Expanding Credit Access:

Using Randomized Supply Decisions to Estimate the Impacts

Dean Karlan & Jonathan Zinman

January 2008

Contributions to this research made by a member of The Financial Access Initiative and Innovations for Poverty Action.

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
Financial Access Initiative

Jonathan Zinman
Dartmouth College
Innovations for Poverty Action

January 2008

ABSTRACT

Expanding access to commercial credit is a key ingredient of development strategies worldwide. There is less consensus on the role of *consumer* credit, particularly when extended at high interest rates. Popular skepticism about "unproductive" and "usurious" lending is fueled by academic work highlighting behavioral biases that induce overborrowing. We estimate the impacts of expanding access to consumer credit at 200% APR using a field experiment and follow-up survey and administrative data. The randomly assigned marginal loans produced significant net benefits for borrowers across a wide range of outcomes. There is also some evidence that the marginal loans were profitable.

---

*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, Michael Anderson, Abhijit Banerjee, Jeff Kling, Doug Staiger, Peter Tufano, and Chris Udry for helpful comments. We are grateful to the National Science Foundation (SES-0424067 and CAREER SES-0547898), BASIS/USAID (CRSP), the Bill and Melinda Gates Foundation through the Financial Access Initiative, 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.

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 positive correlations between depth of access and poverty rates at the macro level (Levine 1997; Honohan 2004), positive impacts of access to microfinance at the micro level (Pitt and Khandker 1998), and positive impacts from expansions of bank branch networks on aggregate poverty (Burgess and Pande 2005). 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 about lending at "usurious" rates.[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. Traditional approaches to microcredit expansion— creating new microfinance institutions, adding branches, using joint

---

[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 (Consumer Credit Research Foundation 2006). The United States recently passed a binding interest rate ceiling on consumer loans to military personnel and their family members. The law was motivated by the concentration of payday lenders, which offer a shorter-term version of the loan product studied in this paper, near military bases; see, e.g., http://www.responsiblelending.org/issues/payday/briefs/page.jsp?itemID=29862357.

[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 (2007b; 2007a) find that consumers who systematically underestimate the interest rate on short-term installment loans also borrow heavily and expensively.

liability mechanisms to overcome high fixed transaction costs and poor screening, monitoring and enforcement capabilities— may not be the most cost-effective method to support efficient expansion. Another way to expand access to credit is simpler: liberalizing screening criteria, either explicitly through credit scoring or implicitly by encouraging loan officers to lower their subjective criteria for approval.[3]

We assess the impacts of liberalizing credit screening criteria by analyzing new data produced by a field experiment, a follow-up survey, and administrative data on creditworthiness over time. The key questions are threefold. First, do credit constraints bind? Second, does relaxing any credit constraints benefit marginal borrowers? Revealed preference logic says it should: a consumer borrows only if she will benefit (weakly, in expectation). Behavioral models say not necessarily: biases in preferences or cognition may lead consumers to overborrow. The third key question is whether the lender profits from making these 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 forthcoming); as Attanasio, Goldberg, and Kyriazidou (forthcoming) 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-

[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.

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 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 added three additional steps to its normal process for new loan applicants. 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." 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. Loan officers were instructed by management to follow the computer's instructions in all cases. But in the third and final step, loan officers had pecuniary incentives to be risk-averse and approved the loan in only 53% of the cases when the computer instructed them to approve.

Our outcome data comes from the Lender's records on repayment and profitability, from credit bureau reports over two years after the start of the experiment, and from household surveys

[4] Studies in developing countries include 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). A growing literature uses natural experiments to study the impact of payday loans in the U.S (Melzer 2007; Morgan and Strain 2007; Morse 2007; Skiba and Tobacman 2007).

conducted by an independent firm at the home or workplace of the marginal applicants six to twelve months after the start of the experiment. The survey measures borrowing activity, loan uses, and a range of proxies for household well-being.

The particulars of the experimental design are critical for interpreting our estimates of the effects of expanding credit access on consumer well-being (and lender profits), and for forming policy prescriptions. The loan officers had discretion over who ultimately received a loan (and only 53% of applicants randomly assigned a "second-look" actually got a loan). This feature of our design mirrors the market mechanism typically used by retail lenders in developing countries to liberalize their screening criteria: senior management encourages front-line personnel to make riskier loans,[5] but then leaves the ultimate credit decisions in the hands of a credit officer, branch manager or loan committee. Thus our experiment is designed to estimate the effects of expanding access through a common approach of screening liberalization.

Estimating the treatment-on-the-treated effects is straightforward under our design. Given that roughly half of those assigned a loan actually got a loan, one can roughly double the average intent-to-treat effects to get an estimate of the effects on those who actually got the loan: the treatment-on-the-treated (TOT) effects. These TOT estimates measure the impacts of interest on a market- and hence policy-relevant sample of interest: those applicants deemed by loan officers to be closest to the margin of creditworthiness.

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.

Measuring the *ultimate* impacts of consumer credit on borrowers presents several challenges. There is no natural summary statistic for household utility; hence we follow evaluations of social policy interventions and measure treatment effects on a range of variables that capture economic

[5] "Encouragement" policies include monetary incentives for front-line personnel, goals (e.g., for portfolio growth), and explicit changes to risk assessment criteria (as in our experiment).

behavior and subjective well-being (Kling, Liebman and Katz 2007). But treatment effect channels may vary across households; e.g., some households may smooth consumption by making critical purchases, others may use loan proceeds to maintain employment in the face of adverse shocks to transportation or family health, others may make investments as more traditionally defined (in self-employment, housing, schooling, or health), while others may benefit in less-tangible ways (becoming more hopeful about future prospects, or acquiring more bargaining power in the household). Consequently we use summary index tests that aggregate across outcomes to address the problem of multiple inference (Anderson 2007; Kling, Liebman and Katz 2007).

We find that expanded access to credit significantly improved outcomes. Over the 6 to 12 month horizon, applicants in the treatment group were significantly more likely to retain their job over the study period, and treatment group incomes were significantly higher. Treated households were also less likely to experience hunger, and had more positive outlooks on their prospects and position. We do find a significant and *negative* impact on other aspects of mental health (depression and stress). But the average treatment effect across all of our economic and subjective outcomes is significant and positive. Over 15 to 27 month horizons, we find a positive impact on having a credit score, and no impact on the score itself. The effects on credit scores cast doubt on the hypothesis that positive treatment effects will turn negative over longer horizons due to debt traps or other delayed realizations of the cost of borrowing.

Perhaps most critically, the confidence intervals for treatment effects on our summary impacts (the overall index of survey outcomes, and credit scores) rule out substantial negative effects. This is important because the default policy regime for consumer credit is restricted access based on the presumption of negative effects on the margin.[6]

[6] This stands in contrast to microenterprise credit, which is often subsidized, and hence raises the issue of evaluating any benefits of expanded access with respect to the opportunity cost of subsidies.

The Lender agreed to implement this experiment because its senior management believed that branch staff applied inefficiently strict underwriting criteria. Prior work suggests that indeed there was little reason to expect that the Lender's risk assessment methods would be fully optimized. See Gross and Souleles (2002) for a specific example, and Allen, DeLong, and Saunders (2004) for a review and discussion of the challenges of retail credit risk assessment and the shortcomings of various methods, including relationship lending and credit scoring. Our estimates of loan profits suggest that the Lender's prior was well-founded. The evidence suggests that the marginal loans were profitable in an absolute sense, although substantially less profitable than inframarginal loans. Exactly how profitable depends on several assumptions about marginal costs and risk-weighting.[7]

In all our results suggest a role for welfare-improving interventions in consumer credit markets but come with important caveats. We only measure some outcomes of interest at 6 to 12 month horizons, and some costs and/or benefits of liberalized access to credit may only materialize over longer horizons. Using a screening liberalization design does not allow us to measure impacts on applicants who "normally" would have been approved; these inframarginal credits have different credit quality than the marginal ones, and hence may have different treatment effects. Nor does our design measure the impacts of another common mechanism for expanding access: the penetration of new markets through, e.g., expanded branch networks (Burgess and Pande 2005). And the external validity of treatment effects in the South African cash loan market is unknown.

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

[7] We cannot simply apply the market test of whether more aggressive underwriting criteria was adopted in steady-state, because the Lender was merged into a bank holding company before the results of the experiment could be applied to company policy, and we do not have access to post-merger data or underwriting policy information.

to evaluate efforts to expand credit access. Most practically, our results suggest that liberalizing screening criteria can benefit both borrowers and lenders, and our methodology demonstrates how lenders can hone in on their sustainability/outreach frontier by taking controlled risks using randomized experimentation.

The paper proceeds as follows. Section II provides background information the applicants, the Lender, and the cash loan market. Section III details the design and implementation of our experiment and data collection methods and empirical strategy. Section IV presents estimates of treatment effects on borrowing and credit access. Section V presents estimates of treatment effects on component and summary index measures of ultimate outcomes of interest. It also presents our estimates of effects on credit scores 15-27 months after treatment, and details our estimates of Lender profits on marginal and inframarginal loans. Section VI concludes with a discussion of external validity and other questions for future research.

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.[8] 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 typically lack the credit rating 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

[8] The Lender was merged into a large bank holding company in 2005 and no longer exists as a distinct entity.

40% of the median borrower's gross monthly income.[9] Our sample for this experiment includes mostly first-time loan applicants of African descent. Table 1 shows some comparative demographics. Table 4 shows 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 30-100% per month. Lenders targeting observably lower risk segments charge as little as 3% per month.[10]

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. The Lender also had a "medium-maturity" product niche in 4-month installment loans. Most other cash

[9] 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.
[10] South Africa has had very low inflation rates in recent years; e.g., 4.35% over our 14-month study period.

lenders focus on 1-month or 12+-month loans.[11] 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).

The Lender did not disclose the APR; South African law does not mandate APR disclosure. Rather interest was charged up front (using the "add-on" practice common in consumer loan markets that lack effective APR disclosure mandates),[12] and the loan was then amortized into 4 equal monthly repayments.[13] But compared to many competitors this pricing was transparent and linear, with no surcharges, application fees, or insurance premiums added to the cost of the loan.

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 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-C), 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.

[11] 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.

[12] See Stango and Zinman (2007b) for evidence on how and why lenders prefer to shroud APRs, and on the mediating impacts of incomplete enforcement of mandated APR disclosure on market outcomes in the U.S.

[13] So a R1,000 loan had monthly repayments of (1000+1000*0.1175*4)/4 = R367.50. Borrowers that prepaid paid add-on interest pro-rated to the time outstanding; e.g., a borrower who stayed current and prepaid her remaining amount at the end of month two would have repaid R367.50 in month one, plus R867.50 at the end of month two, for a total repayment of R1,235 = R1,000 (principal) + R235 (two month's interest).

III. Methodology

Our research design first randomly assigns a "second look" to some marginal rejected applications, and then uses data from the lender, a credit bureau, 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 the lender.

A. Experimental Design and Implementation

<u>Sample and time frame for the experiment</u>

We drew our sample frame from the universe over 3,000 "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.

<u>Experimental Design and Operations</u>

The Lender implemented the experiment in four steps:

First, loan officers evaluated each of about 3,000 new applicants using the Lender's standard 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. During our study period loan officers approved 1,405 new applications based on the standard underwriting criteria. 705 applications were deemed egregious rejects, leaving us with a sample frame of 787 marginally rejected applicants for the experiment.

Second, special “randomizer” software encouraged loan officers to reconsider randomly selected marginal rejects. 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 applications with probabilities that were conditional on the credit score and loan officer assessment. The treatment was simply a message on the computer screen that the application had been “approved” (control applicants remained “rejected”). 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, Panel A, 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.

We describe the randomizer's treatment as "encouragement to reconsider" rather than "randomized approval" because loan officers had pecuniary incentives to be risk-averse and not comply with the randomizer's decision. The Lender deemed it impractical ex-ante to try to align pecuniary incentives with randomizer compliance (note we use the term "compliance" in the econometric sense, not in a layman sense, since the bank officers were not forbidden from refusing the suggestion from the randomizer software). Instead we relied on training and persuasion, and we also monitored the compliance rate in order to gauge how strong this policy change would be in relaxing lending criteria. Table 2, Panel B shows the compliance rates. Not surprisingly compliance was high in the control (still rejected) group: only 2% of these applicants received a loan during the experimental period. But compliance was middling for the treatment (approve) group: only 53% actually received a loan.

Imperfect compliance motivates conducting 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 Sections III-D and III-G for more details).

Accepted applicants were offered an interest rate, loan size, and maturity per the Lender's standard underwriting criteria. Recall that nearly all received the standard contract for first-time borrowers: a 4-month maturity at 200% APR. Loan repayment was monitored and enforced according to normal operations. Branch manager compensation was based in part on loan performance, and as noted above the experiment did not change incentive pay.

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-A, the surveyors asked questions on demographics, resources, recent investments, employment status, income, consumption, and subjective well-being.[14]

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 73 of these cases the targeted respondent (i.e., the loan applicant) could not be located, and someone else from the household was surveyed. In order to avoid potential response bias between the treatment and control groups, neither the survey firm nor the respondents were 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 threefold. First, it avoids one type of 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 avoids another type of mechanical timing bias in favor of finding positive impacts on credit access by evaluating impacts well after the maturity date on the marginal loans. This ensures that we do not simply measure an initial spike of consumption, and that we can observe which marginal borrowers defaulted on their loans. Third, the 6-12 month horizon (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. To measure longer-term effects, after 15 to 27 months we obtained credit reports from a credit bureau for each of the applicants in the experiment (see Section V-B).

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

C. Internal Validity

As noted above, our methodology requires obtaining survey data on both treatment and control households. Hence survey sample attrition would threaten the internal validity of the results from our experiment, since the random assignment 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 2, Panel A, Column 3 corroborates that this condition holds: treatment status is uncorrelated with survey completion. Column 4 highlights that applicant characteristics were balanced across the surveyed and not-surveyed groups. We also have administrative outcome data we can use to measure treatment effects on the not-surveyed: the Lender obtained follow-up credit scores on the entire sample frame of 787 marginal applicants. Table 7 shows that the treatment effects are statistically identical across the surveyed and non-surveyed groups, and we discuss these results in Section V-B.

D. Intention-To-Treat Estimates for Component Outcomes

Imperfect compliance with the random assignment to the treatment group motivates an intention-to-treat (ITT) estimator. ITT produces an unbiased estimate of *average* treatment effects even when there is substantial noncompliance. We implement ITT using the following OLS specification:

$$(1)\quad Y^k_i = \alpha + \beta^k assignment_i + \delta risk_i + \phi appmonth_i + \gamma surveymonth_i + \varepsilon_i$$

Y is a behavior or outcome of interest k for applicant i (or i's household). Examples of Y include measures of borrowing (Table 3), poverty status (Table 5), and loan repayment (Table 8). $Treatment_i$ = 1 if the individual was *assigned* to treatment (irrespective of whether they actually received a loan). $Risk_i$ captures the applicant's credit score; this determined whether the applicant was treated with probability 0.25 or 0.50. $Appmonth_i$ 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.[15]

E. Inference Over Multiple Outcomes: Summary Index Tests

Two concerns arise when using equation (1) to conduct statistical inference over multiple outcomes. One is Type I error(s). The probability that one or more treatment effects is labeled statistically significant due to chance is increasing in the number of outcomes (i.e., in the number of tests preformed). The second concern is evaluating the overall direction and magnitude of the treatment effects when there is a diffuse set of outcomes. We address these concerns using summary index tests.

Following Kling et al (2007), we construct summary indices at two levels: 1) *domains* of related outcomes, and 2) an overall measure that aggregates all of our ultimate outcomes of interest. Our domains are: economic self-sufficiency (income and employment status), food consumption, investment (in housing, education, and self-employment), physical health, mental health, and outlook and control (optimism, intra-household decision power, and self-perception of community status).

We construct indices by first rescaling each outcome Y^k_{ij} (outcome k, for individual i, in domain j) so that higher values map into better outcomes. Next we standardize each outcome into a z-score by subtracting its control group mean, and dividing by its standard deviation.[16] Then we

[15] 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).

[16] Following Kling et al, in constructing indices we impute missing outcomes using the mean of the individual's assigned treatment group. For most outcomes and domains we have few missing values and hence do little imputation; one can see this by comparing the sample sizes for the individual outcomes in Table 5 to our surveyed sample size of 626. As Kling et al note (in their footnote 11), this rule "results in differences between treatment and control means of an index being the same as the average of treatment and control means of the components of that index (when the components are divided by their control group standard deviation and have no missing value imputation), so that the index can be interpreted as the average of results for separate measures scaled to standard deviation units." We do resort to substantial imputation for the mental health outcomes and decision power; see Section V-A for details.

combine outcomes in a domain j by taking the average of equally-weighted standardized components. Then our summary index analog to equation (1) is:

$$\text{(2)} \quad Y^j_i = \alpha + \beta^j \textit{assignment}_i + \delta \textit{risk}_i + \phi \textit{appmonth}_i + \gamma \textit{surveymonth}_i + \varepsilon_i$$

Where Y^j_i is an average z-score: the average of standardized component outcomes in domain j.

F. Heterogeneous Treatment Effects

The average intention-to-treat effect is captured by β^k in equation 1, or β^j in equation 2. 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; Morduch 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 as measured by the Lender's matrix of internal and external credit scores.

G. Treatment-on-the-Treated Effects

As discussed in the introduction, treatment-on-the-treated (TOT) effects are important. They measure the impacts on the marginal borrowers deemed most creditworthy by the credit officer, whereas the ITT estimates the average impact on those "reconsidered" through the objective computer-calculated credit scoring procedure The TOT in this design is easy to calculate: it is the

ITT estimate divided by the difference in the rate of loan officer compliance with the random assignment across the treatment and control groups. As Table 2 Panel B shows, this difference is roughly 0.5 (actually 0.45), so one can obtain a rough estimate of the TOT effect on any outcome or summary index in our study by simply doubling the relevant ITT estimate.

H. External Validity

There are three main external validity issues to consider when interpreting our findings.

One external validity issue is the representativeness of our sample. 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 industry criteria (or new regulatory restrictions). The Conclusion discusses some related markets and policy issues in both developing and developed countries.

The second issue relates to the lender's mechanism we study for expanding access. Our Lender's liberalization of credit screening criteria relied ultimately on loan officer discretion. Consequently our results will extrapolate better to settings where, as is common, firms expand lending through "encouragement" designs. They will not necessarily apply to settings where firms expand by adding branches, on the other hand.

The third external validity issue relates to measuring treatment effects on medium-run outcomes Section III-B detailed why we chose 6-12 months for survey data collection on credit access and well-being measures. We consider 15 and 27 month impacts on credit scores, and address the possibility of time-varying treatment effects, in Section V-B.

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 materially unless 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 3 reports 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 (Panel A, "all sources"). But treated households did respond on the intensive margin of overall borrowing: Panel A shows a significantly higher quantity of loans from all sources (the total number of loans per person rises by 0.141, or 28%).

Both the extensive and intensive margins of borrowing also show a change in the *type* of credit accessed. Treated households were more likely to report borrowing from a microlender (our 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.[17] 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.[18]

Table 3, Panel A 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.

[17] 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.

[18] 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). 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. See the Conclusion for additional discussion of interactions between formal and informal credit markets.

The estimates in Panel A are likely attenuated by systematic underreporting of borrowing. Our companion paper finds that 50% of survey respondents known to have borrowed from the Lender during the 12 months preceding the survey do not report *any* borrowing in the survey (Karlan and Zinman 2008). Consequently this suggests the true ITT impacts on borrowing outcomes are probably twice as large as those estimated in Table 3.

As discussed in the companion paper the most likely explanation for debt underreporting is social stigma; e.g., underreporting is significantly more prevalent among females, and more prevalent yet when female respondents are interviewed by male surveyors. Consequently there is little reason to believe that estimates of the other treatment effects we consider in Section V (on economic and well-being outcomes, such as employment) are also correlated with treatment and thus attenuated (or exacerbated) by underreporting. Recall that we will also measure impacts (on credit scores) based on administrative data that does not rely on self-reports.

Table 3, Panel B 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?" Treated applicants were 12.8 percentage points more likely to report "Microlender or Cash lender" than the control group. Treated households were 11.2 percentage points less likely to report an informal source (friends, family, moneylender, or borrowing circle). Both effects are statistically significant with 99% confidence. 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 Panel B addresses whether the change in marginal source is due (partly) to formal access crowding-out informal access.[19] 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 6.2 percentage points (8.6%), although the result is not statistically significant.

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

Table 3, Panel B 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 4 shows the range of activities households report financing in the survey. These loan uses motivate estimating treatment effects on a particular set of investments and economic outcomes. We then also estimate treatment effects on various measures of subjective well-being. In each case we scale outcomes such that positive coefficients on the intention-to-treat variable (where 1= assigned a loan) indicate positive treatment effects. Details on how we construct outcome measures from the survey data can be found in the Data Appendix. Estimated treatment effects for each "component" survey outcome are reported in Table 5.

As discussed in Section III-E, the large set of component outcomes that could be affected by access to consumer credit motivates aggregating across outcomes and then estimating treatment effects on these summary indices (Table 6). Recall that each index component is a z-score, and that each index value is the average z-score of its component outcomes for the given individual. Consequently our estimate of the treatment effect for index j is an estimate of the average effect on each outcome in j, in standard deviation units.

A. Loan Uses, and ITT Results on Ultimate Outcomes

Table 4 shows that 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.

Measuring total consumption requires far more survey time than we could allot (Deaton and Zaidi 1999), given the many other outcomes of interest, so we focus on measuring two simple measures of food consumption. One is whether anyone in the household experienced hunger in the past 30 days (14% of households in the sample reported some hunger). The other is whether the quality of food consumed by the household improved over the last 12 months (26% reported an improvement). Households randomly assigned a loan were an estimated 5.8 percentage points less likely to report hunger (with a p-value of 0.03), and 3.7 percentage points more likely to report a food quality improvement (although this estimate was not statistically significant, with a p-value of only 0.32). Again, recall that these measures were taken well after the initial loan repayments were due on the marginal loans, so these treatment effects are not simply picking up a very transitory spike in consumption. Combining the two measures of consumption into a summary index (Table 6) produces a significant estimated treatment effect of 0.12 standard deviation units.

Table 4 shows that the next most common purpose for household borrowing is transportation expenses (19.4%); 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.

Two related results (not shown) are consistent with the story that marginal borrowers use loan proceeds to make investments in maintaining wage employment. First, questions on job history reveal that treated applicants were indeed significantly less likely to report leaving a job since entering the experiment. The point estimate (-2.8 percentage points, with a standard error of

1.4pp) is smaller than the estimated effect on employment status, but the confidence intervals do overlap. Second, we find a positive point estimate (+ 2.1 percentage points, with a standard error of 2.5pp) on the likelihood that treated households repaired their car in recent months. And again the confidence interval overlaps with the one for the treatment effect on employment.

The effects on employment, and microfinance's focus on poverty reduction, motivate estimating treatment effects on income as well. Measuring income accurately in developing country settings tends to be difficult (Deaton and Zaidi 1999), and so we focus on relatively discrete measures in hopes of mitigating noise. One measure is the household's percentile in the survey sample distribution of employment earnings since entering the experiment.[20] Another measure is whether total household income exceeds the poverty line. We find positive treatment effects on both measures. Households randomly assigned a loan earned an estimated 5 percentile points more income (p-value = 0.06), which translates to an increase of roughly R3,500 (or 16% of median income) in the middle of the sample distribution. Treated households were 7.4 percentage points (p-value = 0.07) more likely to be above the poverty line, a 12% increase over the sample mean (or equivalently, a 19% reduction in the number of households in poverty).

Table 6 combines our three measures of employment and income into an "economic self-sufficiency" index. The overall treatment effect is positive, large (0.19 standard deviation units), and highly significant (p-value = 0.002). The sub-group estimates (recall that our income split is based on income prior to the treatment) suggest homogeneous treatment effects.

The loan uses table also suggests estimating treatment effects on certain investments.[21] 13.7% of loans are used for educational expenses.[22] Households report almost perfect attendance

---

[20] The functional form of the earnings distribution makes it such that our OLS estimator puts more weight on the bottom part of the income distribution, where the income level difference between percentiles is smaller, than on the rightward part of the income distribution, where starting around the 75th percentile the level difference in income across percentiles increases dramatically.

[21] Many households report financing events, but 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 (results not reported). Given measurement error we have little hope of identifying any treatment effect on the intensive margin (event spending), so we do not include events in our analysis.

among compulsory school-aged children, so we focus on university attendance for households with any member between ages 18 and 26. The estimated treatment effect is essentially zero, and imprecisely estimated. Another frequent use of loan proceeds is housing expenses (11.5%). We estimate that treated households were 4 percentage points (13%) more likely to purchase or improve a house since entering the experiment (Table 5), with a p-value of 0.30. We also estimate the treatment effect on self-employment in the household. It is plausible that cash loans are a viable option for financing self-employment even at 200% APR, since microentrepreneurial credit is very scarce in South Africa, and the returns to microenterprises may be very high for the relatively poor and credit constrained in developing countries (de Mel, McKenzie and Woodruff 2007). Reported prevalence of using loan proceeds to finance business activity is low (3.2%), but may be underreported (since some consumer lenders actively discourage "informal sector" employment), or subsumed in other categories. We estimate that the likelihood of self-employment is 2 percentage points (13%) higher in the full sample, but with a p-value of only 0.5. However, *ex-ante* low-income treated applicants were an estimated 9 percentage points more likely to be self-employed, with a p-value of 0.07. The point estimate is large, given the mean self-employment rate of 15.7% among low-income households in our sample.

Table 6 shows that the estimated treatment effect on our "investment" index—combining self-employment, housing, and university attendance— is small, positive, and insignificant.[23]

We also estimate treatment effects on various subjective measures of well-being. We start by estimating treatment effects on three outcomes we group together as "control and outlook." One outcome is a measure of decision-making power. Many microfinance initiatives seek to increase the intra-household bargaining power of female borrowers.[24] Here we find point estimates that

---

[22] 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.

[23] Here we assume zero education treatment effects on households with no members in the likely university age range of 18 to 26.

[24] For evidence from prior studies see Pitt, Khandker, and Cartwright (2003) on credit program participation, and Ashraf, Karlan, and Yin (2007) on a commitment savings product.

are consistent with positive effects on borrowers of both genders, although the treatment effect on females is imprecisely estimated (Table 5). Our sample size is relatively small here because we asked the decision power questions of married targeted respondents only. We also construct a standard, linear measure of optimism using a battery of questions from the psychology literature. We find insignificant, positive, and small estimated treatment effects: the largest magnitude contained in the 95% confidence interval implies only a 5% increase in optimism. The third outcome is the respondent's perception of her standing on a ladder of socio-economic status in her community/neighborhood. The estimated treatment effect is essentially zero, and the confidence intervals rules out shifts greater than 10%. Combining the three outcomes into a summary index measure produces a positive and highly significant overall treatment effect on control and outlook (Table 6).[25]

We also estimate treatment effects on two measures of self-reported physical health status. The first is based on the question: "Would you say your health at this time is very good, good, fair, bad, or very bad?" Table 5 shows that treated applicants were an estimated 4.7 percentage points more likely to label their own health status as "very good", with a p-value of 0.26. The second outcome is based on questions about recent sickness in the household. The coefficient suggests that treated applicants were slightly more likely to report sickness among household members in the last 30 days, although the point estimate has a p-value of only 0.54. Combining the two measures into a physical health index produces a very small and imprecisely estimated positive treatment effect (Table 6).

[25] Constructing the index requires an assumption about how to impute decision power for the unmarried, since we asked our decision-making power questions only of married respondents. We impute decision power for an unmarried respondent using the mean of the respondent's treatment cell for married respondents; effectively assuming that the treatment effect is the same magnitude (albeit in different intra-household or extra-household domains) for unmarried respondents.

Finally we construct two measures of current mental health status. The first is a standard, linear measure of depression based on a battery of questions from the psychology literature.[26] The estimated treatment effect on depression is very small, and the confidence interval includes a maximum shift of 15%. The second mental health measure is linear stress scale that is again based on a standard battery of questions. Here we find our first hint of a negative treatment effect: the point estimate implies an 8% increase in stress, with a p-value of 0.11.[27] Our sample sizes are relatively small on the mental health measures because the questions were designed to be asked only of targeted respondents, and also were inadvertently skipped for approximately half of the remaining sample due to a survey software bug. Combining the two mental health measures into an index produces an estimated negative treatment effect of 0.15 standard deviation units, with a p-value of 0.06 (Table 6). The estimates by sub-group suggest that there may be heterogeneous treatment effects; e.g., we find significant negative effects on female but not male borrowers.

The final row of results in Table 6 shows estimated treatment effects on the summary index that combines all of our outcome measures. This index captures the estimated average treatment effect on a component outcome. The estimate is highly significant (with a p-value of 0.02), and suggests that access to consumer credit improves the average outcome by 0.07 standard deviation units. In one sense the economic magnitude of this effect is somewhat challenging to put into

[26] The depression scale includes measures of happiness that merit separate mention given the recent interest in using happiness as an outcome measure. We find positive but insignificant treatment effects on the happiness scale, and on a dummy for being happy "most of the time". As in other datasets, our happiness measures correlate strongly and positively with being (self-)employed.

[27] Besides the possibility that servicing debt creates stress (recall that point estimates in Table 3 suggest that treated applicants were more likely to be borrowing at the time of the survey), the survey data suggests two other potential channels. One is that increased decision making power may produce conflict. We asked several questions on intra-household conflict; combining the responses into a linear conflict scale produces a large, but insignificant, estimated increase in conflict. A second possibility is that access to credit permits spending that borrowers regret ex-post. The estimated treatment effect on whether respondents "agree a lot" that "I often find that I regret spending money. I wish that when I had cash, I was better disciplined and saved it rather than spent it" is positive but insignificant.

perspective, given the lack of randomized and outcome-standardized evaluation of microcredit.[28] But in another sense the magnitude matters less than the conclusion that we can rule out negative summary treatment effects over the horizon considered in our survey data (6-12 months). For as discussed at the outset, the default policy approach to consumer credit is to restrict rather than subsidize access.

B. Time-Varying Treatment Effects and Debt Traps? Effects on Credit Scores Over Time

Despite the fact that our survey measures outcomes several months after loans were due to be repaid in full, there may still be some concern that a 6-12 month horizon is too short to capture the full cost of loan repayment in some cases. Similarly, returns to some investments, broadly defined, that are financed with the marginal loans may not be fully realized over 6-12 months. Indeed, some debt trap models imply that marginal borrowing may actually be *counter*productive in the long-run; i.e., that treated applicants may have worse outcomes than untreated applicants over longer horizons.[29] So measuring outcomes and estimating treatment effects over longer horizons is important. But survey data are expensive, and increasingly prone to attrition bias as the treatment grows more distant in time. Thus we address the question of time-varying impacts using administrative data, using credit scores obtained from a leading credit bureau on nearly everyone in our survey sample frame as of two dates: December 31st, 2005 (13-15 months after the initial application), and December 31st, 2006 (25-27 months after the initial application).

Credit scores may be useful outcome measures in three respects. First, credit scores may proxy more directly for ultimate outcomes if they are correlated with said outcomes. The 2005

[28] On the other hand education and other social policy initiatives are more commonly evaluated using these methods. Randomized education treatments are typically thought to have a large impact if they move test scores by 0.2 standard deviation units. The point estimate for the overall effect of the Moving to Opportunity intervention studied in Kling et al was 0.04 standard deviation units on adults (with effects 2-3 times as large on youths). The closest study to ours is Ashraf, Karlan and Yin (2007) in which a commitment savings product in the Philippines led to an increase in decision-making power of 0.50 standard deviations for married females who prior to the experiment had less than median power.

[29] Debt traps refer to a dynamic where borrowers are unable to fully service debt out cash flows, refinance or continue borrowing over longer horizons than the original maturity, and ultimately default or bear extreme costs due to long-term and expensive borrowing.

scores are all measured within 9 months of our survey data, and the December 2005 credit score is actually negatively correlated (-0.10) with the overall summary index—for those with a score. But applicants with a thin credit history are not scored, and having a score is correlated positively (0.12) with our overall index. Second, having a score may not only be privately beneficial (as suggested by its positive correlation with the overall index), but socially beneficial, to the extent it indicates that private information about the borrower's creditworthiness has been made public to lenders. Third, debt traps or other delayed realizations of borrowing costs may ultimately culminate in borrowers defaulting, so we can estimate whether expanding access to credit in the short-run eventually *reduces* creditworthiness on the longer-run (by inducing defaults).

Credit scores are used by consumer lenders in South Africa much as they are in the U.S. Scores can range from 300 to over 850. Our sample had December 2005 and 2006 scores ranging 487 to 817.[30] Our Lender made loan approve/reject decisions with reference to the external credit score (along with an internal score, and soft information collected and assessed by branch personnel). External scores had little if any impact on the loan terms offered conditional on approval. The Lender rarely made loans to applicants with scores below 600, and almost never to applicants below 550. Approval probabilities (based on a matrix of the external and internal scores) were based on 20-30 point external score bands.

But the most important effect of external credit scores on creditworthiness in the cash loan market likely comes from the extensive margin, since many consumers have credit histories that are too thin to be scored. These consumers do not have any score at all, or are assigned a 3-category risk indicator by external score provider. Obtaining an ordinal score increased the probability of loan approval in our sample by 19%, conditional on the Lender's internal score, branch fixed effects, and month of application.

Table 7 provides evidence that our expanding access treatment significantly increased the probability of having a score, and had no effect on the score conditional on having a score. Panel

[30] The 2005 and 2006 scores are correlated 0.50 in our survey sample frame and surveyed samples.

A shows results for the surveyed sample of 626 households (Panel B shows that results on the sample of 787 households that we *attempted* to survey are very similar). Columns 1 and 2 show that marginal applicants who were randomly assigned a loan were an estimated 7.6 and 6.7 percentage points more likely to have an ordinal score after one year and after two years. These are large effects given that 10% and 12% of the sample lacked an ordinal score. On the other hand, we find no evidence that the treatment changed scores conditional on having an ordinal score. The 95% confidence interval bounds the intention-to-treat effect at a small one; e.g., -11 points is a less than 2% change relative to the sample mean. Scores are nearly normal distributed, so results for logged scores produce nearly identical results.

In all, we do not find any evidence that expanding access to consumer credit reduces creditworthiness over a 2-year horizon. If anything the treatment seems to have had a (socially) beneficial impact on creditworthiness by increasing the probability of obtaining a credit score.

C. Impacts on the Lender: Profitability

As noted at the outset, the Lender implemented this experiment based on the prior that its branch staff were overly conservative in applying the risk assessment guidelines provided by senior management. Prior work on retail credit risk assessment suggests that the Lender had every reason to be concerned that its risk assessment model was not fully optimized (Allen, DeLong and Saunders 2004). The particular related questions of interest in our experiment are: were the marginal loans produced by the experiment profitable? And were they less profitable than inframarginal loans?

Table 8 reports our profit estimates for the 172 marginal loans that loan officers originally rejected but decided to approve after our randomized second look (Panel A), and for the 1,405 inframarginal loans to first-time borrowers that loan officers in the experimental branches initially approved during the experimental period (Panel B). Below we refer to the marginal and inframarginal loans together as "study" loans.

We calculate gross revenues on the study loans by discounting all payments made on these loans (including principal, interest, and late fees) back to the start date of the experiment. Since the Lender was not credit constrained—in fact it was highly profitable and financed study loans out of retained earnings-- we discount using a risk-free rate (the South African Treasury security with the most comparable maturity, which was 91 days, with an annual yield of 7.2%, during our study period). Our repayment data ends in May 2005 (due to the merger described above), but by this time nearly all study loans that had not been paid back in full were seriously delinquent (>= 90 days past due). So we assume that no additional payments were collected on study loans after May 20th, 2005.

We then calculate net revenues by subtracting the discounted loan amount advanced to get an estimate of profits, assuming no marginal staff costs.

The question of how to account for marginal staff costs hinges in part on whether there was an opportunity cost of staff time. The Lender did not hire any new staff for this experiment, nor did it incur any additional marketing expense. But there may be a shadow cost if processing, monitoring, and enforcement if marginal loans reduce the amount of staff time allocated to the same activities on inframarginal loans. We estimate this shadow cost using the Lender's estimate of marginal labor costs and quantities for each type of activity.

Whether we account for marginal costs or not, Table 8 suggests two key qualitative findings. First, marginal loans appear to have been substantially less profitable than the inframarginal loans (Column 1). Marginal loans were less likely to have been paid back in full (71.5% vs. 76.4%); the p-value that the inframarginal repayment rate is in fact higher is 0.08. The table also shows that our point estimates for average loan profitability are higher for inframarginal loans. The table reports the p-value for a test of whether the profit difference between inframarginal and marginal loans is different from zero; the probability that it is *greater* than zero is 0.10.

Interestingly, Column 2 suggests that the Lender's screening method did a poor job of distinguishing profitable from unprofitable loans at relatively low ex-ante credit scores (defined based on the Lender's matrix of internal and external scores).

Second, we find substantial, risk-unadjusted profits on marginal and inframarginal loans alike. The question of whether and how much to adjust for risk is important. From the perspective of society, unadjusted profits may be the relevant input into social welfare analysis: one usually assumes that the social planner is risk-neutral. From the perspective of the Lender, some adjustment is probably warranted. Any risk adjustment would presumably increase the profitability gap between inframarginal and marginal loans. Nevertheless we note that the Lender's management concluded that our conservatively estimated profit of R201 ($32) per marginal loan easily exceeded its hurdle. This is unsurprising given that, holding fixed our other assumptions, the Lender's discount rate would need to rise to 87% to make the marginal loans profitable in risk-unadjusted terms (and to 119% if we assume no marginal staff costs.)

The conclusion that the marginal loans were profitable to some degree would likely be strengthened if we had more complete data on additional loans obtained by marginal clients. In principle of course a firm cares about the present value of all expected future transactions with the marginal *client*. Typically the average profitability of the Lender's "follow-on" loans was substantially higher than on the first loan, as loan sizes and maturities rose and default rates fell for more experienced clients. Our data suggests that marginal clients followed the typical pattern, although since the data is truncated at May 2005 we cannot "close the books" on repayment of follow-on loans.

In all the evidence suggests that the marginal loans induced by our experiment were profitable, although substantially less profitable than comparable inframarginal loans. We do not harbor illusions that our profitability estimates are precise, as our calculations are based on several debatable assumptions. We detail our best guesses in Table 8 but emphasize that the magnitudes presented there are speculative. Nevertheless the weight of the evidence suggests that the marginal loans were profitable to some degree, particularly if one takes the risk-neutral perspective of a social planner.

In any case we believe the main implication of our profit estimates is that consumer lenders should seriously consider evaluating their risk assessment models. Taken together with evidence from prior studies that even profitable consumer lenders do not necessarily operate at the frontier, our experiment highlights the potential bottom-line benefits of controlled experimentation with screening criteria.

VI. 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 randomly encouraged loan officers to reconsider marginal applications for market-rate, four-month term loans that they normally would have rejected.[31] 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 6 to 27 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, treatment applicants who were randomly assigned a second look increased their total borrowing, and changed their lender type composition, in the 6-12 months following the experiment. Second, we find that treated applicants benefited from the expanded access. We use household surveys to measure a range of tangible and subjective outcomes 6-12 months

[31] 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.

following the experiment, and find significant and positive effects on job retention, income, food consumption quality and quantity, and household decision-making control and mental outlook. We find negative effects on other aspects of mental health (principally stress). But on net the impacts are significant and positive. We do not find any evidence that the positive 6 to 12 month impacts are transitory and driven by borrowers who have yet to realize the full costs of borrowing. Over 15 to 27 month horizons we find that the treatment increased the likelihood of having an external credit score, and had no effect on the score itself. Third, our evidence suggests that the marginal loans were profitable. This is particularly true if we take the risk-neutral perspective of a social planner.

Most importantly, we do not find any evidence that the net effects of expanded access to expensive consumer credit are negative. The default policy prescription in South Africa and much of the rest of the world (including parts of the U.S.) is to restrict access based on the presumption that vulnerable consumers overborrow in these markets. Our evidence casts doubt n this presumption: consumers who borrowed at 200% in our experiment benefited from doing so, at least relative to their outside options.

As we noted at the outset, there are three types of external validity issues to consider in interpreting our results. One concerns the proper interpretation of the results given the credit delivery mechanism used for expanding access. The second concerns the applicability of results from the South African cash loan market to other credit markets of interest. The third concerns the time horizon for measuring impacts.

We measure the impacts of expanding access through a particular—and common— delivery mechanism. Our lender liberalized its credit screening by encouraging (but not requiring) its branch staff to issue more loans. Thus our design is exactly the right one for measuring the impacts of liberalizing credit screening. But, our estimates should not necessarily be used to predict the impact of other mechanisms for expanding credit access, in particular ones that do not involve a subjective supply decision by front-line lender personnel.

We experimented in a particular setting that is not necessarily representative of other markets, populations, or interventions. However, high-risk consumer lending is growing in many settings across the world. In developing countries entry is both top-down, as incumbent commercial banks and finance companies expand into "sub-prime" segments, and bottom-up, as "micro-lenders" expand their focus from microenterprises to poor households more generally. Our findings are useful because practitioners and policymakers tend to view our setting as one where the deck was mostly stacked against finding beneficial impacts. Our lender was for-profit, targeted consumers rather than entrepreneurs, and did not necessarily disclose an APR (nor was it required to do so by law). Moreover the intervention was blunt, the credit was expensive, and the market was somewhat competitive.

We measure medium-run impacts (6 to 12 months for our well-being measures, 15 to 27 months for credit scores). Measuring longer-term impacts would round out the research and policy picture, but is likely to be difficult (due to attrition from the survey sample frame) and costly.

Replications will be required to determine whether our findings generalize. Future work would also do well to explore some additional mechanisms behind the effects of expanding access to credit. For example, collecting additional data on preferences, cost perceptions and informal sector borrowing would help shed some light on whether marginal borrowers benefit because they have time-consistent preferences and unbiased perceptions of borrowing costs, or because overborrowing borne of present-bias(es) is less costly at formal market rates.

Our main point of generality is methodological. A field experiment followed data collection 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 2007) and how targeted populations make decisions (Bertrand, Karlan, Mullainathan, Shafir and Zinman 2008; Karlan and Zinman forthcoming). Taken together this layered approach can be used to identify markets that are ripe for welfare-improving interventions, to design mechanisms that are most likely improve efficiency, and then to 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.

Data Appendix.
Construction of Component Outcome Measures and Indices

All outcomes described in this appendix are based on data collected from the follow-up household surveys described in Section III-B.

Measuring the Component Outcomes Evaluated in Table 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. We compare households to the poverty line for annual income using a measure of total household income that is constructed by querying for monthly income over the prior 12 months in several different categories of employment, business, property, and program income.

We construct the percentile of total household earnings reported since entering the experiment using questions on the wage and self-employment earnings of each household member, over the prior 12 months. The percentile is based on the distribution of those with non-zero earnings; we set the percentile to zero for 59 households that report zero earnings over the past 12 months.

The decision-making scale 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 spouse 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). The decision-making scale 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.). We could not construct the index for 7 married respondents due to one or more missing components.

The optimism scale ranges from 6 to 30 and is based on the responses to 6 questions. Respondents rank their level of agreement with statements on a 1-5 scale from, and the optimism score is the sum of the responses. See Scheier, Carver, and Bridges (1994) for details on scale construction and validation.

The community socio-economic ladder scale ranges from 1 to 10 and is based on the response to the question “Think of this ladder as representing where people stand in your community or neighbourhood. People define community and neighbourhood in different ways; in this instance we are referring to the people that live around you or with whom you interact on a regular basis. Imagine everyone in your community or neighbourhood is standing somewhere on this ladder. At the TOP of the ladder are the people who are the best off-those who have the most money, the most education, and the most respected jobs. At the BOTTOM are the people who are the worst off-who have the least money, least education, and the least respected jobs or no job. The higher up you are on this ladder, the closer you are to the people at the very top. The lower you are, the closer you are to the people at the very bottom. Where would you place yourself on this ladder, compared to others in your community or neighbourhood?”.

The depression scale ranges from 0 to 60 and is based on the responses to 20 questions. Respondents indicate how often they felt like a certain way during the past week, with “most or all of the time” scoring 3 points and “rarely or none of the time” scoring 0 points. We then sum the scores and multiply the scale by -1 so that higher score reflect less depression. See Radloff (1977) for details on scale construction and validation.

The stress scale ranges from 0-40 and is based on the responses to 10 questions. Respondents indicate how often they felt or thought in a certain way during the last month, with “very often” scoring 4 points and “never” scoring 0 points. We then multiply the scale by -1 so that higher scores reflected less stress. See Cohen and Williamson (1988) for details on scale construction and validation.

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.). The stress, depression, and optimism scales variables are missing 7, 13, and 2 additional 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.

<u>Combining the Component Outcomes into the Indices Evaluated in Table 6</u>
Indices are created by adding related outcome measures together (after imputing missing values and standardizing as detailed in Section III-E), and taking their unweighted average.

Components for each index are listed in Table 5.

The overall index includes each of the component outcomes listed in Table 5.

REFERENCES

Allen, Linda, Gayle DeLong and Anthony Saunders (2004). "Issues of Credit Risk Modeling in Retail Markets." *Journal of Banking and Finance* 28: 727-752.

Anderson, Michael (2007). "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." Working Paper. July.

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

Ashraf, Nava, Dean Karlan and Wesley Yin (2007). "Female Empowerment: Further Evidence from a Commitment Savings Product in the Philippines." Working Paper. June.

Attanasio, Orazio P., Penelopi K. Goldberg and Ekaterini Kyriazidou (forthcoming). "Credit Constraints in the Market for Consumer Durables: Evidence from Micro Data on Car Loans." *International Economic Review*

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

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

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

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

Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir and Jonathan Zinman (2008). "What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment." Working Paper. January.

Burgess, Robin and Rohini Pande (2005). "Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment." *American Economic Review* 95(3): 780-795.

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

Consumer Credit Research Foundation (2006). Personal communication. August.

de Mel, Suresh, David McKenzie and Chris Woodruff (2007). "Returns to Capital in Microenterprises: Evidence from a Field Experiment." World Bank Policy Research Working Paper 4230. May.

Deaton, Angus and Salman Zaidi (1999). "Guidelines for Constructing Consumption Aggregates for Welfare Analysis." Living Standards Measurement Study Working Paper No. 135.

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

Gross, David and Nicholas Souleles (2002). "An Empirical Analysis of Personal Bankruptcy and Delinquency." *The Review of Financial Studies* 15(1): 319-347.

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

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

Karlan, Dean and Jonathan Zinman (2007). "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment." Working Paper.

Karlan, Dean and Jonathan Zinman (2008). "Lying About Borrowing." *Journal of the European Economic Association Papers and Proceedings* 6(2-3) April-May.

Karlan, Dean and Jonathan Zinman (forthcoming). "Credit Elasticities in Less Developed Economies: Implications for Microfinance." *American Economic Review*

Kling, Jeffrey, Jeffrey Liebman and Lawrence Katz (2007). "Experimental Analysis of Neighborhood Effects." *Econometrica* 75(1): 83-120. January.

Laibson, David, Andrea Repetto and Jeremy Tobacman (2005). "Estimating Discount Functions with Consumption Choices Over the Lifecycle." Working Paper. August 11, 2005.

Levine, Ross (1997). "Financial Development and Economic Growth: Views and Agenda." *Journal of Economic Literature* XXXV: 688-726.

McKernan, Signe-Mary (2002). "The Impact of Microcredit Programs on Self-Employment Profits: Do Noncredit Program Aspects Matter?" *Review of Economics and Statistics* 84(1): 93-115. February.

Melzer, Brian (2007). "The Real Costs of Credit Access: Evidence from the Payday Lending Market." Working Paper.

Menon, Nidhiya (2003). "Consumption Smoothing in Micro-credit Programs." *Review of Economics and Statistics* forthcoming

Morduch, Jonathan (1998). Does Microfinance Really Help the Poor? New Evidence on Flagship Programs in Bangladesh., MacArthur Foundation project on inequality. Princeton University, draft.

Morduch, Jonathan (1999). "The Microfinance Promise." *Journal of Economic Literature* 36: 1569-1614. December.

Morduch, Jonathan (2000). "The Microfinance Schism." *World Development* 28(4): 617-629.

Morgan, Donald P. and Michael R. Strain (2007). "Debt Problems and Payday Credit: What's the Connection?" Working Paper. October.

Morse, Adair (2007). "Payday Lenders: Heroes or Villains?" Working Paper. January.

Pitt, Mark and Shahidur 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. October.

Pitt, Mark, Shahidur Khandker and Jennifer Cartwright (2003). "Does Micro-Credit Empower Women? Evidence from Bangladesh." World Bank Policy Research Working Paper. March.

Pitt, Mark, Shahidur Khandker, Omar Haider Chowdhury and Daniel Millimet (2003). "Credit Programs for the Poor and the Health Status of Children in Rural Bangladesh." *International Economic Review* 44(1): 87-118. February.

Porteous, David (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, Marguerite (2001). <u>The Microfinance Revolution: Sustainable Finance for the Poor.</u> 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)

Skiba, Paige and Jeremy Tobacman (2007). "Measuring the Individual-Level Effects of Access to Credit: Evidence from Payday Loans." Working Paper. July 3.

Stango, Victor and Jonathan Zinman (2007a). "Fuzzy Math and Household Finance: Theory and Evidence." Working Paper. November.

Stango, Victor and Jonathan Zinman (2007b). "Fuzzy Math, Disclosure Regulation, and Credit Market Outcomes." Working Paper. November.

Table 1. Demographics

	Sample frame (in experiment, and surveyed)		Applicants with a 25% chance of approval		Applicants with a 50% chance of approval		South Africa	Blacks in South Africa
	Mean	Median	Mean	Median	Mean	Median		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
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)	-
Coloured	25.4%	-	20.0%	-	27.1%	-	8.8% (f)	-
Monthly household income	R 4,359	R 2,153	R 3,348	R 1,713	R 4,646	R 2,200	R 3,750 (c)	R 2,167 (c)
Average individual monthly salary in the formal sector, 2004							R 6,882 (b)	

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 "Coloured" individuals because Capetown branches participated in the experiment.
Average exchange rate during project and survey: 1 US$ = 6.3 Rands.

Notes on monthly household income

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.

Lettered notes:

(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. Experiment Validity and Compliance

Panel A. Orthogonality of treatment to applicant characteristics

Dependent Variable:	*1= Loan Assigned*	*1 = Loan Obtained*	*1 = Surveyed*	*1= Loan Assigned*
sample:	frame	frame	frame	surveyed=1
Mean*(dependent variable):*	0.41	0.23	0.80	0.41
	(1)	(2)	(3)	(4)
Female	0.022	0.039		0.004
	(0.036)	(0.031)		(0.041)
Marital status -- Divorced	0.056	-0.006		0.079
	(0.129)	(0.099)		(0.154)
Marital status -- Married	0.036	0.053		0.023
	(0.045)	(0.039)		(0.051)
Marital status -- Separated	-0.194	0.021		-0.175
	(0.158)	(0.159)		(0.174)
Marital status -- Widow	0.104	0.136		0.010
	(0.118)	(0.111)		(0.131)
Number of dependents	0.000	0.012		0.005
	(0.013)	(0.011)		(0.015)
Non-african race	-0.035	-0.049		-0.053
	(0.040)	(0.034)		(0.044)
Age of applicant	-0.003	-0.004**		-0.002
	(0.002)	(0.002)		(0.002)
Monthly gross income at application (000s)	0.008	0.018**		0.008
	(0.008)	(0.007)		(0.010)
# years at employer	0.005	0.003		0.005
	(0.004)	(0.004)		(0.005)
ITT			-0.006	
			(0.029)	
Observations	786	786	787	625

Huber-White standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Sample contains 787 marginal applicants eligible for the treatment (i.e., for loan approval). Each column reports marginal effects for a single regression 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). Running probits produces qualitatively similar results. Non-african races include white, indian, coloured, and indian/coloured. 'Single' is the omitted marital status category. One observation is dropped from columns (1) and (2) due to missing race.

Panel B. Compliance with treatment assignment

		Full sample		50% treatment probability		25% treatment probability	
Randomizer Says To	Branch Manager Action	Frequency	Proportion Compliance	Frequency	Proportion Compliance	Frequency	Proportion Compliance
Reject	Reject	455		321		134	
Reject	Approve	7	0.98	6	0.98	1	0.99
Approve	Approve	172		144		28	
Approve	Reject	153	0.53	136	0.51	17	0.62

Table 3. Intention-to-Treat Effects on Borrowing and Access

Panel A. Effects on Borrowing and Composition

		Mean depvar for full sample	Full sample	Gender: Female	Gender: Male	Income: High	Income: Low	Credit score: High	Credit score: 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)		618-622	618-622	305-309	309-315	307-311	307-311	279-282	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

Panel B. Effects on Perceptions

	Mean depvar for full sample	Full sample	Gender: Female	Gender: Male	Income: High	Income: Low	Credit score: High	Credit score: Low
Respondent would borrow from microlender if needed a loan	0.201	0.128***	0.142***	0.106*	0.058	0.219***	0.098*	0.155***
		(0.037)	(0.048)	(0.058)	(0.049)	(0.059)	(0.054)	(0.053)
Respondent would borrow from other formal sources (excluding microlenders) if needed a loan	0.535	-0.010	-0.044	0.013	0.016	-0.062	-0.020	-0.001
		(0.045)	(0.062)	(0.067)	(0.062)	(0.064)	(0.067)	(0.062)
Respondent would borrow from informal sources if needed a loan	0.232	-0.112***	-0.099*	-0.105**	-0.082*	-0.132**	-0.083	-0.134***
		(0.036)	(0.053)	(0.048)	(0.046)	(0.057)	(0.054)	(0.048)
Respondent would be able to borrow from friends or family if needed	0.724	-0.062	-0.070	-0.056	-0.155***	0.056	-0.132**	0.005
		(0.040)	(0.056)	(0.059)	(0.056)	(0.059)	(0.060)	(0.054)
Number of observations (range)	538-539	538-539	277-279	260-261	262-268	271-276	244-248	291-294

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. All regressions include controls for: month of application with the Lender, month of survey, and treatment assignment probability. Running probits for the binary outcomes produces qualitatively similar results. The number of observations varies depending on missing values in the survey data. Perception questions were only asked in the 553 cases where the treated applicant could be found (in 73 other case a household member was surveyed). 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 4. 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%	112.1%
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 5. Intention-to-Treat Estimates for Index Components

	Mean depvar for full	Full sample	Gender		Income		Credit score	
			Female	Male	High	Low	High	Low
Consumption Index								
Dummy=1 if household did not experience hunger in past 30 days	0.861	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)
Dummy=1 if quality of food improved over the last 12 months	0.263	0.037	-0.041	0.108**	0.073	0.015	-0.006	0.078
		(0.037)	(0.051)	(0.054)	(0.054)	(0.051)	(0.053)	(0.052)
Number of observations	620-626	620-626	306-311	314-315	310-314	310-312	280-283	340-343
Economic self-sufficiency Index								
Dummy=1 if the borrower is employed	0.804	0.108***	0.107**	0.095**	0.108***	0.086	0.090*	0.103**
		(0.032)	(0.047)	(0.045)	(0.036)	(0.056)	(0.049)	(0.044)
Percentile of household employment earnings since application	45.296	5.008*	3.264	5.666	4.472	3.716	4.543	4.934
		(2.609)	(3.555)	(3.953)	(3.516)	(3.330)	(3.962)	(3.503)
Dummy=1 if the household is above the poverty line	0.606	0.074*	0.093	0.049	0.063	0.061	0.056	0.075
		(0.040)	(0.057)	(0.058)	(0.049)	(0.062)	(0.060)	(0.056)
Number of observations (range)	587-620	587-620	293-307	294-314	298-310	289-310	270-279	317-341
Investment/durables Index								
Dummy=1 if anybody in household is a university student	0.153	-0.011	0.008	-0.044	0.002	-0.035	0.035	-0.060
		(0.037)	(0.052)	(0.050)	(0.058)	(0.046)	(0.059)	(0.048)
Dummy=1 if household bought or improved dwelling since application	0.316	0.040	0.050	0.018	0.087	0.001	0.059	0.037
		(0.039)	(0.055)	(0.057)	(0.055)	(0.056)	(0.056)	(0.055)
Dummy=1 if anybody in the 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)	391-626	391-626	208-311	183-315	189-314	202-312	175-283	216-343
Control and outlook Index								
Decision-making scale	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)
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)
Position on community socio-economic ladder	4.403	0.065	-0.061	0.096	-0.225	0.299	-0.103	0.165
		(0.182)	(0.265)	(0.264)	(0.272)	(0.219)	(0.286)	(0.239)
Number of observations (range)	178-551	178-551	83-285	95-266	116-269	62-282	97-254	81-297
Physical health Index								
Dummy=1 if general health of the borrower is "very good"	0.526	0.047	0.056	0.023	0.046	0.023	-0.005	0.084
		(0.042)	(0.059)	(0.062)	(0.057)	(0.063)	(0.062)	(0.058)
Dummy=1 if no household member was sick in previous 30 days	0.517	-0.026	-0.037	-0.000	-0.018	-0.025	0.085	-0.102*
		(0.042)	(0.059)	(0.061)	(0.058)	(0.061)	(0.061)	(0.058)
Number of observations (range)	610-625	610-625	308-311	302-314	307-313	303-312	277-283	333-342
Mental health Index								
Lack of 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)
Lack of 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)
Number of observations (range)	244-250	244-250	127-133	117	120-122	124-128	112-117	132-133

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. All regressions include controls for: month of application with the Lender, month of survey, and treatment assignment probability. Running probits for the binary outcomes produces qualitatively similar results. Each outcome is scaled so that a higher number represent a better outcome. The Data Appendix provides details on outcome measurement. 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. Intention-to-Treat Estimates for Summary Index Outcome Measures

	Full sample	Gender		Income		Credit score	
		Female	Male	High	Low	High	Low
Consumption Index	0.117**	-0.023	0.232***	0.132	0.094	0.000	0.210***
	(0.058)	(0.082)	(0.083)	(0.081)	(0.085)	(0.087)	(0.080)
Economic self-sufficiency Index	0.190***	0.188**	0.172**	0.175**	0.157*	0.157*	0.188**
	(0.060)	(0.087)	(0.087)	(0.071)	(0.090)	(0.092)	(0.082)
Investment/Durables Index	0.062	0.050	0.041	0.041	0.061	0.095	0.029
	(0.053)	(0.075)	(0.074)	(0.077)	(0.074)	(0.080)	(0.073)
Control and outlook Index	0.172***	0.159**	0.196***	0.098	0.241***	0.110	0.208***
	(0.048)	(0.068)	(0.069)	(0.068)	(0.067)	(0.079)	(0.061)
Physical health Index	0.022	0.018	0.020	0.029	-0.002	0.081	-0.017
	(0.060)	(0.082)	(0.092)	(0.085)	(0.086)	(0.089)	(0.084)
Mental health Index	-0.152*	-0.229**	-0.099	-0.181*	-0.136	-0.105	-0.202*
	(0.079)	(0.112)	(0.114)	(0.108)	(0.117)	(0.115)	(0.109)
Overall Index	0.069**	0.027	0.094**	0.049	0.069*	0.056	0.069*
	(0.030)	(0.040)	(0.044)	(0.040)	(0.041)	(0.045)	(0.041)
Number of observations	626	311	315	314	312	283	343

Huber-White standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Results obtained using OLS to estimate the ITT model detailed in equation (2); each cell presents the estimated treatment effect from a single regression. All regressions include controls for: month of application with the Lender, month of survey, and treatment assignment probability. Indices are created by adding related outcome measures together (after imputing missing values and standardizing as detailed in Section III-E), and taking their unweighted average. The outcome measures contained in each index are listed in Table 5; e.g., the first few rows of Table 5 show that the economic self-sufficiency index is comprised of employment status, employment earnings percentile, and the poverty line variable. 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. See the Data Appendix for more details on the construction of the indices.

Table 7. Treatment Effects on Credit Bureau Scores One and Two Years Later

Panel A. Results on the Surveyed Sample

Dependent variable:	*1= any ordinal score in Dec 2005,* one year impact	*1= any ordinal score in Dec 2006,* two year impact	*score Dec 2005,* one year impact	*score Dec 2006,* two year impact
	(1)	(2)	(3)	(4)
Intent to treat	0.076***	0.067***	-1.097	-1.537
	(0.026)	(0.023)	(5.163)	(5.166)
r-Squared	0.062	0.051	0.021	0.015
mean(dependent variable)	0.88	0.90	629	635
N	626	626	547	561

* p<0.10, ** p<0.05, *** p<0.01. OLS with Huber-White standard errors. All models include controls for randomization probability and month of application. Applicants with a thin credit history do not have an ordinal score: they have no score at all, or a 3-category risk indicator. December 2005 is 13-15 months after the treatment (i.e., after the date of application for those in the experiment). December 2006 is 25-27 months after the treatment.

Panel B. Results on the Entire Sample Frame

Dependent variable:	*1= any ordinal score in Dec 2005,* one year impact	*1= any ordinal score in Dec 2006,* two year impact	*score Dec 2005,* one year impact	*score Dec 2006,* two year impact
	(1)	(2)	(3)	(4)
Intent to treat	0.067***	0.059***	1.456	0.660
	(0.023)	(0.022)	(4.582)	(4.790)
r-Squared	0.061	0.051	0.037	0.015
mean(dependent variable)	0.87	0.88	629	636
N	787	787	682	693

Sample includes everyone who got a treatment assignment and hence who we attempted to survey.

Table 8. Estimated Profitability of Marginal and Inframarginal Loans

		All first loans		Low Credit Score		High Credit Score
Panel A. Marginal Loans		(1)		(2)		(3)
Count		172		85		87
Proportion paid in full by May 2005		0.715		0.753		0.678
NPV of payments made from marginal borrowers	R	221,315.01	R	104,126.21	R	117,188.80
NPV of amount lent to marginal borrowers	R	175,581.39	R	81,893.65	R	93,687.74
NPV of profits, assuming no marginal staff costs	R	45,733.62	R	22,232.56	R	23,501.06
NPV of profits per marginal loan, assuming no marginal staff costs	R	265.89	R	261.56	R	270.13
		(48.09)		(73.95)		(62.21)
NPV of profits, with shadow cost of staff time	R	34,643.62	R	16,739.56	R	17,904.06
NPV of profits per loan, with marginal staff cost	R	201.42	R	196.94	R	205.79
		(48.55)		(74.58)		(62.91)
Panel B. Inframarginal Loans						
Count		1,405		295		1,110
Proportion paid in full by May 2005		0.764		0.692		0.783
NPV of payments made from marginal borrowers	R	2,252,494.30	R	351,566.65	R	1,900,927.70
NPV of amount lent to marginal borrowers	R	1,768,566.20	R	289,515.58	R	1,479,050.60
NPV of profits, assuming no marginal staff costs	R	483,928.10	R	62,051.07	R	421,877.10
NPV of profits per loan, assuming no marginal staff costs	R	344.43	R	210.34	R	380.07
		(21.52)		(32.32)		(25.75)
NPV of profits, with shadow cost of staff time	R	399,181.07	R	43,376.07	R	355,805.01
Profit per inframarginal loan	R	284.11	R	147.04	R	320.55
		(21.67)		(32.76)		(25.9)
Inframarginal Loan - Marginal Loan Profit Difference	R	82.70	R	-49.90	R	114.75
P-value of t-test that profit difference between marginal and inframarginal loan ~=0		0.20		0.49		0.22

Standard errors in parentheses. All loans counted here were to first-time borrowers from the Lender and originated at the 8 experimental branches during our study period: September 21, 2004-November 20, 2004. Marginal loans are those that loan officers originally rejected but decided to approve after our randomized second look. Inframarginal loans are those to first-time borrowers that loan officers initially approved. Average exchange rate during project and survey: 1 US$ = 6.3 Rands. Payments include principal, interest, and late fees. Payments and amount lent discounted to experiment start date using 91-day South African Treasuries, which had an annual yield of 7.2% during our study period. The discount rates at which point estimate on the marginal loans turns unprofitable are 119% (assuming no marginal staff costs) and 87% (assuming costs as detailed below). We assume no payments made after May 20th, 2005 (our data end date), since here we are counting only the first loans made to these borrowers, and those first loans that were not repaid by May 2005 were nearly all seriously delinquent. We do not attempt to adjust profits downward for risk, and note simply that the gap between marginal and infra
The shadow cost of staff time adjusts for the possibility that time spent processing, monitoring, or enforcing any given loan reduces the amount of time spent on productive activities on other loans. This is not necessarily a fair assumption, since there appeared to be nontrivial slack (as evidenced by the fact that the Lender was able to implement this experiment without adding staff). Shadow costs are estimated as follows: (a) processing approved loans: 0.5 hours*R75/hour, (b) monitoring loans: 0.5 hours*R29/hour, (c) enforcement re: delinquent loans: 1 hour*R29/hour, for any loan that goes into default (>= 3 months past due).