	(.052)
Fingerprint * Quintile 3	-0.012	0.002	-0.009	0.004	0.009	0.008	-0.002	-0.001	0.012
	(.044)	(.045)	(.033)	(.022)	(.038)	(.008)	(.007)	(.002)	(.044)
Fingerprint * Quintile 4	0.008	0.015	-0.026	0.009	-0.003	-0.002	-0.003	0.002	-0.008
	(.041)	(.040)	(.034)	(.021)	(.037)	(.009)	(.007)	(.003)	(.041)
Fingerprint * Quintile 5	-0.005	0.043	-0.001	-0.001	-0.018	-0.012	-0.005	-0.002	0.005
	(.044)	(.040)	(.036)	(.023)	(.034)	(.009)	(.006)	(.003)	(.044)
Observations	520	520	520	520	520	520	520	520	520
Mean of dependent variable	0.43	0.15	0.13	0.08	0.19	0.01	0.00	0.00	0.57

Stars indicate significance at 10% (*), 5% (**), and 1% (***) levels.
Notes: Each column presents estimates from three separate regressions: main effect of fingerprinting in Panel A, linear interaction with predicted repayment in Panel B, and interactions with quintiles of predicted repayment in Panel C. All regressions include credit officer * week of initial loan offer fixed effects, baseline characteristics (male, five-year age categories, one-year education categories, and marriage), and baseline risk indicators (dummy for self-reported risk-taking, days of hunger in the previous season, late payments on previous loans, standard deviation of income, years of experience growing paprika, dummy for default on previous loan, and dummy for no previous loans). Standard errors on Panel A coefficients are clustered at the club level, while those in Panels B and C are bootstrapped with 200 replications and club-level resampling. Sample limited to individuals who took out loans in 2008 and who were included in follow-up survey in 2009.

Table 6: **Impact of fingerprinting on agricultural inputs used on paprika crop**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	Seeds (MK)	Fertilizer (MK)	Chemicals (MK)	Man-days (MK)	All Paid Inputs (MK)	KG Manure	Times Weeding
Panel A							
Fingerprint	74.107	733.419	345.328*	-395.501**	757.354	29.649	0.019
	(47.892)	(1211.905)	(190.262)	(181.958)	(1389.230)	(32.593)	(0.147)
Panel B							
Fingerprint	262.116*	11115.814**	466.677	411.043	12255.650**	52.882	0.182
	(146.417)	(5660.459)	(594.037)	(579.097)	(5987.210)	(144.033)	(.466)
Predicted repayment * fingerprint	-234.438	-12946.332**	-151.316	-1005.720	-14337.806**	-28.970	-0.203
	(183.931)	(6245.378)	(701.923)	(732.887)	(6700.416)	(161.334)	(.591)
Panel C							
Fingerprint * Quintile 1	188.703**	5871.126	374.260	106.406	6540.496	78.234	0.445
	(95.018)	(4062.716)	(406.741)	(347.367)	(4210.469)	(111.980)	(.367)
Fingerprint * Quintile 2	78.717	3597.540	244.449	-236.338	3684.368	27.058	-0.443
	(95.343)	(3026.725)	(414.863)	(454.498)	(3362.245)	(81.930)	(.338)
Fingerprint * Quintile 3	124.548	-585.618	500.669	-348.598	-309.000	58.670	-0.191
	(97.766)	(2250.453)	(427.366)	(458.033)	(2602.025)	(94.443)	(.333)
Fingerprint * Quintile 4	-10.190	-1790.213	283.962	-1065.690**	-2582.132	-25.080	-0.254
	(110.489)	(2503.022)	(430.040)	(537.142)	(2952.953)	(73.404)	(.348)
Fingerprint * Quintile 5	18.589	-2444.617	264.620	-315.018	-2476.427	21.879	0.564
	(110.367)	(2201.579)	(445.234)	(572.589)	(2635.638)	(93.481)	(.379)
Observations	520	520	520	520	520	520	520
Mean of dependent variable	247.06	7499.85	671.31	665.98	9084.19	90.84	1.94

Stars indicate significance at 10% (*), 5% (**), and 1% (***) levels.

Notes: Each column presents estimates from three separate regressions: main effect of fingerprinting in Panel A, linear interaction with predicted repayment in Panel B, and interactions with quintiles of predicted repayment in Panel C. All regressions include credit officer * week of initial loan offer fixed effects, baseline characteristics (male, five-year age categories, one-year education categories, and marriage), and baseline risk indicators (dummy for self-reported risk-taking, days of hunger in the previous season, late payments on previous loans, standard deviation of income, years of experience growing paprika, dummy for default on previous loan, and dummy for no previous loans). Standard errors on Panel A coefficients are clustered at the club level, while those in Panels B and C are bootstrapped with 200 replications and club-level resampling. Sample limited to individuals who took out loans in 2008 and who were included in follow-up survey in 2009.

<u>Table 7</u>: Impact of fingerprinting on revenue and profits

	(1)	(2)	(3)	(4)
<u>Dependent variable</u>:	Market sales (Self Report, MK)	Value of Unsold Harvest (Regional Prices, MK)	Profits (market sales + value of unsold harvest - cost of inputs, MK)	Ln(profits)
<u>Panel A</u>				
Fingerprint	7246.174	5270.320	14509.457	0.060
	(8792.055)	(14879.349)	(16679.311)	(0.095)
<u>Panel B</u>				
Fingerprint	69102.211	-29468.424	24207.068	0.651
	(49177.370)	(85252.270)	(90535.890)	(.423)
Predicted repayment * fingerprint	-77131.415	43317.493	-12092.441	-0.737
	(51232.390)	(103316)	(108112.600)	(.501)
<u>Panel C</u>				
Fingerprint * Quintile 1	30766.147	7940.835	31915.287	0.401
	(36850.940)	(50587.570)	(63206.880)	(.363)
Fingerprint * Quintile 2	41981.091	6364.782	45650.027	0.283
	(33084.250)	(75026.680)	(81848.520)	(.264)
Fingerprint * Quintile 3	-20925.441	-14911.454	-26932.651	-0.202
	(17938.730)	(59934.020)	(63400.760)	(.227)
Fingerprint * Quintile 4	-12785.841	7481.854	3609.228	-0.038
	(14733.930)	(57096.050)	(60385.110)	(.231)
Fingerprint * Quintile 5	1053.151	33336.147	34125.843	-0.054
	(15282.460)	(71891.840)	(74254.990)	(.240)
Observations	520	520	520	520
Mean of dependent variable	65004.30	80296.97	117779.16	11.44
Mean of dependent variable (US $)	464.32	573.55	841.28	n.a.

Stars indicate significance at 10% (*), 5% (**), and 1% (***) levels.
Notes: Each column presents estimates from three separate regressions: main effect of fingerprinting in Panel A, linear interaction with predicted repayment in Panel B, and interactions with quintiles of predicted repayment in Panel C. All regressions include credit officer * week of initial loan offer fixed effects, baseline characteristics (male, five-year age categories, one-year education categories, and marriage), and baseline risk indicators (dummy for self-reported risk-taking, days of hunger in the previous season, late payments on previous loans, standard deviation of income, years of experience growing paprika, dummy for default on previous loan, and dummy for no previous loans). Standard errors on Panel A coefficients are clustered at the club level, while those in Panels B and C are bootstrapped with 200 replications and club-level resampling. Sample limited to individuals who took out loans in 2008 and who were included in follow-up survey in 2009. Value of unsold harvest computed using regional prices.

Table 8: Ex post moral hazard

	(1)	(2)	(3)
Dependent variable:	Balance, Sept. 30	Fraction Paid by Sept. 30	Fully Paid by Sept. 30
Panel A			
Fingerprint	424.455	-0.005	0.000
	(565.064)	(0.039)	(0.072)
Panel B			
Fingerprint	-3537.222	0.320	0.399
	(5140.761)	(.237)	(.298)
Predicted repayment * fingerprint	4743.330	-0.389	-0.478
	(5378.012)	(.246)	(.326)
Panel C			
Fingerprint * Quintile 1	-4443.517	0.304	0.314
	(4252.255)	(.211)	(.248)
Fingerprint * Quintile 2	2679.579	-0.164	-0.142
	(1950.827)	(.118)	(.166)
Fingerprint * Quintile 3	-358.930	0.035	0.049
	(978.830)	(.062)	(.108)
Fingerprint * Quintile 4	763.678	-0.026	-0.058
	(909.161)	(.057)	(.113)
Fingerprint * Quintile 5	732.027	-0.019	-0.023
	(955.498)	(.062)	(.119)
Observations	520	520	520
Mean of dependent variable	1439.16	0.89	0.79

Stars indicate significance at 10% (*), 5% (**), and 1% (***) levels.
Notes: Each column presents estimates from three separate regressions: main effect of fingerprinting in Panel A, linear interaction with predicted repayment in Panel B, and interactions with quintiles of predicted repayment in Panel C. All regressions include credit officer * week of initial loan offer fixed effects, baseline characteristics (male, five-year age categories, one-year education categories, and marriage), and baseline risk indicators (dummy for self-reported risk-taking, days of hunger in the previous season, late payments on previous loans, standard deviation of income, years of experience growing paprika, dummy for default on previous loan, and dummy for no previous loans). Standard errors on Panel A coefficients are clustered at the club level, while those in Panels B and C are bootstrapped with 200 replications and club-level resampling. Sample limited to individuals who took out loans in 2008 and who were included in follow-up survey in 2009.

Appendix Table 1: Impact of fingerprinting on attrition from sample

Dependent variable: Indicator for attrition from September 2008 baseline survey to August 2009 survey

	(1)	(2)
Sample:	All respondents	Loan recipients
Panel A		
Fingerprint	-0.057	-0.086
	(0.036)	(0.070)
Panel B		
Fingerprint	-0.042	-0.134
	(.107)	(.197)
Predicted repayment * fingerprint	-0.020	0.059
	(.128)	(.225)
Panel C		
Fingerprint * Quintile 1	-0.023	-0.148
	(.075)	(.136)
Fingerprint * Quintile 2	-0.074	0.035
	(.071)	(.109)
Fingerprint * Quintile 3	-0.069	-0.106
	(.068)	(.105)
Fingerprint * Quintile 4	-0.086	-0.109
	(.076)	(.124)
Fingerprint * Quintile 5	-0.080	-0.115
	(.071)	(.128)
Observations	3206	1147
Mean of dependent variable	0.63	0.55

Stars indicate significance at 10% (*), 5% (**), and 1% (***) levels.
Notes: Each column presents estimates from three separate regressions: main effect of fingerprinting in Panel A, linear interaction with predicted repayment in Panel B, and interactions with quintiles of predicted repayment in Panel C. All regressions include credit officer * week of initial loan offer fixed effects, baseline characteristics (male, five-year age categories, one-year education categories, and marriage), and baseline risk indicators (dummy for self-reported risk-taking, days of hunger in the previous season, late payments on previous loans, standard deviation of income, years of experience growing paprika, dummy for default on previous loan, and dummy for no previous loans). Standard errors on Panel A coefficients are clustered at the club level, while those in Panels B and C are bootstrapped with 200 replications and club-level resampling.

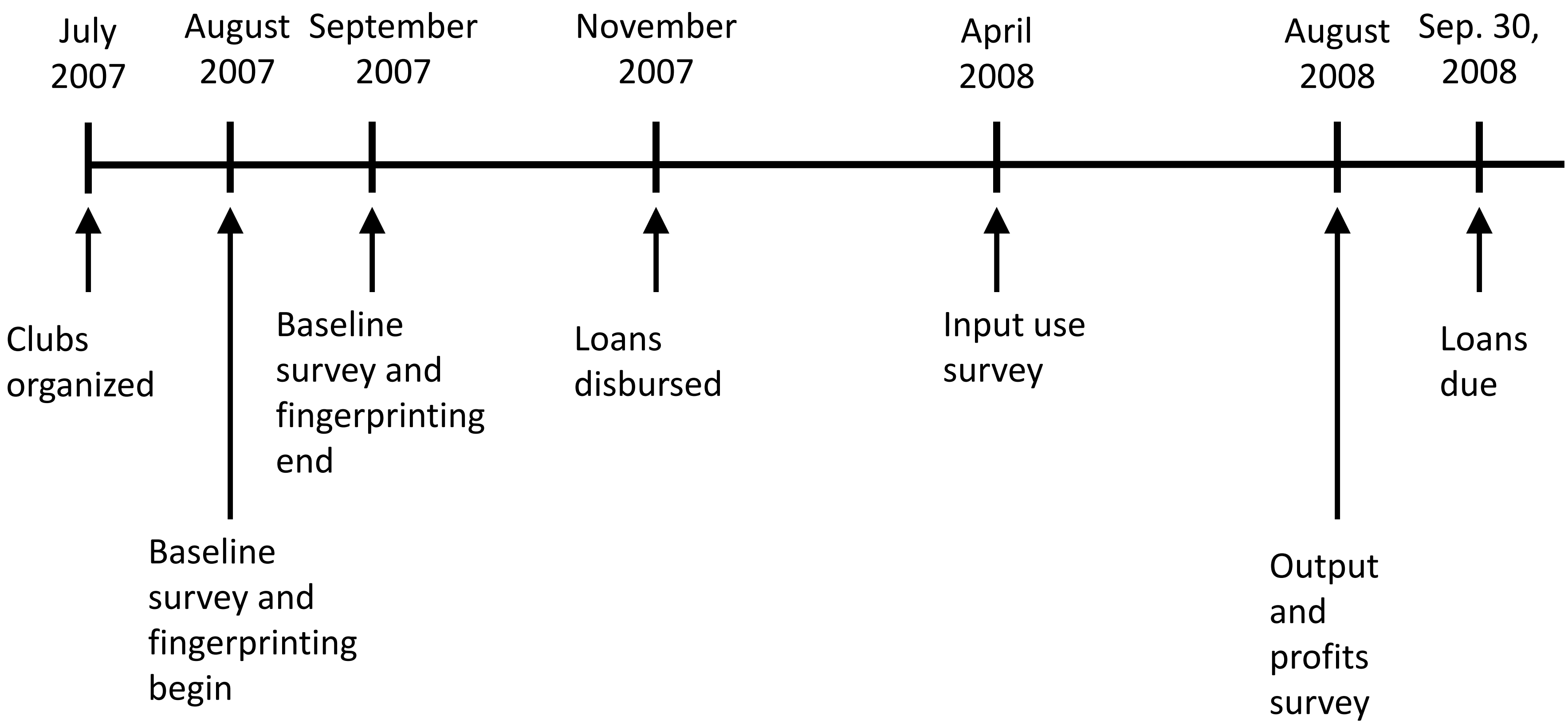
<u>Appendix Table 2</u>: Auxilary regression for predicting loan repayment

Dependent variable:	(1) Fraction Paid by Sept. 30	(2) Fraction Paid by Sept. 30	(3) Fraction Paid by Sept. 30
Male	0.080	0.061	0.058
	(0.073)	(0.048)	(0.048)
Married	-0.071	-0.091	-0.101
	(0.060)	(0.044)**	(0.046)**
Age	0.004	0.001	
	(0.001)***	(0.001)	
Years of education	-0.005	-0.003	
	(0.005)	(0.004)	
Risk taker	-0.078	0.008	0.013
	(0.041)*	(0.031)	(0.031)
Days of Hunger in previous season	0.001	-0.000	-0.001
	(0.002)	(0.001)	(0.001)
Late paying previous loan	-0.058	-0.084	-0.084
	(0.071)	(0.046)*	(0.047)*
Standard deviation of past income	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Years of experience growing paprika	0.005	0.007	0.007
	(0.013)	(0.011)	(0.011)
Previous default	0.088	0.128	0.097
	(0.163)	(0.079)	(0.078)
No previous loan	-0.012	0.015	0.013
	(0.062)	(0.032)	(0.034)
Constant	0.729	0.949	0.982
	(0.114)***	(0.072)***	(0.090)***
Loan officer * week of initial loan offer fixed effects	--	Y	Y
Dummy variables for 5-year age groups	--	--	Y
Dummy variables for each year of education	--	--	Y
Observations	563	563	563
R-squared	0.05	0.46	0.48

Robust standard errors in parentheses

Stars indicate significance at 10% (*), 5% (**), and 1% (***) levels.
Notes: Sample is non-fingerprinted loan recipients from the September 2008 baseline survey. All standard errors are clustered at the club level.

Figure 1: Experimental Timeline



<u>Figure 2</u>: Malawi Study Areas

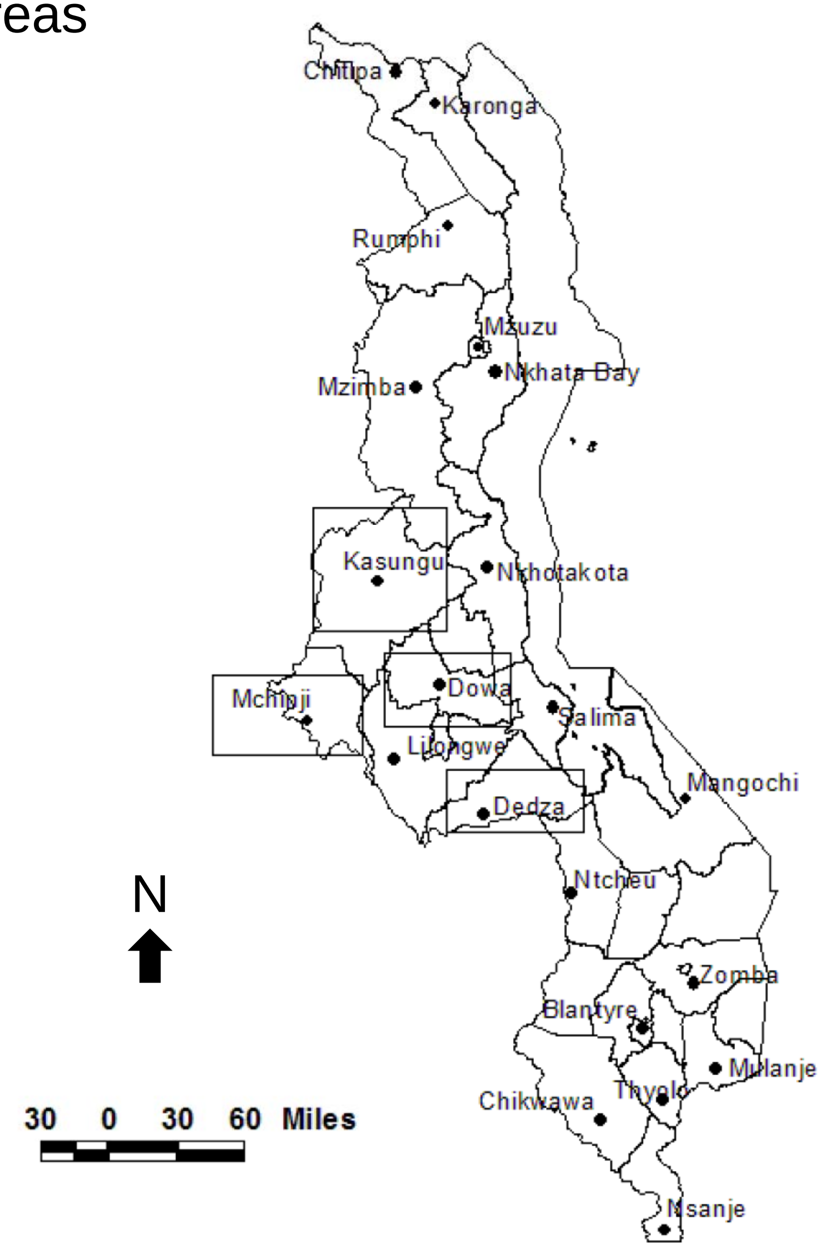
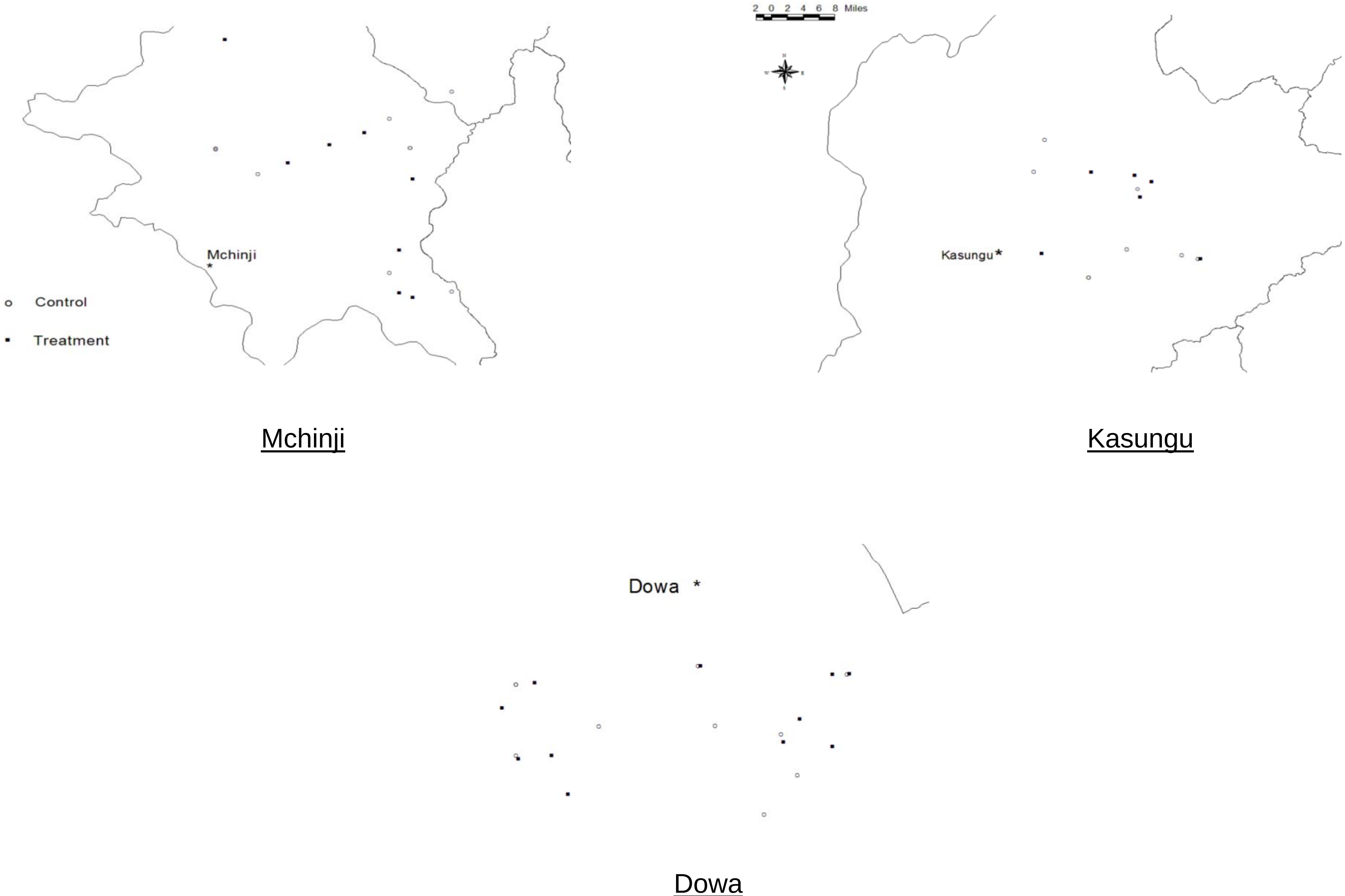


Figure 3: Spatial Distribution of Clubs



Note: GIS location data unavailable for all clubs in Dedza district, 15 clubs in Dowa district, and 6 clubs in Mchinji district.