Expanding Credit Access: Using Randomized Supply Decisions To Estimate the Impacts

Dean Karlan
Yale University
Innovations for Poverty Action
M.I.T. Jameel Poverty Action Lab

Jonathan Zinman
Dartmouth College

December 18th, 2006

ABSTRACT

Expanding credit access is a key ingredient of development strategies worldwide. Microfinance practitioners, policymakers, and donors have ambitious goals for expanding access, and seek efficient methods for implementing and evaluating expansion. There is less consensus on the role of *consumer* credit in expansion initiatives. Some microfinance institutions are moving beyond entrepreneurial credit and offering consumer loans. But many practitioners and policymakers are skeptical about "unproductive" lending. These concerns are fueled by academic work highlighting behavioral biases that may induce consumers to overborrow. We estimate the impacts of a consumer credit supply expansion using a field experiment and follow-up data collection. A South African lender relaxed its risk assessment criteria by encouraging its loan officers to approve randomly selected marginal rejected applications. We estimate the resulting impacts using new survey data on borrower behavior and well-being, and administrative data on loan repayment. We find that the marginal loans increased credit access and produced measurable benefits in the form of increased employment, reduced hunger, and reduced poverty. The marginal loans also appear to have been profitable for the lender. The results must be interpreted with caution but suggest that consumer credit expansions can be welfare-improving.

*dean.karlan@yale.edu, jzinman@dartmouth.edu. Thanks to Jonathan Bauchet, Luke Crowley, Nathanael Goldberg, and Ben Pugsley for excellent research assistance, to Lia Fernald for advice on measures of mental health, and to Sumit Agarwal, Abhijit Banerjee, and Chris Udry for helpful discussions. We are grateful to the National Science Foundation (SES-0424067 and CAREER SES-0547898), BASIS/USAID (CRSP), the Princeton University Center for Migration, the Social Science Research Council Program in Applied Economics, and the Federal Reserve Bank of New York for funding research expenses, and to the Lender for financing the loans. The views expressed in this paper are not necessarily shared by any of the funders, the Lender, or the Federal Reserve System.

Preface

One important theme in the work of the Center for Global Development is the search for ways to make foreign aid agencies more effective. It is a tough problem because aid agencies are not accountable to the people they aim to serve in aid-receiving countries. One symptom of this lack of accountability, noted by CGD's Evaluation Gap Working Group, is that donors too rarely commission rigorous, independent studies of how the programs they back affect clients. This leaves donors vulnerable to development fads and waste.

CGD non-resident fellow Dean Karlan and his co-authors are exemplars of a growing movement within academia to change that. This paper comes out of a program of work that strives to bring the highest scientific standards to the study of microfinance, an area in which public and private donors are heavily involved. Understanding how microfinance affects clients is not straightforward because there are several possible explanations for why, say, a borrower is doing well compared to her non-borrowing peers. The credit may be helping—or perhaps she only borrowed because she was already well-off. This, and other papers in the series, elucidates cause and effect by performing controlled experiments, in which a few parameters are randomly varied and the effects measured. The result is sharper answers, in specific contexts, to questions such as: How sensitive are potential borrowers to high interest rates? At the margin, does access to credit increase their incomes? Does it empower women? In the solidarity group lending method made famous by the Grameen Bank, wherein small groups of borrowers guarantee each other's loans, is that mutual guarantee the essential glue that holds the system together?

This paper contributes both by giving donors insight into the programs they fund, and, more generally, by demonstrating the value of rigorous impact evaluation.

I. Introduction

Expanding access to credit is a key ingredient of development strategies worldwide. The microfinance industry has grown exponentially over the past twenty years under the premise that expanding access to credit will help improve the welfare of the poor (Morduch 1999; Armendariz de Aghion and Morduch 2005). This policy push has been driven by both theoretical and empirical motivations. Theoretical models show that information asymmetries can lead to credit market failures and ensuing poverty traps (Banerjee and Newman 1993). Empirical evidence shows strong negative correlations between depth of access and poverty rates at the macro level (Levine 1997; Honohan 2004), and positive impacts of access to microfinance at the micro level (Pitt and Khandker 1998). Policymakers, practitioners, and funders are committed to continued rapid growth.

There is less consensus on the role of *consumer* credit in expansion initiatives. Some microfinance institutions are moving beyond "traditional" entrepreneurial credit and offering consumer loans. But many practitioners remain skeptical about "unproductive" lending (Robinson 2001). Policy is similarly conflicted, both within and across countries, and over time.[1] Concerns about the development of consumer credit markets are fueled by academic work highlighting behavioral biases that may induce consumers to overborrow.[2]

There is also uncertainty about *how* to expand credit access. Marginal borrowers may require relatively small loan amounts, and thus traditional approaches to microcredit expansion—creating new microfinance institutions, adding branches, designing new joint liability mechanisms— may not be the most cost-effective method to support efficient expansion.

[1] South Africa offers an example of such conflicted policy approaches. South Africa deregulated usury ceilings in 1992 to encourage the development of formal markets in consumer credit. However, recent legislation re-imposed some ceilings, effective in 2007. Another example is the substantial variation across U.S. states in payday lending restrictions (Hanson and Morgan 2005).

[2] For example: Laibson, Repetto, and Tobacman (2005) find that consumers with present-biased preferences would commit $2,000 to not borrow on credit cards; Ausubel (1991) argues that over-optimism produces excess credit card borrowing; Stango and Zinman (2006a; 2006b) find that consumers systematically underestimate the interest rate on non-mortgage installment loans and borrow heavily and expensively as a result.

Another way to expand access to credit is for existing lenders to liberalize their screening criteria.[3]

We assess the impacts of liberalizing credit screening criteria by analyzing new data produced by a field experiment and follow-up survey work. The key questions are threefold. First, do credit constraints actually bind? Second, does relaxing any credit constraints actually benefit marginal borrowers? Revealed preference logic says it should: a consumer borrows only if she will benefit (in expectation). Behavioral models say not necessarily: biases in preferences and cognition may lead consumers to overborrow. The third key question is how much a lender profits or loses from making marginal loans.

The experiment was implemented by a consumer lender in a high-rate, high-risk South African installment loan market where credit constraints appear to bind. First-time applicants are often rejected, even at prevailing real rates of 200% APR. Default rates average about 20% among new borrowers. A prior experiment on experienced borrowers from the same lender found far greater sensitivity to maturity than price (Karlan and Zinman 2006a); as Attanasio, Goldberg, and Kryiazidou (2004) show formally, this pattern of elasticities is further evidence of unmet demand for credit.

Measuring the causal impacts of credit expansion on borrower and lender outcomes is usually complicated by deep identification issues. Two types of endogeneity are particularly problematic: the self-selection of clients into loan contracts, and targeted interventions by lenders and policymakers. These problems make it difficult to draw firm conclusions from non-experimental studies without strong assumptions. A classic example concerns relatively "spunky" individuals selecting or being selected into microcredit borrowing, and thereby confounding any causal effect of access to credit with the causal effects of individual characteristics (including those that may change unobservably over time). Selection can work in

[3] Liberalization of screening criteria is used in directed lending programs (Banerjee and Duflo 2004), semi-directed lending programs (e.g., the Community Reinvestment Act in the United States), and by many microlenders that expand "outreach" while holding their physical capital and risk assessment technology constant.

the opposite direction as well; e.g., if households (lenders) tend to take (target) microcredit in anticipation of needing to smooth upcoming *negative* shocks. Attempts to overcome these problems using quasi-experimental, structural, and control function approaches have yielded mixed results.[4]

We addressed the identification problem by working with a lender to engineer exogenous variation in the loan approval process. Our treatment randomly encouraged loan officers to approve some marginal applications. Specifically, the Lender marketed to and screened new loan applicants using its normal procedures and three additional steps. First, loan officers were required to label rejected applications as either egregiously uncreditworthy or marginally uncreditworthy. Second, the loan officer's computer then instructed the loan officer to reconsider some marginal applications in real-time by randomly producing a message to "approve" or "still reject." Loan officers were instructed by management to follow the computer's instructions in all cases, and compliance was monitored. But in the third and final step, loan officers had pecuniary incentives to be risk-averse and complied in only 53% of the cases when the computer instructed them to approve. Consequently our design identifies treatment-on-the-treated effects of expanding access on the following policy- and strategy-relevant sample: those applicants deemed by loan officers to be closest to the margin of creditworthiness. Neither the treatment (computer said "approve") nor the control (computer said "reject") groups were informed by the Lender that a component of the loan decision was randomized.

We then obtained outcome data from the Lender's records on repayment and profitability, and from household surveys of marginal applicants assigned to the treatment and control groups. An independent research firm conducted surveys at the home or workplace of

[4] See, e.g., Coleman (1999), Kaboski and Townsend (2005), McKernan (2002), Pitt, Khandker, Chowdury, and Millimet (2003), and Pitt and Khandker (1998). These studies focus on micro*entrepreneurial* credit rather than consumer credit. However there may be little economic distinction between small, closely-held businesses and the households that run them, and there is some evidence the microentrepreneurial loans are often used for consumption smoothing (Morduch 1998; Menon 2003).

these marginal applicants six to twelve months after they applied for the loan. The survey measures borrowing activity, loan uses, and a range of proxies for household well-being.

We estimate the *average* impacts of expanding credit access by comparing outcomes across those *assigned* to the treatment and control groups. Our results corroborate the presence of binding liquidity constraints. Control applicants did not simply obtain credit elsewhere; conversely, treated applicants borrowed more overall in the 6-12 months following the experiment, and changed their lender type composition.

We find some evidence that relaxing credit constraints produced statistically significant, tangible benefits by enabling marginal borrowers to make productive investments and smooth consumption. Treatment group applicants were an estimated 11 percentage points more likely than controls to retain wage employment, 6 percentage points less likely to experience severe hunger in their households, and 7 percentage points less likely to be impoverished.[5] Given the level and nature of treatment compliance, one can double these estimates to get a measure of the impact of expanded access on the treated sample: those marginal applicants who were deemed by loan officers to be closest to creditworthy. We find no evidence of significant negative effects on resources or well-being.

We also find that the marginal loans were profitable for the Lender. The average loan earned an estimated $12 (7% of the principal amount). The suggestion that the Lender was "leaving money on the table" is surprising given its long track record of profitability. Explanations include learning about new technologies, and market power that dulls incentives for efficiency. Alternatively, we may not be adjusting adequately for risk, or for loan officer agency problems that make conservative risk assessments optimal. We can not rule out these alternative explanations. As such our finding that the marginal loan is profitable is merely suggestive.

[5] Our employment retention effect is analogous to the "sticking it out" effect whereby access to credit enables small firms to smooth shocks and stay in business (Holtz-Eakin, Joulfaian and Rosen 1994). In our case it appears that access to credit enables consumers to smooth shocks and/or make productive investments in health, uniforms, and transport in order to retain employment. Our hunger reduction effect fits with evidence in Gertler, Levine, and Moretti (2003) that microcredit helps Indonesian families smooth consumption against health shocks. More generally Gertler and Gruber (2002) find very imperfect consumption insurance against illness in Indonesia.

Our results suggest a role for welfare-improving interventions in consumer credit markets but come with other important caveats. The diffuse set of borrower outcomes that could be affected by credit access makes inference challenging: we estimate treatment effects on 10 different sets of outcomes and find significant effects on 4 of them in the full sample.[6] Some of these outcome measures have more variance than others, and as such some results have large standard errors. We follow the practice of clinical trials and report all of our results, including the statistically insignificant ones (Begg 1985; Gerber, Green and Nickerson 2001). We do have sufficient power to detect significant effects on what are arguably the two most important outcomes: credit access and poverty..

Our time horizon for measuring impacts is between six months and one year. This has advantages and disadvantages. The key advantage is that measuring borrower loan uses close in time to the expanding access treatment enables us to "follow the money" with less noise, and thereby learn more about the mechanisms through which marginal loans generate their impacts. The key disadvantage of course is that some effects from relaxing credit constraints may only materialize over the longer-run.

Despite these limitations, our results and methodology offer some novel insights into the motivation, design, and evaluation of credit market interventions. We demonstrate that randomized-controlled trials can be used to help identify the severity of liquidity constraints, and to evaluate efforts to expand credit access. Experiments also can be used to measure whether borrower behavior and outcomes are consistent with models of revealed preference and/or behavioral alternatives. Our results seem more consistent with the former: borrowers in our sample appear to know what is good for them, at least over a 6-12 month horizon. Most practically, our results suggest that liberalizing screening criteria can benefit both borrowers and

[6] The 10 different sets of outcomes are credit access, income, consumption, employment, events, education, housing, well-being, household decision-making, and shocks.

lenders, and our methodology demonstrates how lenders can hone in on their sustainability/outreach frontier by taking controlled risks using randomized experimentation.

II. Market and Lender Overview

Our cooperating Lender operated for over 20 years as one of the largest, most profitable micro-lenders in South Africa.[7] It competed in a "cash loan" or "microloan" industry segment that offers small, high-interest, short-term, uncollateralized credit with fixed repayment schedules to a "working poor" population. Aggregate outstanding loans in the microloan market equal approximately 38% of non-mortgage consumer credit (Department of Trade and Industry South Africa 2003).

Cash loan borrowers generally lack the credit history and/or collateralizable wealth needed to borrow from traditional institutional sources such as commercial banks. Cash loan sizes tend to be small relative to the fixed costs of underwriting and monitoring them, but substantial relative to borrower income. For example, the median loan size made under this experiment ($127) was 40% of the median borrower's gross monthly income.[8] Our sample for this experiment includes mostly first-time loan applicants of African descent. Table 1 shows some comparative demographics. Table 7 and Section IV detail that borrowers finance a variety of different consumption smoothing and investment activities.

Cash lenders arose to substitute for traditional "informal sector" moneylenders following deregulation of the usury ceiling in 1992, and they are regulated by the Micro Finance Regulatory Council. Cash lenders focusing on the observably highest-risk market segment typically make one-month maturity loans at 30% interest *per month*. Informal sector moneylenders charge 30-

[7] The Lender was merged into a large bank holding company in 2005 and no longer exists as a distinct entity.
[8] Throughout the paper we convert all South Africa currency into US dollars using the average exchange rate over our study period of September 21, 2004-November 30, 2005: 6.31 Rand= $1.

100% per month. Lenders targeting observably lower risk segments charge as little as 3% per month.[9]

The cash loan market has important differences and similarities with "traditional" microcredit (e.g., the Grameen Bank, other NGOs, and government lending programs). In contrast to our setting, most microcredit has been delivered by lenders with explicit social welfare and targeting goals. Microlenders typically target female entrepreneurs and often use group liability mechanisms. On the other hand, the industrial organization of microcredit is trending steadily in the direction of the for-profit, more competitive delivery of individual, untargeted credit that characterizes the cash loan market (Robinson 2001; Porteous 2003). This push is happening both from the bottom-up (non-profits converting to for-profits) as well as from the top-down (for-profits expanding into microcredit segments).

Our cooperating Lender's product offerings were somewhat differentiated from competitors. Unlike many cash lenders, it did not pursue collection or collateralization strategies such as direct debit from paychecks, or physically keeping bank books and ATM cards of clients. Its pricing was transparent and linear, with no surcharges, application fees, or insurance premiums added to the cost of the loan. The Lender also had a "medium-maturity" product niche in 4-month loans. Most other cash lenders focus on 1-month or 12+-month loans.[10] In this experiment 98% of the borrowers received the standard loan for first-time borrowers: a 4-month maturity at 11.75% per month, charged on the original balance (200% APR).

Per standard practice in the cash loan market, the Lender conducted underwriting and transactions in its branch network. Its risk assessment technology combined centralized credit scoring with decentralized discretion. The credit scoring model screened out severely unqualified applicants and produced a recommendation on whether to approve the application. Branch personnel made the final decision. The Lender rejected fifty percent of new applications for

[9] South Africa has had very low inflation rates in recent years; e.g., 4.35% over our 14-month study period.
[10] The Lender also had 1, 6, 12, and 18 month products, with the longer maturities offered at lower rates and restricted to the most observably creditworthy customers.

reasons including unconfirmed employment, suspicion of fraud, poor credit rating, and excessive debt burden.

Applicants who were approved often defaulted on their loan obligation (see Section V), despite facing several incentives to repay. Carrots included decreasing prices and increasing future loan sizes following good repayment behavior. Sticks included reporting to credit bureaus, frequent phone calls from collection agents, court summons, and wage garnishments.

III. Methodology

Our research design first randomly encourages loan officers to approve marginal rejected applicants, and then uses data from the lender and household surveys to measure impacts on profitability, credit access, investment, and well-being. The household data are collected by a survey firm with no ties to (or awareness of) the Lender.

A. Experimental Design and Implementation

Sample and time frame for the experiment

We drew our sample frame from the universe of 3,187 "new" applicants who had no prior borrowing from the Lender and applied at any of 8 branches between September 21 and November 20, 2004. The branches were located in the Capetown, Port Elizabeth, and Durban areas. The Lender maintained normal marketing procedures by advertising on billboards, park benches, the radio, and newspapers.

Our sample frame was comprised of "marginal" applicants: new, rejected, but potentially creditworthy. Specifically, applicants were eligible for the loan randomization if they were rejected under the Lender's normal underwriting criteria but not deemed egregiously uncreditworthy by a loan officer. 787 applicants met these criteria.

The motivation for experimenting with credit supply increases on a pool of marginal applicants is twofold. First, it focuses on those who should be targeted by initiatives to expand access to credit. Second, it provides the Lender with information about the expected profitability of changing its underwriting in a way that induces branch personnel to approve more risky loans.

Experimental Design and Operations

The Lender implemented the experiment in four steps:

First, loan officers evaluated each of the 3,187 new applicants using the Lender's normal underwriting process and three additional steps. Under normal operations the loan officer would use a combination of a credit scoring model and her own discretion to make a binary approve/reject decision. The experiment forced loan officers to take the first additional step of dividing the "reject" category into two bins. "Marginal" rejects would be eligible for treatment; "egregious" rejects would not be assigned a loan under any circumstances. Egregious rejects were identified subjectively, based on extremely poor credit history, overindebtedness, suspected fraud, lack of contactability, or legal problems. Loan officers approved 1,695 of the 3,187 (53%) new applications processed by participating branches during our study period. 705 (22%) applications were deemed egregious rejects, leaving us with a sample frame of 787 (25%) marginally rejected applicants for the experiment.

Second, special "randomizer" software randomly assigned a loan to some of the 787 marginal applicants. Loan officers inputted basic information (name, credit history, maximum feasible loan size if approved, and reason for rejection) on each of the 787+705 = 1,492 rejected applications into the randomizer. The randomizer then used the inputted information to treat (i.e., approve) applications with probabilities that were conditional on the credit score and loan officer assessment. The 705 egregious applications had zero probability of being treated. The 787 marginal applicants were divided into two groups based on their credit score. Those with better credit scores were treated with probability 0.50, and those with worse credit scores were treated with probability 0.25 (all analysis controls for this condition of the randomization). Table 2, Column 1 corroborates that randomizer treatment assignments generated observably similar treatment and control groups. In total, 325 applicants were assigned to the treatment group, leaving 462 in the control group.

Last, the branch manager made the final credit decision and announced it to the applicant. The applicant was not privy to the loan officer's initial decision, the existence of the software, or the introduction of a randomized step in the decision-making process. Accepted applicants were offered an interest rate, loan size, and maturity per the Lender's normal underwriting criteria. Recall that nearly all received the standard contract for first-time borrowers: a 4-month maturity at 200% APR.

The branch manager did not have to adhere to the random assignment. For applicants assigned to the treatment (approve) group, 172, or 53%, received a loan. For applicants assigned to the control (reject) group, 7, or 1%, received a loan. Accordingly we conduct our analysis on an "intent-to-treat" basis, since we do not know which control group applicants would have passed the branch manager's final subjective approval step. Hence we compare those *assigned* to treatment to those *assigned* to control, regardless of whether the branch adhered to the random assignment (please see III-C below for more details, and discussion of treatment-on-the-treated effects).

Loan repayment was monitored and enforced according to normal operations. Branch manager compensation was based in part on loan performance, and the experiment did not change the incentive pay of any field personnel.

B. Household Data Collection

Following the experiment we hired a firm to survey applicants in the treatment and control groups. The purpose of the survey was to measure behavior and outcomes that might be affected by access to credit. As detailed in Section V, the surveyors asked questions on demographics, resources, recent investments, employment status, and proxies for well-being.[11]

[11] The survey took an average of 1.5 hours to complete.

The sample frame for the household survey included the entire pool of 787 marginal applicants from the experiment. Surveyors completed 626 surveys, for an 80% response rate. In order to avoid potential response bias between the treatment and control groups, the survey firm and respondents were not informed about the experiment or any association with the Lender. We told the survey firm that the target households' contact information came from a "consumer database in South Africa." Surveyors were trained to conduct a generic household survey, with emphasis on family finances, and the respondent consent form reflected this.

Each survey was conducted within six to twelve months of the date that the applicant entered the experiment by applying for a loan and being placed in the marginal group. Our rationale for this timing is twofold. First, it avoids a mechanical timing bias in favor of finding positive impacts on credit access, by allowing sufficient time for the control group applicants to find credit elsewhere. Second, it (partially) allows for the fact that certain investments have a gestation period before they manifest in outcomes. In short, we have chosen to evaluate "medium-run" rather than immediate impacts. Of course analysis over a longer-run would be interesting as well.

C. Experimental Validity and Empirical Strategy

Our methodology has two experimental validity or interpretation issues. One relates to the possibility of attrition bias. Another relates to imperfect adherence to the random assignment. We describe and address these two issues in turn.

The first experimental validity issue is whether our follow-up survey sampling strategy produces attrition bias. As noted above, our methodology requires obtaining survey data on both treatment and control households. Our experimental variation is sufficient to identify unbiased estimates of the impact of getting a loan on survey outcomes only if treatment assignment is uncorrelated with the probability of completing a survey. Table 3 corroborates that this condition holds: treatment status is uncorrelated with the survey response rate. Table 3 also shows that treatment assignment is uncorrelated with demographics measured in the survey.

The second issue arises due to the non-compliance: cases where the administered treatment deviates from the assigned one. Note here that some non-compliance in the econometric sense (non-treatment in the treatment group) makes good business strategy sense. "Non-compliance" whereby branch personnel cherry-picks the best prospects from a pool of randomly selected marginal applicants is indeed an experiment of interest for lenders where branch personnel have the final say on credit decisions. This setup tests the profitability of "encouraging" branch personnel to make more marginal loans, while still allowing them final subjective decision-making power. Mechanisms for encouragement include incentives, training, and monitoring.

We anticipated substantial non-compliance and sought to maximize econometric power by obtaining the highest feasible approval level (through training and monitoring loan personnel). Table 4 shows the relationship between treatment assignment and administration. 53% of applicants approved by the randomizer actually obtained loans (172 out of 325), and 98% of applicants rejected by the randomizer actually did not get a loan (455 out of 462).

The imperfect treatment in the treatment group motivates an intention-to-treat (ITT) estimator. ITT produces an unbiased estimate of *average* treatment effects even when there is substantial non-treatment. We implement ITT using the following OLS specification:

$$(1) \quad Y_i = \alpha + \beta \mathit{treatment}_i + \delta \mathit{risk}_i + \phi \mathit{appmonth}_i + \gamma \mathit{surveymonth}_i + \varepsilon_i$$

Y is a behavior or outcome of interest for applicant i (or i's household). Examples of Y include measures of borrowing (see Table 5), poverty status (see Table 8), and loan repayment (see Table 10). $Treatment_i$ = 1 if the individual was assigned to treatment (irrespective of compliance). $Risk_i$ captures the applicant's credit score; this determined whether the applicant was treated with probability 0.25 or 0.5. $Appmonth_i$ is the month in which the applicant entered the experiment (September, October, or November 2004), and $surveymonth_i$ is the month in which the survey was completed. These month variables control for the possibility that the lag between application and survey is correlated with both treatment status and outcomes.[12]

The average treatment effect is captured by β. As noted above, using the random assignment (ITT), rather than whether the borrower actually obtained a loan, avoids any bias from noncompliance with the assignment to treatment and control. We also estimate heterogeneous treatment effects by splitting the sample on characteristics of interest. The gender of the borrower is interesting because many microfinance organizations target women, and women are often believed to have differential access to both formal and informal financial services. Household income is interesting because there is often tension in microfinance between "sustainability" (profitability) and "outreach" (expanding credit supply) to the "poorer of the poor" (Morduch 1999; 2000). Little is known about where impacts are strongest. Treatment effects may be stronger on the relatively poor if they are relatively credit constrained. Alternatively, treatment effects may be weaker on the relatively poor if they lack complementary skills or resources. Similarly, we also split the sample by *ex-ante* credit risk.

D. Treatment-on-the-Treated Estimates and Interpretation

Treatment-on-the-treated (TOT) effects can be estimated by doubling the ITT estimates, since the difference in treatment rates between treatment and control groups is 0.5. However any TOT results must be interpreted with care. Heterogeneity in treatment effects (as is likely if manager compliance varied with unobserved applicant characteristics that are correlated with outcomes) imply that the TOT results can *not* be generalized to all individuals who were below the normal creditworthiness threshold. Rather they estimate the impact of credit expansion on *applicants deemed creditworthy enough by branch personnel* to merit compliance with the

[12] This could occur if control applicants were harder to locate (e.g., because we could not provide updated contact information to the survey firm), and had poor outcomes compared to the treatment group (e.g., because they did not obtain credit).

randomization. Effects might differ for expansions implemented through other processes such as centralized credit scoring.

E. Inference in This Sample

Several issues make inference challenging in this implementation of our methodology.

First, the impacts of consumer credit are potentially broad and diffuse. As Section V details, we define "investments" broadly: there are several types of activities that could be financed and then generate benefits for treated households. But how do we measure such benefits? There are few if any generally accepted summary statistics for utility. Consequently we estimate treatment effects on 9 different sets of proxies for household resources and overall well-being.[13]

Second, the 6-12 month time horizon used in our study does not capture some long-run impacts of interest. Poverty traps or debt traps may only become evident over longer horizons, and some investments may have a longer gestation period. We chose a medium-run horizon in order to strike a balance between gestation periods, allowing the control group time to find other credit, and accurately recording the uses of borrowed funds.

Third is external validity. As with most empirical work, our findings are directly applicable to our sample only. Of course our sample is a subset of larger populations of interest: principally, those with physical access to microfinance who are being screened out by current criteria in the industry. We discuss external validity in greater detail in the Conclusion.

IV. Results: Impacts on Borrowing and Credit Access

This section reports treatment effects of the Lender's supply expansion on marginal applicants' overall access to credit. Additional lending by the Lender is unlikely to affect borrowers unless

[13] This stands in contrast to the entrepreneurial credit setting, where the set of investments is somewhat circumscribed by the nature of the business, and there are natural summary statistics for business performance: sales, profitability, and survival.

credit constraints bind. If rejected applicants can simply obtain a loan from a different lender (at similar terms), then we will not find a treatment effect on borrowing, and hence would not expect to find treatment effects on investment or ultimate outcomes.

Table 5 reports the treatment effects on borrowing outcomes. We find no significant effect on the extensive margin of overall borrowing: treated households were not more likely to have obtained a loan in the 6-12 months after applying to the Lender (top panel, "all sources"). But treated households did respond on the intensive margin of overall borrowing: the bottom panel shows a significantly higher quantity of loans from all sources (the total number of loans per person rises by 0.141, or 28%). Both panels also show a change in the *type* of credit accessed. Treated households were more likely to report borrowing from a microlender (the Lender falls into that classification) and less likely to report borrowing from other formal sources (banks, NGOs and retailers). The normative implications of this result are not clear in isolation. We lack good data on loan costs for the individual loans, and rates charged by other formal lenders can vary widely both within and across different source types.[14] But together with data on investments and ultimate outcomes (Section V) we can examine whether the changes in borrowing opportunities produced by the treatment actually benefited households.[15]

Table 5 also shows limited evidence of heterogeneous treatment effects. We find several instances where the treatment effect is significant in one sub-sample but not another. However the differences across males and females, income groups, and credit score bins are not statistically significant.

Table 6 presents treatment effects on what we label "perception of credit access." Specifically, the survey asked: "If you needed a loan tomorrow, where would you go to borrow?"

---

[14] The survey did not ask the respondent to identify the specific lender. Surveyors did ask for the interest rate on each loan, but response rates were very low.

[15] Another limitation of our data is that it almost certainly and dramatically understates the prevalence of informal borrowing (compare to South African Financial Diaries data at www.financialdiaries.com). We believe that most informal loans were not reported due to poor wording and logic in our survey. If, as commonly believed, microloan borrowing serves as a (less expensive) substitute for informal borrowing in South Africa, then this implies that our data: 1) overstates the positive impacts on overall borrowing, and 2) misses a negative impact on informal borrowing.

Treated households were 15.7 percentage points (45%) more likely to report "Microlender or Cash lender" than the control group. Treated households were 11.8 percentage points (23%) less likely to report an informal source (friends, family, moneylender, or borrowing circle). These results are consistent with expanded access to formal credit changing the marginal source of borrowing from informal to formal.

The last row of results in Table 6 addresses whether the change in marginal source is due (partly) to formal access crowding-out informal access.[16] Specifically, the survey asked: "In an emergency could you or your spouse/partner get financial assistance from any friends or relatives?" The point estimate suggests that the treatment did reduce access to informal markets by 5.6 percentage points (7.5%), although the result is insignificant.

Table 6 also shows some heterogeneity in treatment effects on perception of credit access. The results suggest that female, poor, and risky applicants are all relatively more likely to make cash loans their marginal source of credit as a result of the treatment. Relatively wealthier and more creditworthy applicants are more likely to *lose* access to informal credit markets as a result of the treatment. Again, the standard errors are large and do not rule out homogenous treatment effects.

V. Results: Loan Uses, and Ultimate Impacts

Table 7 shows the range of activities financed by household borrowing. These loan uses motivate estimating treatment effects on a particular set of expenditures, activities, and economic outcomes. We then also estimate the treatment effects on a series of summary proxies for well-being that measure stress, depression, optimism, general health, decision-making power, and the incidence of shocks.

[16] This is an old but understudied issue. See Bell (1990) for a discussion and investigation.

A. Loan Uses, and Impacts on Specific Borrower Outcomes

The most common purpose for household borrowing is paying off other debt. This suggests that marginal microloans may be used to economize on interest expenses, and to maintain access to other credit sources by permitting timely repayment. These and other reported uses suggest estimating treatment effects on consumption.

Table 8 shows that we find no effect on average monthly consumption post-treatment; results for other (unreported) measures of total consumption are qualitatively similar (positive signs, imprecisely estimated). The finding that 23.2% of microloans are used to buy or improve food motivates a particular focus on food consumption. Here we find an effect: households randomly assigned a loan were 5.8 percentage points less likely to experience hunger during the past 30 days. This is a large effect on the small base of households (14%) that reported any hunger. The other measures of food consumption also appear to respond positively to credit access, although the estimates are much less precise (perhaps because these intensive measures of food consumption are noisier than binary hunger).

The next most common purpose for household borrowing is transportation expenses (19.5%); this and the clothing category are consistent with work-related investments. Indeed we find large treatment effects on employment: treated applicants were 11 percentage points (13%) more likely to be working at the time of the survey. Since everyone in our sample frame had verified employment at the time they entered the experiment, it appears that the treatment effect operates by enabling households to *maintain* employment by smoothing or avoiding shocks that prevent them from getting to work.

The employment effect, and microfinance's focus on poverty reduction, motivates estimating treatment effects on income as well. We find insignificant effects on the level and percentile of income, although again the point estimates are positive. We find a marginally significant and large decrease in poverty headcount of 7.1 percentage points (17%). Appendix Table 1 explores this further with a simple means comparison, broken out by treatment probability. The results suggest a very large reduction in poverty for the low credit score (25% treatment probability) group, although again the result is only significant at the 90% level.

We also estimate the treatment effect on self-employment. Reported prevalence of using loan proceeds to finance business activity is low (3.2%), but may be underreported (since some consumer lenders actively discourage entrepreneurial activity), or subsumed in other categories. We estimate an increase of 2 percentage points (13%) that is insignificant in the full sample. However, low-income treated applicants were significantly more likely (at the 90% level) to be self-employed. The estimated 9 percentage point increase may seem implausibly large, given the mean self-employment rate of 15.7% among low-income households; however, microentrepreneurial credit is very scarce in South Africa, and the returns to microenterprises may be very high for the relatively poor and credit constrained (McKenzie and Woodruff 2006).

Many households report borrowing for events. The nature of these events—holidays, initiations, funerals, weddings— makes it unsurprising that the extensive margin (the probability of occurrence) is not affected by access to credit (not reported). Table 8 shows that the treatment effect on intensive margin of events spending is insignificant and small.

13.7% of loans are used for educational expenses.[17] Households report almost perfect attendance among compulsory school-aged children, so we focus on the intensive margin of school expenditure. The treatment effect is small and insignificant. The confidence interval on university attendance contains large effects, but the estimate is imprecise.

A final frequent use of loan proceeds is for housing expenses (11.5%). We find a slightly negative treatment effect on home purchase or improvements, but this estimate is very imprecise. The treatment effect on housing expenditure, conditional on making a purchase or improvement, is negative but imprecisely estimated.

[17] Educational expenses may be predictable, but other expenses and income may not; i.e., (treated) households may use credit to smooth educational investment in the aftermath of shocks.

B. Impacts on Summary Measures of Borrower Well-Being

Table 9 reports treatment effects on several summary measures of well-being. In principle we would like to measure utility; in practice we lack a summary statistic for household well-being. The first three rows show that we find no significant treatment effects on standard measures of stress, depression, or optimism. Each outcome is measured on a linear scale,[18] and in each case the confidence intervals suggest that the upper bound on the treatment effect is a 15% change. The fourth row reports the effect on self-reported health. Respondents could choose from one of five categories ranging from very bad (1) to very good (5). The treatment effect is insignificant and bounded above at a small improvement in health.[19]

The second panel of Table 9 reports the treatment effects on decision-making power. Many microfinance initiatives seek to increase the intrahousehold bargaining power of female borrowers. Recent work in the Philippines finds that a commitment *savings* product generated more decision-making power for married females (Ashraf, Karlan and Yin 2006), which in turn led to more purchases of female-oriented durable goods for the household. Here we find point estimates that are consistent with positive effects on borrowers of both genders, although the treatment effect on females is imprecisely estimated.

The bottom panel of Table 9 shows no effect on avoiding shocks on margins other than employment. The table shows that such shocks are prevalent; e.g., 1-0.483 = 51.7% of households had experienced one since entering the experiment. Credit may be used to avoid shocks by making productive investments (e.g., in health), or may result in more exposure to shocks due to induced risk-taking. The treatment effects are insignificant but the standard errors do not rule out substantial changes in either direction.

[18] See Cohen and Willamson (1988), Radloff (1977), and Scheier, Carver, and Bridges (1994) for details on constructing and validating the stress, depression, and optimism scales.

[19] The table reports OLS results. An ordered probit produces qualitatively similar results.

C. Impacts on the Lender

Table 10 estimates positive net profits for the Lender from making the marginal loans assigned by the treatment. The average profit per marginal loan was R74.28 ($11.75), which was 7% of the amount lent. These estimates err on low side: we rule out any interest revenue or principal recovery on loans in default at the time our data feeds ended (May 2005), and also estimate profits on the marginal client's first loan only. Forecasts prepared in consultation with the Lender suggests that the net present value of the marginal client was about $25 (not reported in the table).[20]

The finding that the Lender was "leaving money on the table" is surprising given its long track record returns-on-equity ranging between 30% and 80%. We see at least four potential explanations.

First, to the extent that lending to marginal applicants requires new technology, one could view our experiment as part of a transition to a steady-state of profitable lending to marginal applicants. In other words, our project systematized (and perhaps hastened) a less formal process of learning and innovation.

Second, market power combined with agency problems between managers and shareholders may dull incentives for marginal improvements in efficiency. As discussed in Section II, the Lender seemed to possess a unique market niche, and Table 5 corroborates that its marginal clients faced liquidity constraints.

Third, risk-weighted profits may be negative. Indeed, Table 10 estimates that "inframarginal" loans (to first-time borrowers initially approved under the Lender's normal underwriting criteria during the experimental operations) were nearly twice as profitable. Discussions with the Lender's management indicate that the profit earned by marginal loans was

---

[20] Successive loans are common (though not always taken out immediately upon repay the previous one) and more profitable than the first loan because: 1) default rates fall: the first loan seems to "weed out" unobservably risky types; 2) loan amounts increase over time; 3) maturities increase over time. These patterns are evident both in the Lender's historical data, and in the available data on follow-on borrowing by the borrowers in our experiment.

worth the risk. But we can not see whether this assessment translated into an actual long-run change in strategy, since the Lender was merged into a bank holding company in May 2005, and as far as we know the new management team was not informed of our research.

Fourth, loan officer agency problems may hinder efforts to reach the productive efficiency frontier. The first-best employee incentive contract may be infeasible (due, e.g., to the potential for fraud), and the second-best contract may provide incentives that appear to promote conservatism in risk assessment.[21]

Table 10 provides a bit of additional evidence on operational challenges facing efforts to expand credit supply using existing technology. The results by credit score suggest that the Lender's normal risk assessment process did a poor job of assessing the *relative* profitability of marginal loans as well as the absolute profitability. The most profitable marginal loans were those with the *lowest* ex-ante credit scores; this stands in contrast to inframarginal loans, where the most profitable loans were those with the highest scores. Those with worse credit scores may value their access to this Lender more if they have fewer outside options. Thus the ex-ante worst risks, with more to lose from defaulting, may exert more effort or willingness to repay.

In all the treatment effects on Lender profitability suggest that microlenders should evaluate their productive efficiency. The Lender had a long track record of profitable operations, yet does not seem to have been operating at its profitability frontier in terms of either the quantity or quality of loans.

VI. Discussion: Implications for Theory and Welfare Analysis

The results must be interpreted with caution but have some implications for theory and welfare analysis.

On the theory side, our experiment provides a low-powered test of competing models of consumer intertemporal choice. The results provide some support for the neoclassical prediction

[21] As discussed in Section III, our experiment did not change any employee incentive schemes.

that consumers make themselves (weakly) better off by borrowing when credit constraints are relaxed: our treatment group had better employment stability, experienced less hunger, and was less likely to be below the poverty line. We find no statistically significant evidence that consumers make themselves worse off by borrowing, as some behavioral models would predict. However, our standard errors are large and do not rule out economically meaningful negative impacts on some outcomes.

On the welfare side, our results suggest that the net effect is unambiguously positive, since both borrowers and the Lender appear to benefit from marginal loans. Estimating the magnitude of the welfare gain would require additional assumptions. We surmise that the gains would be substantial under most assumptions, since the treatment effects on borrower outcomes are economically large, and treatment effects on Lender profits are nontrivial.

Again, we should keep in mind that we can only observe costs and benefits over the 6-12 month horizon following the treatment. As discussed in Section V-C, this almost certainly leads to underestimation of Lender profits. Borrower benefits may also be mismeasured if some investments have gestation periods that are longer than our 6-12 month window. On a related note, our results find no significant evidence that marginal borrowing leads to debt traps. The treatment group borrowed more intensively over the full 6-12 months following the experiment, but did not have significantly more debt at the time the survey was conducted (Table 5). This seems to be the case even though the treatment group *could* borrow from microlenders at the time of the survey if desired (Table 5). Consequently the results seem more consistent with a model where the marginal borrower repays her loan from cash flows, rather than refinancing, or defaulting and losing market access.

VII. Conclusion

Measuring the causal impacts of access to credit is critical for evaluating theory and practice, but complicated by basic identification issues. We address the identification problem by engineering exogenous variation in the approval of consumer loans. A lender encouraged loan officers to approve randomly selected marginal rejected applications for market-rate, four-month term loans.[22] Loan officers reconsidered in real-time, and unbeknownst to the applicants. Half of the reconsidered applicants were approved. We then tracked the behavior and outcomes of the treatment (reconsidered) and control (still rejected) groups over the next six to twelve months using administrative data and detailed household surveys.

Our results corroborate the presence of binding liquidity constraints and suggest that expanding credit supply improves welfare. There are three key sets of findings. First, control applicants who were randomly denied by our cooperating lender did not simply obtain credit elsewhere; conversely, treated applicants increased their total borrowing, and changed their lender type composition, in the 6-12 months following the experiment. Second, we find some evidence that treated applicants reaped tangible benefits from being able to make productive investments and smooth consumption: they were an estimated 11 percentage points more likely to retain wage employment, 6 percentage points less likely to experience severe hunger in their households, and 7 percentage points less likely to be below the poverty line. We find no statistically significant negative effects on resources or well-being. In all the results provide little support for behavioral models where biased consumers borrow too much. Third, we find that the marginal loans were profitable for the Lender, with the caveat that true profitability is difficult to measure given the (potential) need for adjustments due to learning, risk, and agency problems.

It is not clear whether these results will extrapolate to other settings. We experimented in a particular setting that is not necessarily representative of other markets, populations, or interventions.[23] But our findings are provocative because practitioners and behavioral theorists view our setting as one where the deck was stacked against finding beneficial impacts. Our

[22] The Lender conducted the experiment on a pool of initially denied applicants and hence did not deny anyone who would have qualified for a loan under standard underwriting criteria. See Section III for details.

[23] Our setting does have parallels to the U.S. payday loan market, and hence could inform related policy debates summarized in, e.g., Flannery and Samolyk (2005). Congress recently imposed a binding interest rate ceiling on loans made to members of the military that was motivated by payday and other "predatory" consumer lending practices; see, e.g., http://www.responsiblelending.org/issues/payday/briefs/page.jsp?itemID=29862357.

Lender was for-profit, the intervention was blunt, the credit was expensive, the market was somewhat competitive, and we targeted consumers rather than entrepreneurs. Yet we find some evidence of benefits and little evidence that consumers harmed themselves by borrowing at 200% APR.

Our main point of generality is methodological. A field experiment followed by a follow-up survey can be used to identify any motivation for, and impacts of, credit market interventions. This approach should build on related work that identifies the presence or absence of specific market failures (Karlan and Zinman 2006b) and how targeted populations make decisions (Bertrand, Karlan, Mullainathan, Shafir and Zinman 2005; Karlan and Zinman 2006a). Taken together this layered approach can identify markets that are ripe for welfare-improving interventions, design mechanisms that are most likely improve efficiency, and then evaluate whether the mechanisms actually work. The layered approach is costly but worth it. Donors, governments, and firms allocate billions of dollars to credit market interventions each year. Even if one takes a pessimistic view of external validity and proceeds market-by-market, a tiny fraction of the resources devoted to large microcredit markets would fund the experiments and surveys needed to generate specific and scientific guidance for practitioners and policymakers.

References

Armendariz de Aghion, B. and J. Morduch (2005). The Economics of Microfinance. Cambridge, MA, MIT Press.

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

Attanasio, O. P., P. K. Goldberg and E. Kryiazidou (2004). Credit Constraints in the Market for Consumer Durables: Evidence from Micro Data on Car Loans. Working Paper.

Ausubel, L. M. (1991). "The Failure of Competition in the Credit Card Market." American Economic Review 81(1): 50-81.

Banerjee, A. and E. Duflo (2004). "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program." M.I.T. Working Paper.

Banerjee, A. and A. Newman (1993). "Occupational Choice and the Process of Development." Journal of Political Economy 101: 274-298.

Begg, C. B. (1985). Publication Bias. The Handbook of Research Synthesis. H. Cooper and L. V. Hedges. New York, Russell Sage Foundation.

Bell, C. (1990). "Interactions between Institutional and Informal Credit Agencies in Rural India." World Bank Economic Review 4: 297-327.

Bertrand, M., D. Karlan, S. Mullainathan, E. Shafir and J. Zinman (2005). What's Psychology Worth? A Field Experiment in the Consumer Credit Market. Working Paper.

Cohen, S. and G. Williamson (1988). Perceived Stress in a Probability Sample of the United States. The Social Psychology of Health. S. Pacapan and S. Oskamp. Newbury Park, CA, Sage.

Coleman, B. (1999). "The Impact of Group Lending in Northeast Thailand." Journal of Development Economics 45: 105-41.

Department of Trade and Industry South Africa (2003). "Credit Law Review: Summary of findings of the Technical Committee."

Flannery, M. and K. Samolyk (2005). "Payday Lending: Do the Costs Justify the Price." Working Paper.

Gerber, A., D. Green and D. Nickerson (2001). "Testing for Publication Bias in Political Science." Political Analysis 9(4): 385-392.

Gertler, P. and J. Gruber (2002). "Insuring Consumption Against Illness." The American Economic Review 92(1): 51-70.

Gertler, P., D. Levine and E. Moretti (2003). "Do Microfinance Programs Help Insure Consumption Against Illness?" Working Paper.

Hanson, S. and D. Morgan (2005). "Predatory Lending?" Working Paper.

Holtz-Eakin, D., D. Joulfaian and H. Rosen (1994). "Sticking it Out: Entrepreneurial Survival and Liquidity Constraints." Journal of Political Economy 102(1): 53-75.

Honohan, P. (2004). "Financial Development, Growth and Poverty: How Close are the Links?" World Bank Policy Research Working Paper 3203.

Kaboski, J. and R. Townsend (2005). "Policies and Impact: An Analysis of Village-Level Microfinance Institutions." Journal of the European Economic Association 3(1): 1-50.

Karlan, D. and J. Zinman (2006a). "Credit Elasticities in Less Developed Economies: Implications for Microfinance." Working Paper.
Karlan, D. and J. Zinman (2006b). "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment." Working Paper.
Laibson, D., A. Repetto and J. Tobacman (2005). Estimating Discount Functions with Consumption Choices Over the Lifecycle. Working Paper.
Levine, R. (1997). "Financial Development and Economic Growth: Views and Agenda." Journal of Economic Literature XXXV: 688-726.
McKenzie, D. and C. Woodruff (2006). "Do entry costs provide an empirical basis for poverty traps? Evidence from Mexican microenterprises." Economic Development and Cultural Change forthcoming.
McKernan, S.-M. (2002). "The Impact of Microcredit Programs on Self-Employment Profits: Do Noncredit Program Aspects Matter?" Review of Economics and Statistics 84(1): 93-115.
Menon, N. (2003). "Consumption Smoothing in Micro-Credit Programs." Working Paper.
Morduch, J. (1998). Does Microfinance Really Help the Poor? New Evidence on Flagship Programs in Bangladesh., MacArthur Foundation project on inequality. Princeton University, draft.
Morduch, J. (1999). "The Microfinance Promise." Journal of Economic Literature 36: 1569-1614.
Morduch, J. (2000). "The Microfinance Schism." World Development 28(4): 617-629.
Pitt, M. and S. Khandker (1998). "The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter?" Journal of Political Economy 106(5): 958-96.
Pitt, M., S. Khandker, O. H. Chowdhury and D. Millimet (2003). "Credit Programs for the Poor and the Health Status of Children in Rural Bangladesh." International Economic Review 44(1): 87-118.
Porteous, D. (2003). "Is Cinderella Finally Coming to the Ball: SA Microfinance in broad perspective." Micro Finance Regulatory Council working paper.
Radloff, L. S. (1977). "The CES-D scale: A self-report depression scale for research in the general population." Applied Psychological Measurement 1: 385-401.
Robinson, M. (2001). The Microfinance Revolution: Sustainable Finance for the Poor. Washington, DC, IBRD/The World Bank.
Scheier, M. F., C. S. Carver and M. W. Bridges (1994). "Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A re-evaluation of the Life Orientation Test." Journal of Personality and Social Psychology 67(1063-1078).
Stango, V. and J. Zinman (2006a). "Fuzzy Math and Red Ink: Payment/Interest Bias, Intertemporal Choice, and Wealth Accumulation." Working Paper.
Stango, V. and J. Zinman (2006b). "How a Cognitive Bias Shapes Competition: Evidence from Consumer Credit Markets." Working Paper.

Table 1. Comparative demographics

		Experiment sample		Applicants with a 25% chance of approval		Applicants with a 50% chance of approval		Average salary in the formal sector, 2004	South Africa	Blacks in South Africa
		Mean	Median	Mean	Median	Mean	Median			
Head of household employed		68.2%	-	75.0%	-	66.3%	-		73.8% (a)	68.9% (a)
Female head of household		37.7%	-	31.8%	-	39.4%	-			
Years of education of head of household		9.8	11	9.7	11	9.8	11			
Age of head of household		44.4	42	41.0	39	45.3	43			
Number of kids in household		1.9	2	1.6	1	2.0	2			
Number of household members		5.4	5	4.8	4	5.6	5		3.8 (d)	3.9 (d)
Any member of household is self-employed		16.7%	-	13.3%	-	17.7%	-		15.7% (e)	17.7% (e)
Race of loan applicant										
	African	65.0%	-	70.6%	-	63.4%	-		79.3% (f)	-
	White	4.8%	-	4.4%	-	5.0%	-		9.5% (f)	-
	Indian	4.7%	-	5.0%	-	4.6%	-		2.4% (f)	-
	Colored	25.3%	-	20.0%	-	26.9%	-		8.8% (f)	-
Monthly household income		R 4,117	R 1,945	R 3,160	R 1,600	R 4,389	R 2,100	R 6,882 (b)	R 3,750 (c)	R 2,167 (c)

The experiment sample varies from 578 to 626 depending on missing values in the survey.

Race varies a lot by province in South Africa; e.g., our sample includes relatively high proportion of mixed race "Colored" individuals because Capetown branches participated in the experiment.

Average exchange rate during project and survey: 1 US$ = 6.3 Rands.

<u>Notes on monthly household income:</u>

Respondents were asked separately about:

- permanent employment salary and bonuses,
- casual employment salary and bonuses,
- income from self-employment,
- many different grants and pensions (unemployment, old age, disability, child rearing, etc.),
- rent and remittances received,
- agriculture income, and
- any other type of income.

<u>Lettered notes:</u>

(a) Employment rate of the active population. Source: Labour force survey, September 2004.

(b) Average earnings for non-agriculture formal employees, November 2004. Source: Quarterly Employment Statistics, Statistics South Africa, November 2005.

(c) In Rands of 2000. Inflation for the period 2000-November 2004: 25%.

(d) Average household size. Census 2001.

(e) Calculated from the Labour Force Survey, September 2004.

(f) South African population. Source: Mid-year population estimates, South Africa 2004, Statistics South Africa.

Table 2. Orthogonality of Treatment to Applicant Character

Dependent Variable:	*1= Loan Assigned*	*1 = Loan Obtained*
	(1)	(2)
Female	0.031	0.049
	(0.036)	(0.030)
African	0.073	0.183**
	(0.086)	(0.074)
Colored	0.047	0.196*
	(0.092)	(0.107)
Indian	0.130	0.218
	(0.121)	(0.145)
Age of applicant	-0.002	-0.002
	(0.002)	(0.002)
Monthly gross income at application (in '000)	0.010	0.019***
	(0.008)	(0.006)
# months at employer	0.000	0.000
	(0.000)	(0.000)
Observations	785	783

Huber-White standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%
Sample contains 785 of the 787 marginal applicants eligible for the treatment (i.e., for loan approval). Each column reports marginal effects for a single probit of the dependent variable listed in the column heading on a set of covariates comprised of: 1) the right-hand-side variables listed in the row headings; 2) the credit score categories that determined the treatment assignment probability (these are not shown). 'White' is the omitted race category. Two observations are dropped due to missing race.

Table 3. Experiment and Survey validity

	Full sample	
	ITT=1	ITT=0
Female head of household	41.3%	35.1%
Number of years of education of head of household	9.7	9.8
Average age of head of household	44.8	44.1
Average number of children (<18) in the household	1.9	1.9
Average number of household members	5.4	5.4
Survey response rate	79.7%	79.4%

Sample size varies from 578 to 626 depending on missing values in survey.
Sample size for the survey response rate: 787 (includes 161 applicants not found for the survey).

Table 4. Compliance with Treatment Assignment

Randomizer Says	Lender Actually	Frequency
Reject	Rejects	455
Approve	Approves	172
Reject	Approves	7
Approve	Reject	153

Table 5. Treatment Effects on Borrowing

		Mean depvar for full sample	Full sample	Gender		Income		Credit score	
				Female	Male	High	Low	High	Low
Dummy 'got a loan'									
Since date of application	All sources	0.352	0.041	0.023	0.078	0.009	0.079	0.030	0.064
			(0.040)	(0.056)	(0.059)	(0.056)	(0.059)	(0.060)	(0.056)
	Microlender	0.184	0.125***	0.121***	0.129***	0.127***	0.131**	0.155***	0.107**
			(0.034)	(0.046)	(0.050)	(0.046)	(0.052)	(0.050)	(0.046)
	Other formal sources	0.172	-0.055*	-0.098**	0.010	-0.077*	-0.040	-0.106**	-0.015
			(0.032)	(0.044)	(0.045)	(0.047)	(0.040)	(0.046)	(0.044)
	Informal sources	0.032	0.011	0.027	-0.001	-0.002	0.030	0.016	0.014
			(0.015)	(0.020)	(0.024)	(0.018)	(0.026)	(0.023)	(0.021)
At time of survey	All sources	0.333	0.027	0.028	0.059	-0.034	0.067	0.015	0.050
			(0.040)	(0.057)	(0.057)	(0.056)	(0.055)	(0.059)	(0.055)
	Microlender	0.150	0.118***	0.129***	0.119***	0.094**	0.142***	0.122***	0.128***
			(0.031)	(0.044)	(0.045)	(0.044)	(0.045)	(0.045)	(0.044)
	Other formal sources	0.198	-0.047	-0.083*	0.008	-0.088*	-0.026	-0.090*	-0.007
			(0.033)	(0.048)	(0.047)	(0.050)	(0.042)	(0.050)	(0.046)
	Informal sources	0.015	-0.001	0.005	-0.004	0.000	-0.000	0.013	-0.013*
			(0.009)	(0.015)	(0.013)	(0.010)	(0.016)	(0.019)	(0.008)
Sample size		626	626	311	315	314	312	283	343
Number of observations (range)		618-624	614-624	305-310	309-315	307-313	307-311	279-283	335-341
Number of loans									
Since date of application	All sources	0.506	0.141**	0.141	0.178*	0.086	0.225**	0.160	0.130
			(0.069)	(0.096)	(0.101)	(0.088)	(0.109)	(0.101)	(0.096)
	Microlender	0.230	0.211***	0.216***	0.202***	0.185***	0.254***	0.263***	0.173***
			(0.051)	(0.074)	(0.072)	(0.062)	(0.086)	(0.080)	(0.067)
	Other formal sources	0.210	-0.069*	-0.101*	-0.004	-0.081	-0.065	-0.127**	-0.026
			(0.041)	(0.057)	(0.058)	(0.057)	(0.057)	(0.056)	(0.060)
	Informal sources	0.053	0.010	0.039	-0.016	-0.003	0.039	0.028	-0.000
			(0.025)	(0.026)	(0.045)	(0.018)	(0.043)	(0.029)	(0.039)
At time of survey	All sources	0.421	0.077	0.042	0.156*	0.014	0.114	0.059	0.113
			(0.057)	(0.077)	(0.086)	(0.084)	(0.075)	(0.085)	(0.079)
	Microlender	0.166	0.133***	0.129**	0.149***	0.114**	0.148***	0.148***	0.137***
			(0.036)	(0.051)	(0.055)	(0.056)	(0.046)	(0.054)	(0.048)
	Other formal sources	0.229	-0.057	-0.104**	0.018	-0.101*	-0.039	-0.119**	0.005
			(0.041)	(0.053)	(0.061)	(0.060)	(0.052)	(0.057)	(0.059)
	Informal sources	0.018	0.001	0.014	-0.009	0.000	0.004	0.022	-0.018
			(0.012)	(0.021)	(0.017)	(0.011)	(0.022)	(0.025)	(0.011)
Sample size		626	626	311	315	314	312	283	343
Number of observations (range)		609-621	609-621	303-309	306-312	304-311	305-310	278-282	331-339

Huber-White standard errors in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%

All results obtained using OLS to estimate the ITT model detailed in equation (1); each cell presents the estimated treatment effect from a single regression. Running probits for the binary outcomes produces qualitatively similar results.

The number of observations varies depending on missing values in the survey data.

The income cutoff point is the median income measured at application.

The credit score represents the quality of the application, along two dimensions: (1) the credit bureau score, and (2) an internal score computed by the Lender. The credit score cutoff point separates applicants in the two lowest categories from applicants in the three higher categories.

Table 6. Treatment Effects on Perception of Credit Access

	Mean depvar for full sample	Full sample	Gender		Income		Credit score	
			Female	Male	High	Low	High	Low
Respondent would borrow from microlender if needed a loan	0.348	0.157***	0.213***	0.081	0.089	0.252***	0.115	0.189***
		(0.048)	(0.064)	(0.072)	(0.066)	(0.072)	(0.073)	(0.066)
Respondent would borrow from other formal sources (excluding microlenders) if needed a loan	0.685	0.006	-0.034	0.041	0.046	-0.064	0.017	-0.009
		(0.044)	(0.062)	(0.062)	(0.054)	(0.069)	(0.067)	(0.060)
Respondent would borrow from informal sources if needed a loan	0.504	-0.118**	-0.149**	-0.052	-0.132*	-0.091	-0.161**	-0.063
		(0.049)	(0.068)	(0.074)	(0.067)	(0.071)	(0.075)	(0.066)
Respondent would be able to borrow from friends or family if needed	0.746	-0.056	-0.047	-0.065	-0.153***	0.069	-0.110*	-0.003
		(0.040)	(0.055)	(0.059)	(0.056)	(0.058)	(0.060)	(0.054)
Number of observations (range)	434-530	434-530	216-272	218-258	218-265	216-265	187-244	247-286

Huber-White standard errors in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%

All results obtained using OLS to estimate the ITT model detailed in equation (1); each cell presents the estimated treatment effect from a single regression. Running probits produces qualitatively similar results.

The number of observations varies depending on missing values in the survey data.

The income cutoff point is the median income measured at application.

The credit score represents the quality of the application, along two dimensions: (1) the credit bureau score, and (2) an internal score computed by the Lender. The credit score cutoff point separates applicants in the two lowest categories from applicants in the three higher categories.

Table 7. Loan Uses

	All loans since application	Microlender loans since application	Other formal loans since application	Informal loans since application
Pay other debts	28.3%	31.7%	27.7%	15.2%
Transportation	19.4%	12.7%	9.2%	24.2%
Events	16.9%	15.5%	17.7%	21.2%
School/university	13.7%	15.5%	12.3%	9.1%
Improve/build house	11.5%	6.3%	18.5%	6.1%
Buy/improve food	9.9%	23.2%	6.9%	0.0%
Bills	7.3%	7.0%	8.5%	6.1%
Durable goods	6.7%	4.2%	10.8%	0.0%
Health care	5.1%	5.6%	3.8%	24.2%
Other personal uses	4.5%	3.5%	6.9%	6.1%
Buy clothes	3.5%	4.9%	3.1%	0.0%
Business uses	3.2%	2.8%	4.6%	0.0%
Total	129.9%	133.1%	130.0%	112.2%
Number of observations (i.e. number of loans)	314	142	130	33

The columns sum to more than 100% because respondents could state more than one use of the loan proceeds.

The number of observations for all loans (314) is not equal to the sum of the number of observations of the sub-samples due to 9 missing values in the variable 'loan source'.

'Transportation' includes buying/repairing a car, and public transport.

'Events' include cultural and religious ceremonies (Christmas, funeral, young men initiation, etc.), and holidays and parties.

'Other personal uses' include helping families and friends, and miscellaneous expenses.

Table 8. Treatment Effects on Consumption, Income, and Expenditure

	Mean depvar for full sample	Full sample	Gender		Income		Credit score	
			Female	Male	High	Low	High	Low
Consumption								
Log(average consumption since application)	7.878	0.023	-0.081	0.074	0.014	-0.045	-0.096	0.098
		(0.083)	(0.117)	(0.124)	(0.110)	(0.106)	(0.120)	(0.118)
Number of observations (range)	626	626	311	315	314	312	283	343
Food consumption								
Change in food quality over last 12 months	2.905	0.085	-0.156	0.278*	0.128	0.052	0.022	0.164
		(0.098)	(0.139)	(0.142)	(0.140)	(0.139)	(0.141)	(0.138)
Dummy=1 if anybody in household went to bed hungry in last 30 days	0.139	-0.058**	-0.016	-0.085**	-0.044	-0.058	-0.006	-0.094**
		(0.027)	(0.039)	(0.038)	(0.034)	(0.044)	(0.039)	(0.038)
Number of observations (range)	604-626	604-626	297-311	307-315	309-314	295-312	275-283	329-343
Employment								
Dummy=1 if the respondent is employed	0.802	0.108***	0.107**	0.096**	0.108***	0.085	0.090*	0.104**
		(0.032)	(0.047)	(0.045)	(0.036)	(0.056)	(0.049)	(0.044)
Dummy=1 if anybody in household is self-employed	0.167	0.022	-0.015	0.051	-0.057	0.090*	-0.008	0.046
		(0.033)	(0.043)	(0.049)	(0.045)	(0.050)	(0.048)	(0.047)
Number of observations (range)	605-606	605-606	300-307	299-305	306	299-300	273-275	330-333
Income and Poverty								
Post-treatment income percentile	50.000	2.571	0.392	4.652	2.658	0.708	0.883	2.969
		(2.450)	(3.328)	(3.753)	(3.424)	(3.120)	(3.623)	(3.388)
Dummy=1 if household total income below poverty line	0.416	-0.071*	-0.090	-0.050	-0.054	-0.065	-0.055	-0.073
		(0.041)	(0.058)	(0.058)	(0.051)	(0.062)	(0.061)	(0.057)
Log(1+average income since application)	6.965	0.182	-0.184	0.593*	0.426	-0.105	-0.169	0.379
		(0.212)	(0.287)	(0.340)	(0.299)	(0.301)	(0.302)	(0.299)
Number of observations (range)	620-622	620-622	306-309	313-314	310-314	308-310	279-283	339-341
Events								
Log(expenditures on events since application), conditional on having experienced such events	7.480	0.128	0.020	0.158	0.253	-0.059	0.191	0.081
		(0.216)	(0.318)	(0.304)	(0.306)	(0.312)	(0.305)	(0.277)
Number of observations	183	183	87	96	101	82	67	116
Education spending								
Log(1+school expenditures since application) - Households with kids 7 to 15 years old	6.679	-0.125	-0.103	-0.029	-0.220	-0.341	-0.393	0.232
		(0.256)	(0.341)	(0.430)	(0.355)	(0.350)	(0.342)	(0.387)
Dummy=1 if anybody in the household is a university student	0.118	0.020	0.016	0.010	0.035	-0.008	0.046	-0.022
		(0.028)	(0.042)	(0.038)	(0.042)	(0.035)	(0.045)	(0.035)
Number of observations (range)	269-602	269-602	149-301	120-301	139-305	130-297	133-271	136-331
Housing expenses								
Dummy=1 if bought or improved residence/house since application	0.241	-0.010	-0.001	-0.022	0.028	-0.042	0.054	-0.051
		(0.036)	(0.050)	(0.053)	(0.051)	(0.051)	(0.053)	(0.050)
Log(amount spent for buying or improving residence/house since application), conditional on having purchased a house or made house improvements.	6.988	-0.218	-0.205	0.024	-0.912*	0.236	-0.109	-0.295
		(0.324)	(0.452)	(0.495)	(0.458)	(0.352)	(0.542)	(0.400)
Number of observations (range)	151-626	151-626	74-311	77-315	84-314	67-312	68-283	83-343

Huber-White standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%
All results obtained using OLS to estimate the ITT model detailed in equation (1); each cell presents the estimated treatment effect from a single regression.
Running probits for the binary outcomes produces qualitatively similar results.
The number of observations varies depending on missing values in the survey data.
The income cutoff point is the median income measured at application.
The credit score represents the quality of the application, along two dimensions: (1) the credit bureau score, and (2) an internal score computed by the lender.
The credit score cutoff point separates applicants in the two lowest categories from applicants in the three higher categories.
Change in food quality scale: much worse (1)-much better (5).
The poverty line is the household size-specific 'minimum living level', as computed by the Bureau of Market Research of the University of South Africa (UNISA) in 2001.
Events include wedding, dowry, young men initiation, holiday, parties, ceremonies, and other.

Table 9. Treatment Effects on Measures of Well-Being

	Mean depvar for full sample	Full sample	Gender		Income		Credit score	
			Female	Male	High	Low	High	Low
Well-being								
Stress scale	18.580	1.414	1.245	1.452	2.178	0.632	0.703	1.926
		(0.882)	(1.186)	(1.313)	(1.383)	(1.187)	(1.399)	(1.222)
Depression scale	18.828	-0.264	1.249	-2.749	-0.161	-0.056	-0.639	0.197
		(1.571)	(2.140)	(2.429)	(2.259)	(2.430)	(2.663)	(2.116)
Optimism scale	21.969	0.362	0.176	0.566	0.102	0.654	0.030	0.704
		(0.339)	(0.466)	(0.502)	(0.485)	(0.493)	(0.481)	(0.502)
General health scale	4.344	0.067	0.038	0.069	0.073	0.019	-0.002	0.113
		(0.069)	(0.103)	(0.092)	(0.080)	(0.116)	(0.106)	(0.090)
Number of observations (range)	244-610	244-610	127-308	117-302	120-307	124-303	112-277	132-333
Decision-making								
Decision-making index	13.719	0.865	1.158	1.355*	0.348	1.246	1.135	0.271
		(0.695)	(1.057)	(0.808)	(0.836)	(1.486)	(1.053)	(0.939)
Number of observations	178	178	83	95	116	62	97	81
Shocks (not including job loss)								
Dummy=1 if no shock since application	0.483	-0.014	0.019	-0.045	-0.014	-0.012	0.018	-0.027
		(0.042)	(0.059)	(0.059)	(0.058)	(0.061)	(0.062)	(0.057)
Dummy=1 if no shock in 30 days prior to survey	0.621	0.004	0.015	0.004	-0.028	0.047	0.084	-0.061
		(0.041)	(0.058)	(0.059)	(0.058)	(0.059)	(0.061)	(0.057)
Number of observations (range)	617-619	617-619	309	308-310	309-311	308	282	335-337

Huber-White standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%
All results obtained using OLS to estimate the ITT model detailed in equation (1); each cell presents the estimated treatment effect from a single regression. Running probits for the binary outcomes produces qualitatively similar results.
The number of observations varies depending on missing values in the survey data.
The income cutoff point is the median income measured at application.
The credit score represents the quality of the application, along two dimensions: (1) the credit bureau score, and (2) an internal score computed by the lender. The credit score cutoff point separates applicants in the two lowest categories from applicants in the three higher categories.
A shock is the occurrence of a surprise funeral, birth (if pregnancy was a surprise), theft, catastrophe, loss of livestock in the household, or sickness that required a household member to stay in bed (sickness was only measured for the 30 days prior to the survey). We exclude job loss because Table 8 presents treatment effects on employment.
Perceived stress scale range: 0-40; Depression scale range: 0-57; Optimism scale range: 6-30. Higher scores reflect: higher stress; more depression; more optimism. For details on scale construction and validation see Cohen and Williamson (1988), Radloff (1977), and Scheier, Carver, and Bridges (1994). General health scale range: very bad (1)-very good (5).
Stress, depression, and optimism questions were not asked in the 73 surveys answered by a household member who was not the marginal applicant (this occurred when the marginal applicant was unavailable/had moved out/etc.). Hence the maximum number of observations for these variables is 553. The stress (depression) variables, are missing 7 (13) observations because one or more of the scale components is missing. Due to a survey software bug, we are also missing stress and depression variables for the 46% of the sample that was randomly assigned to be asked stress and depression questions after questions on borrowing.
The decision-making index was based on questions asked to married marginal applicants about how the household decides about: routine purchases, expensive purchases, giving assistance to family members, family purchases, recreational use of money, personal use of money, number- of children, use of family planning, method of family planning, assistance given to relatives, decision to borrow, amount to borrow, and where/who to borrow from. The value for each item takes zero if the decision-making is done by the respondent's house or someone else in the household, one if the decision-making is done by the couple, and two if decision-making is done by the respondent. The index is the sum of the 13 responses (range: 0-26). We could not construct the index for 7 married respondents due to one or more missing components.

Table 10. Estimated Profitability of Marginal and Inframarginal Loans

	All first loans			Loan to borrowers with low credit score			Loan to borrowers with high credit score		
	n		NPV	n		NPV	n		NPV
Marginal Loans									
Total revenue on all marginal loans	151	R	48,888.29	71	R	22,758.36	80	R	26,129.92
Loan losses on marginal loans in default	39	R	27,166.90	15	R	11,002.70	24	R	16,164.20
Cost of funds on all marginal loans	151	R	1,522.67	71	R	604.62	80	R	918.06
Marginal operating costs per loan:									
Processing/Monitoring of all loans	151	R	7,852.00	71	R	3,692.00	80	R	4,160.00
Enforcement of loans in default	39	R	1,131.00	15	R	435.00	24	R	696.00
Total profitability of all marginal loans	151	R	11,215.72	71	R	7,024.04	80	R	4,191.66
Profit per marginal loan	151	R	74.28	71	R	98.93	80	R	52.40
Inframarginal Loans									
Total revenue on all inframarginal loans	1,399	R	550,283.40	298	R	87,598.84	1,101	R	462,684.60
Loan losses on inframarginal loans in default	303	R	267,489.30	91	R	63,487.04	212	R	204,002.30
Cost of funds on all inframarginal loans	1,399	R	17,829.23	298	R	3,346.55	1,101	R	14,482.67
Marginal operating costs per loan:									
Processing/Monitoring of all loans	1,399	R	72,748.00	298	R	15,496.00	1,101	R	57,252.00
Enforcement of loans in default	303	R	8,787.00	91	R	2,639.00	212	R	6,148.00
Total profitability of all inframarginal loans	1,399	R	183,429.87	298	R	2,630.25	1,101	R	180,799.63
Profit per inframarginal loan	1,399	R	131.11	298	R	8.83	1,101	R	164.21

"n" indicates: the number of first-time applicants who were randomly assigned a loan and actually received one (in the "all marginal loans" rows); the number of borrowers who defaulted on their first loan (in the "in default" rows); or the number of first-time applicants who were approved under the Lender's normal criteria during the experimental period (in the "all inframarginal loans" rows).

Average exchange rate during project and survey: 1 US$ = 6.3 Rands.

Revenue is based on actual interest payments through May 2005. Average loan size was R1,044 for marginal loans and R1,260 for inframarginal loans.

A loan is considered in default if the borrower was 3 or more payments late as of the last month for which we have data, May 2005. We assume that loans in default produced no additional revenue or recovery of principal after May 2005.

The cost of funds is the principal lent times a measure of the Lender's opportunity cost: the yield on 91-day South African Treasury bills at the time of the experiment, adjusted to reflect the loan's duration (i.e., for regular principal payments and prepayments).

The marginal operating cost per loan is based on the following assumptions:

- # of hours per loan screening and processing a loan:		0.5
- # of hours per loan monitoring loans:		0.5
- # of hours per loan enforcing bad debt:		1
- actual hourly cost of labor -- Branch Managers:	R	75
- actual hourly cost of labor -- Tellers:	R	29

Appendix Table 1. Treatment Effect on Poverty status: Comparison of Means

	Percentage of households below the poverty line					
	Treatment	Control	t-test of the means difference (p value)	Treatment	Control	t-test of the means difference (p value)
Full sample						
Applicants with probability of approval of 25%	31.4%	50.0%	0.057*			
Sample size	35	102	137			
Applicants with probability of approval of 50%	37.8%	42.9%	0.259			
Sample size	222	261	483			
Gender sub-samples		Female			Male	
Applicants with probability of approval of 25%	21.1%	46.9%	0.067*	43.8%	51.4%	0.585
Sample size	19	32	51	16	70	86
Applicants with probability of approval of 50%	42.9%	48.3%	0.393	32.7%	36.4%	0.558
Sample size	112	143	255	110	118	228
Income sub-samples		High income			Low income	
Applicants with probability of approval of 25%	15.8%	41.0%	0.056*	50.0%	55.6%	0.695
Sample size	19	39	58	16	63	79
Applicants with probability of approval of 50%	25.2%	27.8%	0.641	52.4%	58.6%	0.350
Sample size	119	133	252	103	128	231

* significant at 10%; ** significant at 5%; *** significant at 1%

Income data is missing for 6 borrowers.

Marginal applicants were assigned a treatment probability based on their credit score, with higher scores dictating the 50% probability.

The poverty line is the 'minimum living level', as computed by the Bureau of Market Research of the University of South Africa (UNISA) in 2001.

The sub-samples "high income" and "low income" group applicants with pre-treatment income higher or lower than the median, respectively.