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

Dean Karlan
Innovations for Poverty Action, Yale University

Jonathan Zinman
Innovations for Poverty Action, Dartmouth College

Expanding access to commercial credit is a key ingredient of financial development strategies. There is less consensus on whether expanding access to *consumer* credit helps borrowers, particularly when loans are extended at high interest rates. Popular skepticism about "unproductive," "usurious" lending is fueled by research highlighting behavioral biases that may induce overborrowing. We estimate the impacts of expanding access to consumer credit at a 200% annual percentage rate (APR) using a field experiment and follow-up data collection. The randomly assigned marginal loans produced significant net benefits for borrowers across a wide range of outcomes. There is also some evidence that the loans were profitable. (*JEL* D12, D14, G21, I32)

Introduction

Expanding access to credit is a key ingredient of financial development strategies worldwide. The small business administration and comparable small- and medium-enterprise (SME) initiatives target billions of dollars of commercial credit in developed economies. The microcredit industry targets billions of dollars of commercial credit in developing economies. A widely shared presumption of these efforts is that expanding access to "productive" credit makes entrepreneurs and small business owners (weakly) better off.

There is less consensus on whether expanding access to *consumer* credit does borrowers more good than harm. 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

Thanks to Jonathan Bauchet, Luke Crowley, Nathanael Goldberg, Ben Pugsley, Bram Thuysbaert, and Nouhoum Traore for excellent research assistance; to Lia Fernald for advice on measures of mental health; to a referee, the editor Laura Starks; and to Sumit Agarwal, Michael Anderson, Abhijit Banerjee, Greg Fischer, 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 article are not necessarily shared by any of the funders, the Lender, or the Federal Reserve System. Send correspondence to Jonathan Zinman, Department of Economics, Dartmouth College, HB 6106, Hanover, NH 03755; telephone: (603) 667-5068. E-mail: jzinman@dartmouth.edu.

doi:10.1093/rfs/hhp092 Advance Access publication November 28, 2009

consumers to overborrow.[1] Indeed, practitioners and policymakers have long been skeptical about "unproductive" lending at "usurious" rates in "subprime" markets.[2]

A growing empirical literature on the effects of access to expensive credit on U.S. borrowers has added fuel to this debate. Several studies find evidence suggesting that access to payday loans has negative effects on household financial condition and well-being (Campbell, Martinez-Jerez, and Tufano 2008; Carrell and Zinman 2008; Melzer 2009; Skiba and Tobacman 2008a). But several other studies find evidence suggesting that payday loan access has positive effects on households (Morgan and Strain 2008; Morse 2009; Wilson et al. 2008; Zinman forthcoming).

We estimate the effects of expanding access to expensive consumer credit using variation produced by a field experiment in which a single finance company (the "Lender") randomly liberalized screening criteria on four-month, unsecured, 200% APR consumer installment loans in South Africa.[3] We then measure borrower outcomes using a follow-up household survey and administrative data on applicant creditworthiness over time. The key questions are threefold. First, do credit constraints bind?[4] If they do not (e.g., if all rejected marginal applicants can obtain comparable credit from other lenders), then one would not expect to find any impacts of expanding access from one particular lender. Second, does relaxing any credit constraints benefit marginal borrowers? Third, are the marginal loans profitable? The Lender agreed to implement our experiment because its senior management believed that branch staff applied inefficiently strict underwriting criteria; indeed, the literature on retail credit risk assessment suggests that there was little reason to expect that the Lender's risk assessment methods would be fully optimized ex ante.[5]

Our field experiment approach is motivated by the deep endogeneity issues that complicate identification of the causal impacts of credit expansion

[1] Behavioral biases may produce borrowing that is excessive relative to a normative benchmark. For example, Laibson, Repetto, and Tobacman (forthcoming) find that consumers with present-biased preferences would commit $2000 to not borrow on credit cards; Ausubel (1991) argues that overoptimism produces excess credit card borrowing; Stango and Zinman (2009; forthcoming) find that consumers tend to underestimate the interest rate on short-term loans and borrow more expensively and heavily as a result.

[2] On practice, see, e.g., Robinson (2001). On policy, South Africa, the location of our study, offers an interesting case. South Africa deregulated usury ceilings in 1992 to encourage the development of formal markets in consumer credit. However, recent legislation reimposed some ceilings, effective in 2007. An example from the United States is the substantial variation, both across states and within states across time, in payday lending restrictions (Carrell and Zinman 2008).

[3] The Lender competed in a "cash loan" market similar to the U.S. payday loan market (see Section 2 for details). The product offered in the experiment was the Lender's standard one for first-time applicants.

[4] Prior evidence suggested strongly that credit constraints would 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 2008a); as Attanasio, Goldberg, and Kyriazidou (2008) show formally, this pattern of elasticities is further evidence of unmet demand for credit.

[5] 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.

on borrower and lender outcomes.[6] 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 nonexperimental studies without strong assumptions. A classic example concerns relatively “spunky” individuals selecting or being selected into loan contracts 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 or policymakers) tend to take (target market or deregulate) expensive credit 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.[7]

We addressed the identification problem by working with the Lender to engineer exogenous variation in the loan approval process. Our design randomly encouraged branch staff (loan officers and branch managers) to approve some marginal applications. Specifically, the Lender added three additional steps to its normal process for new loan applicants. First, branch staff was required to label rejected applications as either egregiously uncreditworthy or marginally uncreditworthy. Second, the branch computer then instructed staff to take a second look at some marginal rejected applications in real time, by randomly assigning marginal applicants to treatment (“approve”) or control (still “reject”).[8] In the third and final step, branch staff were encouraged but not required to follow the treatment assignment. Branches actually made loans to 53% of the marginal rejected applicants randomly assigned to treatment (i.e., to be approved).

Granting branch staff discretion in the final step gives the experimental design ecological validity by permitting estimation of treatment effects on those applicants who are closest to a real-world margin of creditworthiness.[9] Retail lenders often liberalize their screening criteria[10] through a mechanism much

6 Field experiments conducted collaboratively with lenders are a new but growing approach to studying credit markets. See, e.g., Ausubel (1999) and Karlan and Zinman (2008a).

7 See above for several comparable impact studies from the U.S. payday loan market. Studies in developing countries have focused on access to microentrepreneurial credit rather than consumer credit; see, for example, Coleman (1999); Kaboski and Townsend (2005); McKernan (2002); Morduch (1998); Pitt et al. (2003); and Pitt and Khandker (1998). There may be little economic distinction between small, closely held businesses and the households that run them, and there is some evidence that microentrepreneurial loans are often used for household expenses (Banerjee et al. 2009; Karlan and Zinman 2009; Morduch 1998).

8 The Lender did not inform applicants that a component of the credit decision was randomized.

9 The downside of branch staff discretion is that it sacrifices statistical power, since one must use the treatment *assignment*–to “approve” or still “reject”–to identify exogenous variation in credit access. See Section 2.4 for details.

10 Liberalization of screening criteria is used by retail lenders seeking market share and/or untapped profits, by mission-oriented microlenders that expand “outreach” while holding their physical capital and risk assessment technology constant, by directed lending programs (Banerjee and Duflo 2008), and by semidirected lending programs (e.g., the Community Reinvestment Act in the United States).

like the one used in our experiment: senior management encourages front-line personnel to make riskier loans,[11] but then leaves the ultimate credit decisions in the hands of a loan officer, branch manager, or credit committee. Our design permits estimation of both intention-to-treat (ITT) and treatment-on-the-treated (TOT) effects of liberalizing screening criteria by using the applicant treatment *assignment* (to "approve" or still "reject") to generate the required exogenous variation in credit access. Regressing an outcome on the treatment assignment delivers the ITT effect; i.e., the average effect of liberalizing screening criteria on the population of viable marginal applicants (those who pass the initial filter of not being egregiously uncreditworthy). Then one can obtain the TOT effect by roughly doubling the ITT effect (since roughly half of the experimental pool assigned "approve" actually were approved). These TOT estimates measure impacts on a market- and hence policy-relevant sample of interest: those applicants deemed by branch staff to be closest to the creditworthiness bar. But our estimates do not measure impacts on applicants who are farther below or well above the bar.

Our results corroborate the presence of binding liquidity constraints. Treated applicants borrowed more overall than controls in the six to twelve months following the experiment and also shifted their borrowing composition toward loans like the Lender's. This opens up the possibility that our Lender's marginal loans affected borrower real activity and well-being.

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 household survey variables that capture economic 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 2008; Kling, Liebman, and Katz 2007).

We find that expanded access to credit significantly improved overall borrower outcomes. Economic self-sufficiency (employment and income), food consumption, and one measure of subjective well-being (an index of intra-household control, community status, overall optimism) were each higher for treated applicants than for controls six to twelve months after the treatment. We find a negative effect on another measure of subjective well-being (an index

[11] Other "encouragement" strategies besides the type used in our experiment (a credit scoring model-based recommendation) include monetary incentives for front-line personnel and goals (e.g., for portfolio growth).

of depression and stress). But the average treatment effect across all of our economic and subjective outcomes is significant and positive. Over fifteen- to twenty-seven-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 borrowing costs.

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.[12]

We also find that the marginal loans were profitable for the Lender (although substantially less profitable than inframarginal loans). Exactly how profitable depends on several assumptions about marginal costs and risk weighting.[13]

In all, our results suggest a role for welfare-improving interventions in consumer credit markets but come with important caveats. We measure some of the outcomes of interest only at six- to twelve-month horizons, and some costs and/or benefits of liberalized access to credit may materialize only over longer horizons. Our screening liberalization design does not allow us to measure impacts on some applicant segments of interest: those who "normally" would have been approved, and those who are well below the bar of creditworthiness. Nor do we 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 our treatment effects to markets with different competitive structures or consumers with different characteristics is uncertain.

While we can only speculate about the implications of our findings for the optimal (de)regulation of microcredit and subprime credit in other settings, the findings are noteworthy nonetheless because they will be so surprising to the many policymakers and microfinance practitioners who operate under the strong presumption that expanded access to expensive credit does consumers more harm than good. We find evidence to the contrary in a setting where the likelihood of "productive" lending seemed slim ex ante: our Lender was for-profit (with no social mission), targeted employed individuals rather than entrepreneurs, and did not necessarily disclose an APR (nor was it required to do so by law). Hence, we hope our findings will shift some programmatic and policy debates a bit toward the agnosticism that seems warranted in light of the available evidence.

12 As noted at the outset, this stands in contrast to microenterprise credit, which is often assumed to be beneficial, and in fact sometimes subsidized with aims of poverty alleviation.

13 We cannot simply apply the market test of whether more aggressive underwriting criteria were 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 postmerger data or underwriting policy information.

Our study has more clear-cut implications for the methodology of accumulating the additional evidence needed to optimize policy and practice. We demonstrate that randomized-control trials can be used to help identify the severity of liquidity constraints and 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 article proceeds as follows. Section 1 provides background information on the applicants, the Lender, and the cash loan market. Section 2 details the design and implementation of our experiment and data collection methods and empirical strategy. Section 3 presents estimates of treatment effects on borrowing and credit access. Section 4 presents estimates of treatment effects on summary index measures of ultimate outcomes of interest. It also presents our estimates of effects on credit scores fifteen to twenty-seven months after treatment, and details our estimates of Lender profits on marginal and inframarginal loans. Section 4 concludes with a discussion of external validity and other questions for future research.

1. Market and Lender Overview

Our cooperating Lender operated for over twenty years as one of the largest, most profitable microlenders in South Africa.[14] 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 nonmortgage 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 40% of the median borrower's gross monthly income.[15] Our sample for this experiment includes mostly first-time loan applicants of African descent. Table 1 shows some comparative demographics. In Section 4 we show that borrowers finance a variety of different consumption smoothing and investment activities through these loans.

Cash lenders arose to substitute for traditional "informal sector" money lenders following deregulation of the usury ceiling in 1992, and they are

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

[15] Throughout the article we convert all South Africa currency into U.S. dollars using the average exchange rate over our study period of 21 September 2004–30 November 2005: 6.31 Rand = $1.

Table 1
Demographics

	Sample frame (in experiment, and surveyed)		Applicants with a 25% chance of approval		Applicants with a 50% chance of approval			
	Mean (1)	Median (2)	Mean (3)	Median (4)	Mean (5)	Median (6)	South Africa (7)	Blacks in South Africa (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)	–
Colored	25.4%	–	20.0%	–	27.1%	–	8.8% (f)	–
Monthly household income	R4359	R2153	R3348	R1713	R4646	R2200	R3750 (c)	R2167 (c)
Average individual monthly salary in the formal sector, 2004							R6882 (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 a relatively high proportion of mixed race "colored" individuals because Capetown branches participated in the experiment.
Average exchange rate during project and survey: 1 US$ = 6.3 Rand.
Income variables are means.
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 nonagriculture formal employees, November 2004. Source: Quarterly Employment Statistics, Statistics South Africa, November 2005.
(c) In Rand 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.

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.[16]

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 automatic teller machine (ATM) cards of clients. The Lender also had a "medium-maturity" product niche in four-month installment loans. Most other cash lenders focus on one-month or twelve-month-plus loans.[17] In this experiment 98% of the borrowers received the standard loan for first-time borrowers: a four-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),[18] and the loan was then amortized into four equal monthly repayments.[19] But compared with the pricing of 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 50% of new applications for reasons such as unconfirmed employment, suspicion of fraud, poor credit rating, and excessive debt burden.

Applicants who were approved often defaulted on their loan obligation (see Section 4.3), despite facing several incentives to repay. Carrots included decreasing prices and increasing future loan sizes following good repayment behavior (demand for repeat loans is high, as evidenced by the fact that more than 50% of first-time borrowers borrowed again from the Lender). Sticks

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

[17] The Lender also had one-, six-, twelve-, and eighteen-month products, with the longer maturities offered at lower rates and restricted to the most observably creditworthy customers.

[18] See Stango and Zinman (2009) for evidence on how and why lenders prefer to shroud APRs, and on the mediating impacts of incomplete enforcement of mandated APR disclosure outcomes in the United States.

[19] So an R1000 loan had monthly repayments of $(1000 + 1000^{*}0.1175^{*}4)/4 =$ R367.50. Borrowers that prepaid paid add-on interest prorated to the time outstanding; e.g., a borrower who stayed current and prepaid her remaining amount at the end of month 2 would have repaid R367.50 in month 1, plus R867.50 at the end of month 2, for a total repayment of R1235 = R1000 (principal) + R235 (two months' interest).

included reporting to credit bureaus, frequent phone calls from collection agents, court summons, and wage garnishments.

Overall, the cash loan market has both differences and similarities with other markets for expensive credit. It clearly differs from many U.S. subprime markets in that cash loans are unsecured and typically held by originators. The cash loan target market (the working poor and lower middle class) and products look something like the U.S. payday loan market, although our Lender offers a somewhat larger (relative to income), longer-term, and cheaper version of the standard U.S payday loan.[20] Unlike many "traditional" microlenders (e.g., the Grameen Bank and other NGOs), cash lenders do not have explicit social welfare and targeting goals, and do not use group liability mechanisms. But the industrial organization of microcredit is trending toward something that looks more like the cash loan market: for-profit, more competitive delivery of untargeted, individual liability loans (Porteous 2003; Robinson 2001). This evolution is happening from both the bottom-up (nonprofits converting to for-profits) and the top-down (for-profits expanding into subprime and consumer segments).

2. 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.

2.1 Experimental design and implementation

2.1.1 Sample and time frame for the experiment. We drew our sample frame from the universe of over 3000 "new" applicants who had no prior borrowing from the Lender and applied at any of eight branches between 21 September and 20 November 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 comprised 787 "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.

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

[20] The typical payday loan is for \$300 or less, and costs "\$15 per \$100" for two weeks, or 390% APR (Stegman 2007).

information about the expected profitability of changing its underwriting in a way that induces branch personnel to approve more risky loans.

2.1.2 Experimental design and operations. The Lender implemented the experiment in three steps.

First, branch staff (loan officers and/or branch managers) evaluated each of about 3000 new applicants using the Lender's standard underwriting process and three additional steps. Under normal operations branch staff would use a combination of a credit scoring model and discretion to make a binary approve/reject decision. The experiment forced branch staff 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 branch staff approved 1405 new applications based on the standard underwriting criteria. Seven hundred and five applications were deemed egregious rejects, leaving us with a sample frame of 787 marginally rejected applicants for the experiment.

Second, special "randomizer" software encouraged branch staff to take a second look at 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 $787 + 705 = 1492$ 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 the branch staff's 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.[21] 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 take a second look" rather than "randomized approval" because loan officers and branch managers 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

[21] Thus, the branch manager had the ultimate discretion to comply or deviate from the computer's randomization.

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)
No. of years at employer	0.005	0.003		0.005
	(0.004)	(0.004)		(0.005)
1 = Loan assigned			−0.006	
			(0.029)	
Observations	786	786	787	625

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

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 comprising (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, Colored, and Indian/Colored. "Single" is the omitted marital status category. One observation is dropped from columns (1) and (2) due to missing race.

pecuniary incentives with randomizer compliance (note that we use the term "compliance" in the econometric sense, not in a layman sense, since the branch staff 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 (approved) 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 (see Sections 2.4 and 2.7 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 four-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.

2.2 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 below, the surveyors asked questions on demographics, resources, recent investments, employment status, income, consumption, and subjective well-being.[22]

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 seventy-three 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

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

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 six- to twelve-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 fifteen to twenty-seven months we obtained credit reports from a credit bureau for each of the applicants in the experiment (see Section 4.2).

2.3 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.[23]

2.4 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 an OLS specification:

$$Y_i^k = \alpha + \beta^k assignment_i + \delta\ risk_i + \phi\ appmonth_i + \gamma\ surveymonth_i + \varepsilon_i, \tag{1}$$

where Y is a behavior or outcome of interest k for applicant i (or i's household). Examples of Y include measures of borrowing, poverty status, and loan repayment.[24] $Assignment_i = 1$ if the individual was *assigned* to treatment (irrespective of whether they actually received a loan). $Risk_i$ captures the applicant's

[23] Table 5 shows that the treatment effects on credit scores are statistically identical across the surveyed and non-surveyed groups, and we discuss these results in Section 4.2.

[24] Tables 3, 4, and 6 show the results on borrowing, poverty status, and loan repayment, respectively.

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.[25]

2.5 Inference over multiple outcomes

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 are labeled statistically significant due to chance is increasing in the number of outcomes (i.e., in the number of tests performed). The second concern is evaluating the overall direction and magnitude of the treatment effects when there is a diffuse set of outcomes. Following Kling, Liebman, and Katz (2007), we address these concerns using two approaches.

The first approach is to 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, intrahousehold decision power, and self-perception of community status).

We construct indices by first rescaling each outcome Y_{ij}^k (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.[26] Then we combine outcomes in a domain j by taking the average of equally weighted standardized components. Then our summary index analog to Equation (1) is

$$Y_i^j = \alpha + \beta^j assignment_i + \delta\ risk_i + \phi\ appmonth_i + \gamma\ surveymonth_i + \varepsilon_i, \tag{2}$$

where Y_i^j is an average z-score: the average of standardized component outcomes in domain j.

[25] 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 with the treatment group (e.g., because they did not obtain credit).

[26] Following Kling, Liebman, and Katz (2007), 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 4.1 for details.

The second approach is to construct familywise *P*-values. This is a way of effectively correcting standard errors on the individual outcome (i.e., on the index component) results for multiple inference. We do this correction within each outcome domain.

2.6 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. Looking at 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. Looking at household income is interesting because there is often tension in microfinance between "sustainability" (profitability) and "outreach" (expanding credit supply) to the "poorer of the poor" (Morduch 1999; 2000). Little is known about where impacts are strongest. Treatment effects may be stronger on the relatively poor if they are relatively credit constrained. Alternatively, treatment effects may be weaker on the relatively poor if they lack complementary skills or resources. Similarly, we also split the sample by ex ante credit risk as measured by the Lender's matrix of internal and external credit scores.

2.7 Treatment-on-the-treated effects

As discussed in the Introduction, TOT effects are important. They measure the impacts on the marginal borrowers deemed most creditworthy by branch staff, 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 branch staff 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 simply by doubling the relevant ITT estimate.

2.8 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, the findings may apply only to the sample used to estimate them. Our sample is a subset of larger populations of interest: principally, those with physical access to subprime unsecured credit who are being screened out by industry criteria (or regulatory restrictions). The Conclusion section discusses some related markets and policy issues in both developing and developed countries.

The second issue relates to the mechanism we study for expanding access. Our Lender's liberalization of credit screening criteria relied ultimately on branch staff 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.

The third external validity issue relates to measuring treatment effects on medium-run outcomes. Section 2.2 details why we chose six to twelve months for survey data collection on credit access and well-being measures. We consider fifteen- and twenty-seven-month impacts on credit scores, and address the possibility of time-varying treatment effects in Section 4.2.

3. 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. While borrowing behavior is not necessarily an ultimate outcome of interest, estimating these treatment effects is an important step in the analysis because 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 look back at borrowing over the entire period elapsed since the treatment, since the effects of borrowing on ultimate outcomes of interest may have a gestation period and/or be durable. 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 six to twelve months after applying to the Lender (panel A, "all sources"). But treated households did borrow more intensively than controls: 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.[27] But together with data on investments and ultimate outcomes

[27] 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.

(Section 4), we can examine whether the changes in borrowing opportunities produced by the treatment actually benefited households.[28]

Table 3, panel A, also shows limited evidence of heterogeneous treatment effects. We find several instances where the treatment effect is significant in one subsample but not another. However, the differences across males and females, income groups, and credit score bins are not statistically significant.

The estimates in panel A are likely attenuated by systematic underreporting of borrowing. A companion paper finds that 50% of survey respondents known to have borrowed from the Lender during the twelve months preceding the survey do not report *any* borrowing in the survey (Karlan and Zinman 2008b). Consequently, this suggests that 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 4 (on economic and well-being outcomes, such as employment) are also correlated with treatment and thus attenuated (or amplified) by underreporting. Recall that we will also measure impacts on credit scores that do 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. There is little evidence of heterogeneous treatment effects.

4. Results: Loan Uses, and Ultimate Impacts

As detailed below and further in the Web Appendix Table 1, households report using loans to finance a range of investment and consumption smoothing activities in the survey.[29] These loan uses motivate estimating treatment effects on a particular set of investments and economic outcomes. We then also estimate

[28] Another limitation of our data is that they almost certainly and dramatically understate the prevalence of informal borrowing (compare with 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) overstate the positive impacts on overall borrowing; and (2) miss a negative impact on informal borrowing. See the Conclusion section for additional discussion of interactions between formal and informal credit markets.

[29] Web appendices are available at http://www.dartmouth.edu/~jzinman/Papers/ExpandingAccess_WebAppendix.pdf.

Table 3
Intention-to-treat effects on borrowing and access

		Mean (dependent variable) for the full sample	Full sample	Gender		Income		Credit score	
				Female	Male	High	Low	High	Low
Panel A: Effects on borrowing and composition									
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									
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)
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 cases 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
Intention-to-treat estimates for summary index outcome measures

		Gender		Income		Credit score	
	Full sample	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 the 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 2.5), and taking their unweighted average. The outcome measures contained in each index are listed in Web Appendix Table 2; e.g., the first few rows of that table show that the economic self-sufficiency index comprises 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.

treatment effects on various measures of subjective well-being. In each case we report the intention-to-treat estimates and 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 Web Data Appendix. Estimated treatment effects for each "component" survey outcome are reported in the Web Appendix Table 2, with unadjusted standard errors reported in column 1, and familywise adjusted *P*-values reported in column 2.

As discussed in Section 2.5, 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 4). 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.

4.1 Loan uses, and ITT results on ultimate outcomes

The most common purpose for household borrowing is paying off other debt (28.3%). 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 (e.g., 9.9% of households report using loans to buy or improve food) 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 thirty 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 past twelve months (26% reported an improvement). Table 4 shows that the "consumption" summary index for these two outcomes increased significantly by an estimated 0.12 standard deviation units for treated households (those with an applicant that was randomly assigned a second look) relative to control households (with an applicant that remained rejected). The subsamples show significant increases for males but not females, and for low but not high credit score applicants. The consumption measures were taken well after the initial loan repayments were due on the marginal loans, and hence these treatment effects are not simply picking up a very transitory spike in consumption.

The next most common purpose for household borrowing is transportation expenses (19.4%); this and other uses (e.g., health care 5.1%, clothes 3.5%) are consistent with work-related investments. These uses help motivate an "economic self-sufficiency" index that includes current employment status and two measures of income over the past year.[30] The ITT effect on this index is positive, large (0.19 standard deviation units), and highly significant (P-value = 0.002). The subgroup estimates (recall that our income split is based on income prior to the treatment) are significant across the board and suggest homogeneous treatment effects.

The positive impact of credit access on economic self-sufficiency is arguably our most important result, and consequently we explore some mechanisms behind it. Web Appendix Table 2 shows significant positive effects on all three index components. The employment treatment effect seems to operate by enabling households to *maintain* employment by smoothing or avoiding shocks that prevent them from getting to work. Everyone in our sample frame had verified employment at the time they entered the experiment. Questions on job history reveal that treated applicants were significantly less likely to report leaving a job since entering the experiment. The point estimate (−2.8 percentage points, with a standard error of 1.4 pp) is smaller than the estimated effect on employment status, but the confidence intervals do overlap. We do

[30] 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. 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. The other measure is whether total household income exceeds the poverty line.

not find a significant effect on getting a new job. We also find a positive point estimate (+ 2.1 pp, with a standard error of 2.5 pp) 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 treatment effect on employment is large (albeit with a large confidence interval that includes much smaller effects); given the 53% compliance rate with the random assignment, it implies a TOT effect of a 20 percentage point increase in the likelihood of employment. A treatment effect this large makes sense if many marginal applicants are trying to borrow precisely because they are at risk of losing their job. The reported prevalence of using loan proceeds for transportation and/or clothing is consistent with the size of the effect we find.

The reported uses also suggest estimating treatment effects on certain investments.[31] For example, 13.7% of loans are used for educational expenses.[32] Households report almost perfect attendance among compulsory school-aged children, so we focus on university attendance for households with any member between ages 18 and 26. Another frequent use of loan proceeds is housing expenses (11.5%). We also include self-employment or business activity in the investment index. 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. And 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 2008). Nevertheless, the estimated treatment effects on our "investment/durables" index—combining self-employment, housing, and university attendance—are small and insignificant.[33]

We also estimate treatment effects on various subjective measures of well-being. We start with a "control and outlook" index. One component is a measure of decision-making power.[34] Many microfinance initiatives seek to

[31] 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.

[32] 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.

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

[34] 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 intrahousehold or extrahousehold domains) for unmarried respondents.

increase the intrahousehold bargaining power of female borrowers.[35] Another component is a standard measure of optimism using a battery of questions from the psychology literature. The third component is the respondent's perception of his/her standing on a ladder of socioeconomic status in his/her community/neighborhood. The ITT effect on the index is positive and highly significant in the full sample. Higher income and higher credit score applicants are the only subgroups that did not experience a significant increase.

The next row of Table 4 shows that we do not find significant effects on an index of two measures of self-reported physical health status.

Our final domain-specific index contains two measures of current mental health status: depression[36] and stress, as measured by standard batteries of questions from the psychology literature. Here we find our first evidence of a negative treatment effect: the index drops by 0.15 standard deviation units for treated relative to control, with a *P*-value of 0.06. The estimates by subgroup suggest that there may be heterogeneous treatment effects; we find significant declines only for female, relatively high income, and relatively low credit score applicants.[37]

The final row in Table 4 shows estimated ITT 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. The point estimates are positive for each subsample but significant only for male, relatively low income, and relatively low credit score applicants.

In one sense the economic magnitude of these treatment effects is somewhat challenging to put into perspective, given the lack of randomized and outcome-standardized evaluation of microcredit.[38] But in another sense the

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

[36] 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.

[37] Fernald et al. (2008) explore this result in detail. 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 suggest two other potential channels. One is that increased decision-making power may produce conflict. We asked several questions on intrahousehold 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.

[38] In contrast, 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. (2007) was 0.04 standard deviation units on adults (with effects two to three times as large on youths). The closest study to ours is Ashraf, Karlan, and Yin (forthcoming), 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.

magnitude matters less than the conclusion that we can rule out negative summary treatment effects over the horizon considered in our survey data (six to twelve months). For, as discussed at the outset, the default policy approach to consumer credit is to restrict rather than subsidize access.

4.2 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 six- to twelve-month horizon is too short to capture the full cost of loan repayment in some cases. Similarly, returns to some investments that are financed with the marginal loans may not be fully realized over six to twelve months. Indeed, some debt trap models imply that marginal borrowing may actually be *counter* productive in the long run— i.e., treated applicants may have worse outcomes than untreated applicants over longer horizons.[39] 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: 31 December 2005 (thirteen to fifteen months after the initial application), and 31 December 2006 (twenty-five to twenty-seven 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 scores are all measured within nine 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 in the longer run (by inducing defaults).

Credit scores are used by consumer lenders in South Africa much as they are in the United States. Scores can range from 300 to over 850. Our sample had December 2005 and 2006 scores ranging 487 to 817.[40] Our Lender made

[39] Debt traps refer to a dynamic where borrowers are unable to fully service debt out of 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.

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

Table 5
Treatment effects on credit bureau scores one and two years later

Dependent variable:	1 = any ordinal score in December 2005 One-year impact (1)	1 = any ordinal score in December 2006 Two-year impact (2)	Score December 2005 One-year impact (3)	Score December 2006 Two-year impact (4)
Panel A. Results on the surveyed sample				
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
Panel B. Results on the entire sample frame				
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

Panel A: $^{*}P < 0.10$, $^{**}P < 0.05$, $^{***}P < 0.01$. 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 three-category risk indicator. December 2005 is thirteen to fifteen months after the treatment (i.e., after the date of application for those in the experiment). December 2006 is twenty-five to twenty-seven months after the treatment.
Panel B: Sample includes everyone who got a treatment assignment and hence who we attempted to survey.

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 twenty- to thirty-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 three-category risk indicator by an 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 5 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 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. In contrast, 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 two-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.

4.3 Impacts on the Lender: Profitability

As noted at the outset, the Lender implemented this experiment based on the prior assumption 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 6 reports our profit estimates for the 172 marginal loans that branch staff originally rejected but decided to approve after our randomized second look (panel A), and for the 1405 inframarginal loans to first-time borrowers that staff 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 ninety-one days, with an annual yield of 7.2%, during our study period). Our repayment data end 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 20 May 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 of processing, monitoring, and enforcement if marginal loans reduce the amount of staff time allocated to the same activities on

Table 6
Estimated profitability of marginal and inframarginal loans

	All first loans (1)	Low credit score (2)	High credit score (3)
Panel A: Marginal loans			
Count	172	85	87
Proportion paid in full by May 2005	0.715	0.753	0.678
NPV of payments made from marginal borrowers	R221,315.01	R104,126.21	R117,188.80
NPV of amount lent to marginal borrowers	R175,581.39	R81,893.65	R93,687.74
NPV of profits, assuming no marginal staff costs	R45,733.62	R22,232.56	R23,501.06
NPV of profits per marginal loan, assuming no marginal staff costs	R265.89 (48.09)	R261.56 (73.95)	R270.13 (62.21)
NPV of profits, with shadow cost of staff time	R34,643.62	R16,739.56	R17,904.06
NPV of profits per marginal loan, with marginal staff cost	R201.42 (48.55)	R196.94 (74.58)	R205.79 (62.91)
Panel B: Inframarginal loans			
Count	1405	295	1110
Proportion paid in full by May 2005	0.764	0.692	0.783
NPV of payments made from marginal borrowers	R2,252,494.30	R351,566.65	R1,900,927.70
NPV of amount lent to marginal borrowers	R1,768,566.20	R289,515.58	R1,479,050.60
NPV of profits, assuming no marginal staff costs	R483,928.10	R62,051.07	R421,877.10
NPV of profits per inframarginal loan, assuming no marginal staff costs	R344.43 (21.52)	R210.34 (32.32)	R380.07 (25.75)
NPV of profits, with shadow cost of staff time	R399,181.07	R43,376.07	R355,805.01
NPV of profits per inframarginal loan, with marginal staff cost	R284.11 (21.67)	R147.04 (32.76)	R320.55 (25.9)
Inframarginal loan – Marginal loan profit difference	R82.70	R−49.90	R114.75
P-value of *t*-test that profit difference between marginal and inframarginal loan $\sim = 0$	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 eight experimental branches during our study period: 21 September 2004–20 November 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 Rand. Payments include principal, interest, and late fees. Payments and amount lent discounted to experiment start date using ninety-one-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 20 May 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 inframarginal profits in column 1 would be larger if we did adjust for risk.

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).

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 6 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 unprofitable in risk-unadjusted terms (and to 119% if we assume no marginal staff costs).

In principle, of course a firm cares about the present value of all expected future transactions with the marginal *client*, and the conclusion that the marginal loans were profitable would likely be strengthened if we had more complete data on future loans. Typically, the average profitability of the Lender's repeat 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 suggest that marginal clients followed the typical pattern, with prevalent and relatively profitable repeat borrowing,[41] although since the data are truncated at May 2005 we cannot "close the books" on repayment of repeat loans.

In all, the evidence suggests that the marginal loans induced by our experiment were profitable, although substantially less profitable than comparable inframarginal loans.

However, 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 6 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.

Thus, 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

[41] Fifty-six percent of marginal borrowers and 61% of inframarginal borrowers in our sample borrowed again from the Lender by May 2005.

the potential bottom-line benefits of controlled experimentation with screening criteria.

5. 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.[42] Branch staff reconsidered in real time, and unbeknownst to the applicants. Half of the reconsidered applicants were approved. We then tracked the behavior and outcomes of the treatment (reconsidered) and control (still rejected) groups over the next six to twenty-seven 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 in the six to twelve 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 six to twelve months following the experiment, and find significant and positive effects on food consumption, economic self-sufficiency, and some aspects of mental health and outlook. We do 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 six- to twelve-month impacts are transitory and driven by borrowers who have yet to realize the full costs of borrowing. Over fifteen- to twenty-seven-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.

Most important, 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 United States) is to restrict access based on the presumption that vulnerable consumers overborrow in these markets. Our evidence casts doubt on this presumption: consumers who borrowed at 200% in our experiment benefited from doing so, at least relative to their outside options.

[42] 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 2 for details.

Replications and extensions will be required to determine whether our findings generalize. In particular, future work would do well to focus on the mechanisms behind the effects of expanding access to credit. This is critical for reconciling the apparent conflict between studies like ours that find positive effects of access to expensive consumer credit (see also Morgan and Strain 2008; Morse 2009; Wilson et al. 2008; Zinman forthcoming), and the studies that find negative effects (Campbell, Martinez-Jerez, and Tufano 2008; Carrell and Zinman 2008; Melzer 2009; Skiba and Tobacman 2008a). Are the differences due to methodology or to subject heterogeneity? For example, collecting additional data on preferences, cost perceptions, and informal sector borrowing would help illuminate 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.

A final point is methodological. A field experiment followed by 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 (Ausubel 1999; Karlan and Zinman forthcoming) and how targeted populations make decisions (Bertrand et al. forthcoming; Karlan and Zinman 2008a; Skiba and Tobacman 2008b). 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 to improve efficiency, and then to evaluate whether the mechanisms actually work. 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 could fund the experiments and surveys needed to generate specific and scientific guidance for policymakers, donors, and investors to learn the impact of such programs, and for practitioners to learn how to employ credit-scoring techniques to screen more effectively and profitably.

References

Allen, L., G. DeLong, and A. Saunders. 2004. Issues of Credit Risk Modeling in Retail Markets. *Journal of Banking and Finance* 28:727–52.

Anderson, M. 2008. Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association* 103:1481–95.

Ashraf, N., D. Karlan, and W. Yin. Forthcoming. Female Empowerment: Impact of a Commitment Savings Product in the Philippines. *World Development*.

Attanasio, O. P., P. K. Goldberg, and E. Kyriazidou. 2008. Credit Constraints in the Market for Consumer Durables: Evidence from Micro Data on Car Loans. *International Economic Review* 49:401–36.

Ausubel, L. M. 1991. The Failure of Competition in the Credit Card Market. *American Economic Review* 81:50–81.

Ausubel, L. M. 1999. Adverse Selection in the Credit Card Market. Working Paper, University of Maryland.

Banerjee, A., and E. Duflo. 2008. Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program. Working Paper, MIT.

Banerjee, A., E. Duflo, R. Glennerster, and C. Kinnan. 2009. The Miracle of Microfinance? Evidence from a Randomized Evaluation. Working Paper, MIT.

Bertrand, M., D. Karlan, S. Mullainathan, E. Shafir, and J. Zinman. Forthcoming. What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment. *Quarterly Journal of Economics*.

Burgess, R., and R. Pande. 2005. Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment. *American Economic Review* 95:780–95.

Campbell, D., A. Martinez-Jerez, and P. Tufano. 2008. Bouncing Out of the Banking System: An Empirical Analysis of Involuntary Bank Account Closures. Working Paper, Harvard University.

Carrell, S. E., and J. Zinman. 2008. In Harm's Way? Payday Loan Access and Military Personnel Performance. Working Paper, University of California-Davis.

Coleman, B. 1999. The Impact of Group Lending in Northeast Thailand. *Journal of Development Economics* 45:105–41.

de Mel, S., D. McKenzie, and C. Woodruff. 2008. Returns to Capital in Microenterprises: Evidence from a Field Experiment. *Quarterly Journal of Economics* 123(4):1329–72.

Deaton, A., and S. Zaidi. 1999. Guidelines for Constructing Consumption Aggregates for Welfare Analysis. Working Paper No. 135, Living Standards Measurement Study.

Department of Trade and Industry South Africa. 2003. Credit Law Review: Summary of Findings of the Technical Committee, Pretoria, South Africa.

Fernald, L., R. Hamad, D. Karlan, E. Ozer, and J. Zinman. 2008. Small Individual Loans and Mental Health: A Randomized Controlled Trial among South African Adults. *BMC Public Health* 8:409.

Gross, D., and N. Souleles. 2002. An Empirical Analysis of Personal Bankruptcy and Delinquency. *Review of Financial Studies* 15:319–47.

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

Karlan, D., and J. Zinman. 2008a. Credit Elasticities in Less Developed Economies: Implications for Microfinance. *American Economic Review* 98:1040–68.

Karlan, D., and J. Zinman. 2008b. Lying about Borrowing. *Journal of the European Economic Association Papers and Proceedings* 6:510–21.

Karlan, D., and J. Zinman. 2009. Expanding Microenterprise Credit Access: Using Randomized Supply Decisions to Estimate the Impacts in Manila. Working Paper, Yale University.

Karlan, D., and J. Zinman. Forthcoming. Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment. *Econometrica*.

Kling, J., J. Liebman, and L. Katz. 2007. Experimental Analysis of Neighborhood Effects. *Econometrica* 75:83–120.

Laibson, D., A. Repetto, and J. Tobacman. Forthcoming. Estimating Discount Functions with Consumption Choices over the Lifecycle. *American Economic Review*.

McKernan, S.-M. 2002. The Impact of Microcredit Programs on Self-Employment Profits: Do Noncredit Program Aspects Matter? *Review of Economics and Statistics* 84:93–115.

Melzer, B. 2009. The Real Costs of Credit Access: Evidence from the Payday Lending Market. Working Paper, Northwestern University.

Morduch, J. 1998. Does Microfinance Really Help the Poor? New Evidence on Flagship Programs in Bangladesh. Working Paper, Princeton University, Woodrow Wilson School of Public and International Affairs.

Morduch, J. 1999. The Microfinance Promise. *Journal of Economic Literature* 36:1569–614.

Morduch, J. 2000. The Microfinance Schism. *World Development* 28:617–29.

Morgan, D., and M. R. Strain. 2008. Payday Holiday: How Households Fare after Payday Credit Bans. Federal Reserve Bank of New York Staff Report no. 309.

Morse, A. 2009. Payday Lenders: Heroes or Villains? Working Paper, University of Chicago.

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

Pitt, M., S. Khandker, and J. Cartwright. 2003. Does Micro-Credit Empower Women? Evidence from Bangladesh. Working Paper, World Bank Policy Research.

Pitt, M., S. Khandker, O. H. Chowdhury, and D. Millimet. 2003. Credit Programs for the Poor and the Health Status of Children in Rural Bangladesh. *International Economic Review* 44:87–118.

Porteous, D. 2003. Is Cinderella Finally Coming to the Ball: SA Microfinance in Broad Perspective. Working Paper, Micro Finance Regulatory Council.

Robinson, M. 2001. *The Microfinance Revolution: Sustainable Finance for the Poor*. Washington, DC: IBRD/The World Bank.

Skiba, P., and J. Tobacman. 2008a. Do Payday Loans Cause Bankruptcy? Working Paper, Vanderbilt University.

Skiba, P., and J. Tobacman. 2008b. Payday Loans, Uncertainty, and Discounting: Explaining Patterns of Borrowing, Repayment, and Default. Working Paper, Vanderbilt University.

Stango, V., and J. Zinman. 2009. Fuzzy Math, Disclosure Regulation, and Credit Market Outcomes: Evidence from Truth in Lending Reform. Working Paper, University of California-Davis.

Stango, V., and J. Zinman. Forthcoming. Exponential Growth Bias and Household Finance. *Journal of Finance*.

Stegman, M. 2007. Payday Lending. *Journal of Economic Perspectives* 21:169–90.

Wilson, B., D. Findlay, J. J. Meehan, C. Welford, and K. Schurter. 2008. An Experimental Analysis of the Demand for Payday Loans. Working Paper, George Mason University.

Zinman, J. Forthcoming. Restricting Consumer Credit Access: Household Survey Evidence on Effects around the Oregon Rate Cap. *Journal of Banking and Finance*.