ECONOMIC GROWTH CENTER

YALE UNIVERSITY

P.O. Box 208629
New Haven, CT 06520-8269
http://www.econ.yale.edu/~egcenter/

CENTER DISCUSSION PAPER NO. 911

Observing Unobservables:
Identifying Information Asymmetries
with a Consumer Credit Field Experiment

Dean S. Karlan
Yale University and Princeton University

and

Jonathan Zinman
Federal Reserve Bank of New York

May 2005

This paper can be downloaded without charge from the Social Science Research Network electronic library at: http://ssrn.com/abstract=725563

An index to papers in the Economic Growth Center Discussion Paper Series is located at:
http://www.econ.yale.edu/~egcenter/research.htm

Observing Unobservables:
Identifying Information Asymmetries
with a Consumer Credit Field Experiment

Dean S. Karlan and Jonathan Zinman

ABSTRACT

Information asymmetries are important in theory but difficult to identify in practice. We estimate the empirical importance of adverse selection and moral hazard in a consumer credit market using a new field experiment methodology. We randomized 58,000 direct mail offers issued by a major South African lender along three dimensions: 1) the initial "offer interest rate" appearing on direct mail solicitations; 2) a "contract interest rate" equal to or less than the offer interest rate and revealed to the over 4,000 borrowers who agreed to the initial offer rate; and 3) a dynamic repayment incentive that extends preferential pricing on future loans to borrowers who remain in good standing. These three randomizations, combined with complete knowledge of the Lender's information set, permit identification of specific types of private information problems. Specifically, our setup distinguishes adverse selection from moral hazard effects on repayment, and thereby generates unique evidence on the existence and magnitudes of specific credit market failures. We find evidence of both adverse selection (among women) and moral hazard (predominantly among men), and the findings suggest that about 20% of default is due to asymmetric information problems. This helps explain the prevalence of credit constraints even in a market that specializes in financing high-risk borrowers at very high rates.

I. Introduction

Information asymmetries are often believed to cause credit market failures. Stiglitz and Weiss (1981) sparked a large literature of theoretical papers on the role of asymmetric information in credit markets; this literature has influenced economic policy and practice worldwide. These theories show that information frictions and ensuing credit market failures can produce harmful real consequences at both the micro and the macro level, via underinvestment (Gale 1990; Banerjee and Newman 1993; Hubbard 1998), overinvestment (de Meza and Webb 1987; Bernanke and Gertler 1990), or poverty traps (Mookherjee and Ray 2004).

Yet empirical evidence on the existence and importance of specific information frictions is relatively thin (Chiappori and Salanie (2003)).[1] Distinguishing between adverse selection and moral hazard is difficult even when precise data on underwriting criteria and clean variation in contract terms are available, as a single interest rate (or insurance contract) may produce independent, conflated selection and incentive effects.[2] For example, a positive correlation between default and a randomly assigned interest rate, conditional on observable risk, could be due to adverse selection *ex-ante* (those with relatively high probabilities of default will be more likely to accept a high rate) or moral hazard *ex-post* (because those given high rates have greater incentive to default).

More generally, despite widespread interest in liquidity constraints and their real effects, empirical evidence on the existence of any *specific* credit market failure is lacking. Consequently

---

[1] Note also that the 2001 Nobel Prize Committee's citation for pioneering work on asymmetric information did not cite any empirical work on credit markets, while citing six empirical papers on labor markets and four on insurance markets (Bank of Sweden 2001).

[2] See Ausubel (1999) for a related discussion of the problem of disentangling adverse selection and moral hazard in a consumer credit market. See Chiappori and Salanie (2000) and Finkelstein and McGarry (2003) for approaches to the analogous problem in insurance markets.

there is little consensus on the importance of liquidity constraints for individuals.[3] Empirical work typically has examined this issue indirectly,[4] either through accounting exercises which calculate the fixed and variable costs of lending, or by inferring credit constraints by from an agent's ability to smooth consumption and/or income (e.g., Morduch (1994)). Work studying the impact of credit market failures on the real economy tends to take some reduced-form credit constraint as given (e.g., Wasmer and Weil (2004)), or as a hypothesis to be tested (e.g., Banerjee and Duflo (2004)), without evidence of a specific friction that may (or may not) actually produce a sub-optimal allocation of credit. Our work provides microfoundations for studying the real effects of credit constraints by identifying the presence (or absence) and magnitudes of two specific credit market failures: adverse selection and moral hazard.

We test for the presence of distinct types of hidden information problems using a new experimental methodology that disentangles adverse selection from moral hazard effects on repayment. This market field experiment was implemented by a South African firm specializing in high-interest, unsecured term lending to poor workers. The experiment identifies information asymmetries by randomizing loan pricing along three dimensions: first on the interest rate offered on a direct mail solicitation, second on the actual interest rate on the loan contract, and third on the interest rate offered on future loans.[5]

A stylized example, illustrated in Figure 1, captures the heart of our methodology. Potential borrowers with the same observable risk are randomly offered a high or a low interest rate on a direct-mail solicitation. Individuals then decide whether to borrow at the solicitation's "offer" rate. Of those that respond to the high rate, half are randomly given a new lower "contract" interest rate, while the remaining half continue to receive the high rate (i.e., their contract rate equals the

---

[3] The empirical importance of credit market failures for firms is also debated; see, e.g., Hurst and Lusardi (2004) and Banerjee and Duflo (2004).
[4] See Morduch and Armendariz de Aghion (2005) for a discussion of this literature.
[5] The Lender assumed all of the revenue and repayment risk from these pricing changes. Some implementation and operational costs were shared with the authors, e.g., training and project management. Although the Lender typically employs direct mail solicitation to market to former clients, they have not in the past included price in their letters.

offer rate). Individuals do not know beforehand that the contract rate may differ from the offer rate. Any selection effect is identified by considering the sample that received the low contract rate, and comparing the repayment performance of those who responded to the *high offer* interest rate with those who responded to the *low offer* interest rate. This follows from the fact that although everyone in this hypothetical sample was randomly assigned identical contracts, they selected in at varying, randomly assigned rates, so any difference in repayment is attributable to selection on unobservables. Similarly, any effect of repayment burden (which includes moral hazard)[6] is identified by considering the sample that responded to the high offer interest rate and comparing the repayment performance of those who received the *high contract* interest rate with those who received the *low contract* interest rate. These borrowers selected in identically, but ultimately received randomly different interest rates on their contract, and any difference in default is attributable to the resulting repayment burden.

Next, after all terms (loan amount, length of loan, and interest rate) are finalized, we randomize whether the contract interest rate applies to the initial loan only or instead to all loans within a one year period, conditional on good repayment. The yearlong assignment explicitly raises the benefits of repaying the current loan on time (in the 98% of cases where the contract rate is less than the Lender's standard rate). Moreover, this "dynamic repayment incentive" does not change the costs of repaying the initial loan; initial debt burden is unperturbed. Any correlation between this incentive and default therefore must be driven by *choices*—pure moral hazard. The dynamic repayment incentive thus yields our sharpest test for the presence of moral hazard.

---

[6] We define moral hazard as the effect of repayment burden on repayment that stems from *ex-post* behavioral changes driven by the incentives of the contract. Repayment burden also includes a mechanical wealth or income effect: those with positive (negative) shocks to wealth or income will be more (less) able to repay higher interest debt. Section IV discusses this in more detail.

Our approach to estimating the extent and nature of asymmetric information is thus most similar in intent to Edelberg (2004), and in methodology to Ausubel (1999).[7] Edelberg estimates a structural model to disentangle the effects of adverse selection and one type of moral hazard (in effort) in collateralized consumer credit markets in the United States. She finds evidence consistent with both phenomena. Ausubel uses market experiments conducted by a large American credit card lender to estimate the extent and nature of adverse selection. He does not attempt to account for moral hazard separately, arguing that any such effect must be trivially small over the range of interest rates (800 basis points per annum) contracted on in his data.

We also examine whether the pattern of information asymmetries varies with the gender of the borrower, since 75% of the clients in lending programs targeting the poor in developing countries are women (Microcredit Summit Campaign 2003). Microlending institutions have targeted women for several reasons, including their relatively reliable repayment behavior. Women are better microcredit risks on average, and some of this advantage seems to persist even after conditioning on observables used in credit scoring models (Armendariz de Aghion and Morduch 2005). Most interesting to us is the widespread belief that women are less likely to engage in moral hazard of various types. Armendariz de Aghion and Morduch (2005) notes that immobility, in particular, may make default relatively costly for female borrowers and monitoring relatively cheap their lenders. The only evidence we know of that addresses gender differences in moral hazard is Kevane and Wydick (2001): they find that female borrowing groups were less likely to misuse funds than their male counterparts at a group lending institution in Guatemala. Practitioners do not seem to hold priors that adverse selection differs by gender, although if social sanctions bind more for women, producing risk aversion, it follows that women may also be less

---

[7] At least two other papers endeavor to disentangle adverse selection from moral hazard in a developing country credit market. Karlan (2004) finds evidence of social capital mitigating moral hazard effects. Klonner and Rai (2004) uses institutional features of rotating credit associations in India, and finds evidence for adverse selection. Other papers estimating the prevalence of private information in credit markets include Calem and Mester (1995), Crook (2002), Drake and Holmes (1995), and Cressy and Toivanen (2001).

likely to choose *ex-ante* risky gambles when faced with high interest rates; i.e., adverse selection would be less severe among women.

We find evidence of moral hazard among male borrowers and adverse selection among female borrowers.[8] In the full sample, these findings aggregate to robust evidence for moral hazard and weak evidence for adverse selection. The magnitudes of these information problems appear to decrease with the length of the prior lending relationship with the Lender. Where statistically significant, the effects of private information are economically important, and overall our results indicate that adverse selection and moral hazard explain about 20% of default in our sample. Information asymmetries thus help explain the prevalence of credit constraints even in a market that specializes in financing high-risk borrowers at very high rates.

The paper proceeds as follows. Section II provides background on South African consumer credit markets and our cooperating Lender. Section III lays out the experimental design and implementation. Section IV details how specific theoretical models motivate the design. Section V maps the experimental design and related theory into our empirical strategy. Section VI presents the empirical results. Section VII concludes with a brief discussion of some implications, unresolved questions, and future work.

II.   Market and Lender Overview

Our cooperating Lender competes in a "cash loan" industry segment that offers small, high-interest, short-term credit with fixed repayment schedules to a "working poor" population. Cash loan borrowers generally lack the credit history and/or collateralizable wealth needed to borrow from traditional institutional sources such as commercial banks. Cash lenders arose to

---

[8] Several other studies have found evidence of differential behavior by gender in financial markets. Male private equity investors are more likely trade excessively and consequently earn lower returns (Barber and Odean 2001). Women with time inconsistent preferences (but not men) are more likely to use a new commitment savings product (Ashraf, Karlan and Yin 2004). Access to microloans increases household consumption more when a woman is the household's program participant (Pitt and Khandker 1998).

substitute for traditional "informal sector" moneylenders following deregulation of the usury ceiling in 1992, and they are regulated by the Micro Finance Regulatory Council (MFRC). Aggregate outstanding loans equal 38% of non-mortgage consumer credit (Department of Trade and Industry South Africa 2003).[9]

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 Lender's median loan size of R1000 ($150) is 32% of its median borrower's gross monthly income. Cash lenders focusing on the observably highest-risk market segment typically make one-month term loans at 30% interest *per month*. Lenders targeting observably lower risk segments charge as little as 3% per month.[10] Rejection is prevalent even at these high rates: the Lender denies 50% of new loan applicants.

The Lender has been in business over 20 years and is one of the largest micro-lenders in South Africa, with over 100 branches throughout the country. Our experiment took place in a mix of 86 urban and rural branches throughout the provinces of Kwazulu-Natal, Eastern Cape, Western Cape, and Gauteng. All loan underwriting and transactions are conducted face-to-face in the branch network, with the risk assessment technology combining centralized credit scoring with decentralized loan officer discretion. The Lender's product offerings are somewhat differentiated from competitors. Unlike many cash lenders, it does not pursue collection or collateralization strategies such as direct debit from paychecks, or physically keeping bank books and ATM cards of clients. Its pricing is transparent and linear, with no surcharges, application fees, or insurance premiums added to the cost of the loan. The Lender also has a "medium-term" product niche, with a 90% concentration of 4-month loans (Table 1a). Most other cash lenders focus on 1-month or

---

[9] The prevalence of for-profit institutional players makes the consumer credit market in South Africa distinct from most other developing countries (Porteous 2003).

[10] There is essentially no difference between these nominal rates and corresponding real rates. For instance, South African inflation was 10.2% *per year* from March 2002-2003, and 0.4% per year from March 2003-March 2004.

12+-month loans.[11]   The Lender's standard 4-month rates, absent this experiment, ranged from 7.75% to 11.75% per month depending on observable risk, with 75% of clients in the high risk (11.75%) category.

Borrowers face several incentives to repay these high-interest loans.   Carrots include decreasing prices and increasing future loan sizes following good repayment behavior.   Sticks include reporting to credit bureaus, frequent phone calls from collection agents, court summons, and wage garnishments.

III.   Experimental Design and Implementation

We identify specific types of asymmetric information by incorporating the random assignment of interest rates into the day-to-day operations of the Lender.   This section outlines the experimental design and implementation, describes related data collection, and validates the integrity of the random assignments using several statistical tests.   The methodology is implemented in a consumer credit market, but is applicable to other market settings as well.

The experiment was pilot-tested in July 2003, and then fully executed in two additional waves launched in September and October 2003.   We begin with a brief overview of the experiment, and then describe each step in detail below.

A.   Design Overview

First the Lender randomized interest rates attached to "pre-qualified," limited-time offers that were mailed to 57,533 former clients with good repayment histories.[12]   Two rates were

---

[11] The Lender also has 1, 6, 12, and 18 month products, with the longer terms offered at lower rates and restricted to the most observably creditworthy customers.

[12] Private information may be less prevalent among past clients than new clients if hidden information is revealed through the lending relationship (Elyasiani and Goldberg 2004). Hence, if this experiment were conducted on individuals with no prior borrowing history, one may expect larger information asymmetries. To estimate the information asymmetries on such individuals, we also sent soliciations to 3,000 individuals from a mailing list purchased from a consumer database,

assigned to each client: an "offer rate" (r^o) included in the direct mail solicitation, and a "contract rate" (r^c) that was weakly less than the offer rate and revealed only *after* the borrower had accepted the solicitation and applied for a loan. Clients did not know beforehand that the contract rate may be lower than the offer rate. For 59% of the clients, the contract rate was identical to the offer rate. Final credit approval (i.e., the Lender's decision on whether to offer a loan after updating the client's information) and the loan size and term offered to the client were orthogonal to the experimental interest rates by construction. Therefore the two interest rate randomizations enable us to cleanly distinguish selection effects from repayment burden effects, since some clients will select on different interest rates *ex-ante*, but then have identical repayment burdens *ex-post*, while other clients will select on the same rate *ex-ante*, but have different repayment burdens *ex-post*. [13]

We also randomly assigned differential dynamic repayment incentives (D) in order to cleanly identify a moral hazard component of any repayment burden effects. Some clients were assigned eligibility to receive r^c on all future loans taken within the next year (D=1), conditional on repayment performance, while others obtained r^c for just the first loan (D=0). D=1 provides favorable pricing on *future* borrowing (in the 98% of cases where r^c was less than the Lender's standard rate) but does not shift the cost of repaying the borrower's *initial* loan taken at r^c. Thus D enables us to test whether a marginal *incentive*, access to future financing at preferable rates, induces better *choices* related to repayment; i.e., whether D reduces any moral hazard found in this market. Clients were informed of D by the branch manager after all paperwork had been completed and all other terms of the loan were finalized. Figure 2 shows the experimental operations, step-by-step. Figure 3 shows a scatter plot of the offer rate against the contract rate for all individuals (41% of sample) who were assigned $r^c < r^o$.

---

but only 1 person from this list borrowed. A subsequent list was purchased and 5,000 letters sent (but without randomized interest rates) and only 2 people responded. We return to this issue in Section VI.

[13] As detailed in Section IV, we define "repayment burden" as the reduced-form combination of several underlying moral hazard parameters and a wealth effect.

B.   *Sample Frame*

The sample frame consisted of all individuals from 86 branches who had borrowed from the Lender within the past 24 months, were in good standing, and did not have a loan outstanding in the thirty days prior to the mailer. Tables 1a and 1b present summary statistics on the sample frame and the sub-sample of clients who obtained a loan at r^c by applying before the deadline on their mailer. Most notably, clients differ in observable risk as assessed by the Lender. The Lender assigns prior borrowers into "low", "medium", and "high" risk categories, and this determines the borrower's loan pricing and term options under normal operations. The Lender does not typically ask clients why they seek a loan, but the experimental protocol included a survey that indicates the following self-reported uses: education (19%), housing renovations (11%), payoff other debt (11%), household consumption and/or family event (13%), funeral and medical (4%) and miscellaneous/unreported (32%).

C.   *The Randomizations*

Each client was assigned three random variables: an offer interest rate (r^o), a contract interest rate (r^c), and a binary variable for whether the contract rate would be valid for up to one year (D=1) or one loan (D=0). Rates varied from 3.25 percent per *month* to 11.75 percent per month.[14] 41% of the sample was chosen randomly and unconditionally to receive $r^c < r^o$ (Table 1a). At the time of the randomization, we verified that the assigned rates were uncorrelated with other known information, such as credit report score. Table 2 shows that the randomizations were

---

[14] Appendix Table 1 shows the resulting r^o and r^c distributions conditional on the three observable risk categories. Note these are "add-on" rates, where interest is charged upfront over the original principal balance, rather than over the declining balance. We adopt the cash loan market's convention of presenting rates in add-on, monthly form.

successful, *ex-ante*, in this fashion; i.e., conditional on the observable risk category, r^o and r^c were uncorrelated with other observable characteristics.[15]

Lastly, each individual was assigned to receive r^c either for one full year (D=1) or for only the first loan (D=0). In the pilot and the second wave of the experiment, this randomization was conducted at the branch level, such that 14 branches were assigned D=0, and 10 branches were assigned D=1. In the third wave, this randomization was done at the individual level.[16]

D.   *The Offer and Loan Application Process*

The Lender mailed solicitations featuring the offer rate to 57,533 former clients. Each letter had a deadline by which the individual had to respond in order to obtain r^o. The deadline ranged from 2 weeks to 6 weeks, and is discussed in related research (Bertrand, Karlan, Mullainathan, Shafir and Zinman 2005).[17] The Lender routinely mails teasers to former borrowers but had never promoted specific interest rate offers before this experiment.

Clients accepted the offer by entering a branch office and filling out an application in person with a loan officer. Loan applications were taken and assessed as per the Lender's normal underwriting procedures. Specifically, loan officers: a) updated observable information and decided whether to offer *any* loan based on their updated risk assessment; b) decided the maximum loan size for which applicants qualified; and c) decided the longest loan term for which applicants qualified. Each decision was made "blind" to the experimental rates. Table 2 Column 5 verifies that the Lender's rejection decision was in fact uncorrelated with the contract interest rate and

---

[15] Column 3 shows that the dynamic repayment incentive was predicted by number of months since the last loan, the number of prior loans and age. However, including controls for these variables in the primary specifications do not change the estimates of the effect of the dynamic incentive on default.

[16] The dynamic repayment incentive randomization was done initially at the branch level because operations personnel at the Lender were concerned that it would be complicated to communicate D on a case-by-case basis. Once the branches were more comfortable with the experimental design, this was relaxed for the third (and largest) wave of offers.

[17] The solicitations also incorporated randomized decision frames and cues, inspired by findings from marketing and psychology literatures, that were designed to estimate the impact of these "behavioral" effects on consumer demand. These randomizations were orthogonal to the pricing randomizations examined here, by construction.

dynamic repayment incentive. 5,028 (8.7%) clients applied for a loan under this experiment, and of those 4,348 (86.5%) were approved.

In determining maximum loan size, the Lender relies on a debt service ratio: the monthly payment of a loan may not exceed a certain percentage of their net monthly income. A lower interest rate would thus allow for a larger loan. A larger loan might then generate a repayment burden effect, which could cause a higher default rate (and bias against finding moral hazard with respect to the interest rate). In order to mitigate this potential confound, the maximum allowable loan size was established based on the *normal*, not experimental, interest rates.

The contract rate r^c was kept secret from both the loan officer and borrower until after the officer approved the loan application and a loan amount and term were established.[18] Special operations software was developed to facilitate and control this process, and we verify that this condition held in practice in Table 2, column 4 by testing that the offer rate, and not the contract rate (once controlling for the offer rate), predict take-up of the loan. Once the other loan contract features were agreed upon, the software then revealed r^c, which was weakly less than r^o. If the rates were the same, no mention was made of the second rate. If $r^c < r^o$, the loan officer told the client that the actual interest rate was in fact lower than the initial offer. Loan officers were instructed to present this as simply what the computer dictated, not as part of a special promotion or anything particular to the client.

Clients then were permitted to adjust their desired loan size L following the revelation of r^c. In theory endogenizing L in this fashion has implications for identifying moral hazard effects (since a lower r^c strengthens repayment incentives *ceterus paribus*, but might induce choice of a

---

[18] There are several reasons to implement the contract rate assignment "double-blind". Most importantly, we did not want the contract rate to contaminate any selection effects (by influencing either credit approval, or the applicant's decision whether to accept the loan offer). The double blind device also elicits two points on the credit demand curve for each consumer who received $r^c < r^o$ (Karlan and Zinman 2005).

higher L that weakens repayment incentives), as discussed below. But in practice only 10% of borrowers changed their loan demand after r^c was revealed (Karlan and Zinman 2005).[19]

Finally, the software informed the loan officer whether the individual's r^c was valid for one year (47% of borrowers obtained D=1) or for one loan (53% obtained D=0).

E.   Default Outcomes

We use three measures of default: (1) Monthly Average Proportion Past Due (the average default amount in each month divided by the total debt burden), (2) Proportion of Months in Arrears (the number of months with positive arrearage divided by the number of months in which the loan was outstanding), and (3) Account in Collection Status (typically, the Lender considers a loan in collection status if there are three or more months of payments in arrears). Table 1a presents the summary statistics on the default measures. These measures were chosen in consultation with the Lender as proxies for the credit risk, collection costs, and ultimate bad debt incurred by the firm.

IV. Theoretical Overview

We begin by discussing the specific models of private information that motivate our experimental setup, and then describe how the experimental design maps into these models. The Theory Appendix provides a more formal derivation.

We test across two models of selection on unobservables: the Stiglitz-Weiss (1981) *adverse selection* model (hereafter "SW") and the de Meza and Webb (1987; 2001) *advantageous selection* model (hereafter "DW").[20] Specifically, r^o can produce either adverse or advantageous

---

[19] On the other hand, project clients *did* exhibit significant interest rate elasticities with respect to r^o on both the extensive and intensive margins (Karlan and Zinman 2005).
[20] Klonner and Rai (2004) provides a clear comparison of the Stiglitz-Weiss and de Meza-Webb models of selection.

selection, depending on the relationship between borrower risk and return.[21] If risk, defined from the Lender's perspective as the probability of default, and borrower returns are positively correlated, then SW implies that higher rates induce unobservably less risky borrowers to drop out of the applicant pool. Thus under adverse selection, repayment would decrease in r^o as we move away from the initial equilibrium. If risk and return are negatively correlated, then DW implies that higher rates induce unobservably riskier borrowers to drop out of the applicant pool. Thus under advantageous selection, repayment would increase in r^o as we move away from equilibrium.[22] In a consumer credit context, this would hold if borrowers with unobservably relatively unstable income view high interest rates as unaffordable. One limitation of our setup is that if there are heterogeneous selection effects, such that some borrowers select adversely and others advantageously, then the effect of r^o will obscure the true magnitude of selection on unobservables. We explore this in Section VI, although empirically, we cannot distinguish heterogeneous intensity of adverse selection from offsetting adverse and advantageous effects.[23]

The second randomly assigned interest rate, r^c, identifies the impact of repayment burden via a combination of several underlying structural parameters of interest. Repayment burden incentives operate through the borrower's project management and repayment choices. Project management choices are defined as those that impact returns. Higher interest rates will produce moral hazard in *project choice* (conditional on effort) if borrowers prefer mean-preserving spreads in project returns under limited liability (Stiglitz and Weiss 1981). Similarly, higher interest rates

---

[21]There is potentially a third type of selection based on private information, a "lemons" effect, which is unlikely to be important in our setting. As described in Ausubel (1991) and elsewhere, given a setting with competitive bargaining and the presence of private information generated from lending relationships, a single deviating lender would find that reducing rates attracts *ex-ante* unobservably worse repayment risks, since competing lenders will match the rate reduction only for the better risks. But survey evidence on pricing practices in the cash loan market shows that lenders as a rule do not make price concessions, even for good customers. Note that while the lemons effect is commonly described as *adverse* selection, in our setting it is analogous to *advantageous* selection in the sense that reducing interest rates decreases profitability on the margin.

[22] de Meza and Webb (2000) shows that advantageous selection can persist in equilibrium if moral hazard prevents lenders from raising interest rates to clear the market.

[23] In principle the nonprice marketing randomizations discussed in Bertrand et. al. (2005) also could produce unobservable selection effects, but empirically they do not.

reduce effort (conditional on project choice), by producing *debt overhang* that reduces borrower returns in successful states (Ghosh, Mookherjee and Ray 2000). Repayment choice simply refers to the fact that *voluntary default* (conditional on project returns) becomes more attractive under limited enforcement as repayment burden increases (Eaton and Gersovitz 1981; Ghosh and Ray 2001). In contrast, the *income effect* of repayment burden has nothing to do with choice: it works mechanically, by simply increasing the probability that a borrower with uncertain cash flow will be unable to repay. Note that each of these hypothesized incentive and income effects works in the same direction — a higher repayment burden decreases the probability of repayment.

These four components of repayment burden all have intuitive salience in this setting, and hence our priors are agnostic regarding their relative importance. *Project choice* may be relatively limited (compared to say a pure commercial loan market), or may not — anecdotal reports suggest the possibility of "hidden" investment in entrepreneurial projects, and survey evidence reveals cross-sectional variation in the deployment of funds consistent with a range of consumption smoothing and human capital investment opportunities. *Debt overhang* might also be less salient in a consumer rather than commercial credit setting, but then again the relevant effort in the consumer case might be related to maintaining one's wage employment, or to obtaining credit from the informal sector in the event of a negative outcome. *Voluntary default* might be mitigated by reputation effects (repeat contracting opportunities) and aggressive (if imperfect) enforcement, but to what extent? The size of the *income effect* depends critically on the variance of borrower cash flows, which we do not observe.

V. Empirical Strategy

We now present the empirical strategy used to test the theoretical models and interpret the results of the experiment. Recall that we identify any selection and repayment burden effects by

randomly assigning separate offer and contract interest rates to borrowers, conditional on observable risk, and then estimating the relationship between loan repayment and these rates. We employ five approaches to analyzing the results: stylized comparison of means, a base-case OLS specification, a semi-parametric matching estimator, instrumental variables, and non-parametric plots.

Abstracting from functional form considerations for the moment, our basic empirical model takes the form:

(1) $Y_i = f(r_i^o, r_i^c, D_i, X_i)$

where i indexes borrowers. Y is a measure of repayment; r^o is the rate offered on the mail solicitation; and $r^c \leq r^o$ is the rate actually contracted upon loan approval. D is the randomly dynamic repayment incentive (or lack thereof), with D=1 if r^c is valid for up to one year (if the borrower stays current), and D=0 if r^c applies to one loan only. X always includes the Lender's summary measure of observable risk (since the interest rates were randomized conditional on this measure), and also may include other readily observable characteristics that the Lender *could* use for screening.

The Theory Appendix shows formally that:

- r^o identifies the selection effect conditional on r^c -- with $dY/dr^o>0$ if there is adverse selection on net, and $dY/dr^o<0$ if there is advantageous selection on net.

- r^c identifies the repayment burden effect conditional on r^o — with $dY/dr^c>0$ if there is such an effect.

Stylized Comparison of Means: Table 3

We classify rates into "high" and "low" groups, *a la* Figure 1. This is done by setting cutoffs at the median experimental rates for each observable risk category. Table 3 presents mean

comparisons using this method. The findings preview the regression results: we find occasional evidence of adverse selection and repayment burden in the full sample, and more robust evidence for adverse selection among female borrowers and repayment burden among male borrowers.

OLS Specification: Tables 4 and 5

Our base specification is a linear model estimated using OLS:

$$(2) \quad Y_i = \alpha + \beta_o r_i^o + \beta_c r_i^c + \beta_w D_i + \chi X_i + \varepsilon_{ib}$$

i again indexes borrowers, and β_o, β_c, and β_w are the estimates of the selection effect, the repayment burden effect, and the dynamic incentive effect, respectively. X need include only the Lender's summary measure of observable risk since the randomizations conditioned only on this variable. We also include fixed effects for the month in which the offer letter was sent (June, September, or October 2003). The error term, ε_{ib}, is corrected for clustering at the branch level, *b*. The model is estimated on the takeup sample of 4,348 observations since these are the only project clients for whom we observe repayment behavior. The OLS results are robust to including loan size and term as control variables, which is noteworthy since loan size responds to the interest rate (Karlan and Zinman 2005) and could have an independent effect on default. Results obtained with the OLS estimator are similar to the comparison of means, and are presented in Tables 4 and 5 and discussed in Section VI.

Semi-Parametric Matching Estimator: Table 6

Next we develop a semi-parametric approach that resembles a matching estimator, but with a continuous treatment variable. This is motivated by the concern that the OLS specification will impose incorrect functional form assumptions if any of the three different random variables interact to influence default. We address this possibility using three related specifications that impose

functional form assumptions on only one random variable at a time, while conditioning non-parametrically on all combinations of the other two random variables and observable risk. For example, to identify the selection effect (r^o), we group observations that match perfectly on the other random variables (r^c and D) and the risk level, and employ a fixed effect model that de-means the dependent variable (default) and the variable of interest (in this case r^o) within these perfectly-matched groups. This model is shown is equation (3) below, and then we present the analogous models for identifying the repayment burden effect and dynamic incentive effect in equations (4) and (5), respectively. In each of these formulas the subscript "c" represents contract interest rate, "o" represents offer interest rate, "d" represents the dynamic incentive, and "r" represents the lender-defined categorical risk level (high, medium or low). The unit of observation is *i*, the individual borrower. The month of the offer letter, *M,* is a categorical variable (either July, September or October, 2003).

Match on c,d,r to identify the offer rate effect: (3) $Y_{icdr} - \overline{Y}_{cdr} = \beta(r^o_{icdr} - \overline{r}^o_{cdr}) + \delta M_i + \varepsilon_{icdr} - \varepsilon_{cdr}$

Match on o,d,r to identify the contract rate effect: (4) $Y_{iodr} - \overline{Y}_{odr} = \beta(r^c_{iodr} - \overline{r}^c_{odr}) + \delta M_i + \varepsilon_{iodr} - \varepsilon_{odr}$

Match on c,o,r to identify the dynamic incentive effect: (5) $Y_{icor} - \overline{Y}_{cor} = \beta(W_{icor} - \overline{W}_{cor}) + \delta M_i + \varepsilon_{icor} - \varepsilon_{cor}$

Table 6 presents the estimates obtained using these three specifications. The results are discussed in Section VI.

Instrumental Variables Estimator: Table 7

Lastly, we consider the case where repayment burden is defined as total interest due on the loan, not merely the marginal cost of debt (the interest rate). Total interest due includes an *endogenous* component (loan size) multiplied by a random variable (the interest rate), so we employ an instrumental variables approach to identify the effect of loan pricing on default.

Specifically, we use the random variables to instrument for endogenous variable of interest, debt burden. The instrumental variable specification is:

(6) $\quad Y_i = \alpha + \beta_0 \hat{I}_i^o + \beta_c \hat{I}_i^c + \delta X_i + \varepsilon_i$

where I^0 and I^C are the endogenous variables, total interest due under the offer and contract rates, respectively. The first stage specifications are:

(7) $\quad I_i^o = \alpha^1 + \beta^1_o r_i^o + \beta^1_c r_i^c + \chi^1 X_i + \upsilon^1_i$

and

(8) $\quad I_i^c = \alpha^2 + \beta^2_o r_i^o + \beta^2_c r_i^c + \chi^2 X_i + \upsilon^2_i$

with r_i^o and r_i^c serving as instrumental variables for I^0 and I^C.

These results are presented in Table 7 and discussed in the next section.

Plot of Coefficients: Figures 4-11

In figures 4-11 we relax the linear treatment effect assumptions imposed by (2)-(8) using a nonparametric graphical approach described in Section VI, part D.

VI. Empirical Results

A. *Overview of Base Specification*

Our primary analysis estimates equation (2) using ordinary least squares, tobit, or probit on the "full" sample containing all three observable risk categories (Table 4).[24] The dependent variables are the three default measures described above. In all specifications, the interest rate

---

[24] Results are robust to including the log of loan size, and loan term, to address the possibility of interest rate effecting repayment through its effect on loan size and term. An alternative, instrumental variables approach is presented in Table 7. This too yields qualitatively similar results. Nor do results change if we include branch fixed effects to control for any differences in experimental implementation and/or the mechanical influence of varying mailer dates (staggered by groups of branches) on repayment measures. Pooling the risk categories implies that the full sample model lacks a common support for interest rates exceeding 7.75 (since, e.g., low risk borrowers are never offered a rate above 7.75). We re-estimate our base specifications over a common support and report qualitatively similar results in Appendix Table 2, although much power is lost on the smaller sample size.

units are in monthly percentage points (e.g., 7.50 for 7.50% per month), and we report marginal

effects where applicable. Results on interest rate variables therefore capture the effect of a one

percentage point (100 basis points) increase in the monthly rate. We find evidence of moral hazard

through the dynamic pricing incentive D, but little suggestion of any selection on unobservables in

the full sample.

B.   *Primary Results with the Base Specification*

Row 1 of Table 4 presents estimates of β_o, the response of repayment behavior to the offer

rate. This coefficient identifies any net selection on unobservables, with $\beta_o > 0$ indicating adverse

selection, and $\beta_o < 0$ indicating advantageous selection. We find no robust evidence in either

direction on the full sample.[25]

Row 2 of Table 4 presents estimates of β_c, the response of repayment behavior to the

contract rate. This coefficient identifies any effect of repayment burden, with $\beta_c > 0$ indicating some

combination of moral hazard and wealth effects. Similar to the adverse selection results, we find

consistently positive effects that are typically insignificant statistically. The one marginally

significant result (column 2) implies that a 400 basis point cut would reduce the average number of

months in arrears by 13%.[26]

Results on D, the dynamic repayment incentive variable, deliver robust evidence of moral

hazard (Table 4, row 3). Recall that clients with D=1 face a marginal *incentive* to repay — if they

maintain good standing with the Lender they are eligible to borrow at r^c for up to a year. D

consequently has no effect on the debt burden of the current loan, but does increase the benefit of

---

[25] Recall from Section IV that the offer rate coefficient will understate the true extent of private information problems if there are concurrent adverse and advantegeous selection effects in the sample. This would manifest as parameter heterogeneity; e.g., if part of the sample has $\beta_o < 0$, and part of the sample has $\beta_o > 0$. We explored this possibility by interacting the offer rate variable with various demographics (see, e.g., Table 8), and found some evidence that adverse selection was decreasing in income. Splitting the sample into income quintiles or at the median (not reported), we found that only one of the six offer rate coefficients had the negative sign consistent with advantageous selection, and that this coefficient was not statistically significant (t-stat = 0.6).
[26] Coefficient * 4 / mean outcome = 0.007 * 4 / 0.219 = 13%

maintaining a good relationship with the lender by reducing the interest rate on future loans. D therefore is free of the noise (bad shocks) and confound (endogenous loan size) that might bias estimates of the reduced-form repayment burden effect toward zero. D thus identifies pure *moral hazard* that is alleviated by the provision of a marginal, dynamic repayment incentive. D's effect is large and significant, with the incentive producing decreases in the various default measures ranging from 1.1 to 1.9 percentage points in the OLS specifications on the full sample. These magnitudes imply that D=1 clients defaulted 7 to 16 percent less often than the mean borrower. Table 4 (Columns 2, 4 & 6) also shows that this effect is increasing in and driven by the size of the discount on future loans, as each 100 basis point decrease in the price of future loans reduces default by 4% in the full sample.

Table 5 shows the primary results from Table 4, but by gender. Recall from the introduction that gender differences in response to liquidity and incentives have been documented in several other settings, and that differential response to interest rates are of particular interest to development economists and microfinance practitioners. In keeping with conventional wisdom, we find evidence of moral hazard (but not adverse selection) for males. Contrary to conventional wisdom, we find adverse selection (but not moral hazard) for females.

Table 5, Columns 1 through 6 show the results for male borrowers. We find evidence of moral hazard but no selection on unobservables. Both experimental instruments for moral hazard, the contract rate and the dynamic repayment incentive (D), are large and sometimes significant determinants of default. The coefficient on the offer rate is relatively small and always insignificant.

Table 5, Columns 7 through 12 show a different pattern for female borrowers. We find evidence of adverse selection but no repayment burden or moral hazard effects. The offer rate coefficient is always large and positive for females, and statistically significant in 5 of the 6 regressions reported, indicating adverse selection. On the other hand, the contract rate coefficient

is now wrong-signed (and significant in one case). The dynamic repayment incentive results are insignificant, but signed negatively (evidence for moral hazard) and often large economically. On balance then there is no statistically significant evidence for repayment burden or moral hazard influencing the repayment behavior of female borrowers.

In specifications not shown, we test whether the gender effects are significantly different from each other by including gender interaction terms in the Table 4 specifications. The interaction term is significant statistically for both the offer and contract interest rates, but not for the dynamic repayment incentive variable. This is not surprising, since the coefficient on the dynamic repayment incentive for women is negative and similar (slightly smaller) to that for men, just not significant statistically.

C.   *Primary Results with Matching Estimator and IV Specification*

Table 6 presents the results of the semi-parametric matching estimator specified in equations (3)-(5). The point estimates remain largely the same as in the OLS results shown in Table 5, but the standard errors are slightly larger due to the loss of hundreds of degrees of freedom.

Table 7 presents the results of the IV estimator outlined in equations (6)-(8). These follow a similar pattern to the results obtained from the base specification (Tables 4 and 5), with the exception that we now find more weakly significant evidence for adverse selection. (As before, however, the significant full sample results appear to be driven by females.) Results from estimating the endogenous version of the IV estimator (second panel of Table 7) suggest that using total interest cost (or debt burden) as the regressor can lead to misleading inferences in cases where exogenous variation is not available. We surmise that the difficulty stems from the fact that lower interest rates-- i.e., a lower *marginal* cost of borrowing-- produce demand for larger loan sizes

(Karlan and Zinman 2005), and hence a larger *total* cost of borrowing. Recall that our IV estimator surmounts this problem by instrumenting for the total cost of borrowing with our randomly assigned marginal cost of borrowing.

D.   Graphs

Figures 4-11 show smoothed plots of coefficients from a non-parametric estimator that relaxes the linear treatment effect assumption maintained in all of our regression specifications. Each plot has one random variable of interest; e.g., in Figure 4 this variable is the offer rate. The estimates of interest are the coefficients on the each individual offer rate (e.g., one cell for 7.25, another for 7.5). The specifications include controls (again, non-parametric) for the contract interest rate, risk category and month of the offer. The plots reproduce the qualitative patterns (sign and magnitudes) found in our regressions, but are not conclusive in the sense that the confidence intervals leave open the possibility of nontrivial nonlinearities.[27]

E.   Gender Analysis

Next, we explore whether the gender effects are actually driven by systematic variation in household demographics rather than something unobservable and/or fundamental to gender *per se*. Such compositional confounds could be driven by differential demand or supply, although in the case of supply, the Lender does not explicitly use gender in their screening of clients. Tables 1a and 1b show that there are minimal observable compositional differences across gender. A parsimonious test (shown in Table 8) for compositional confounds is to add interactions between the interest rate variables and observable characteristics to our base specification. The significant interaction terms between gender and the rate variables, *conditional on other demographics* (and their interactions with rates and gender), suggest that gender differences are not merely driven by

---

[27] The confidence bands on these plots were created by bootstrapping with 500 replications.

mechanical composition effects. We continue to find strong repayment burden effects among males, and strong adverse selection effects among females. None of the demographic interaction effects are significant, nor do they detract from the simple gender interaction effects with rates. If, for instance, the gender effect was masking a "female and married" effect, then we would observe a significant "married and offer rate" interaction and the gender interaction with the offer rate would go to zero.

A related check is to predict default using observable demographic information, and then test whether observable selection on the interest rate differs across males and females. If unobservably riskier women are more likely to borrow at high interest rates, one should also find that observably riskier women are also more likely to borrow at high interest rates. Observable risk is defined here as the predicted probability of default. Table 9 shows that the effect of the interaction between observable risk and the offer rate on the decision to apply is significant and positive for women, but not for men. This means that the price elasticity of the application decision is decreasing in observable risk for women but not for men, and provides further confirmation of a gender difference that is not driven purely by other observable characteristics.

Of course we do not observe all of the observable characteristics of interest; e.g., we do not have independent measures of education and occupation, nor do we observe health status, unmarried co-habitation or head-of-household. Hence the above tests do not completely rule out the possibility that relatively mechanical demographic differences generate the observed gender pattern.

On balance, however, the evidence suggests that male and female borrowers do pose different types of private information problems for the Lender, with strong evidence that females select adversely and some evidence that male repayment is relatively sensitive to repayment burden. The question of what drives this pattern (e.g., demographics, outside credit options, social norms, hard-wiring....) merits further, more systematic, exploration.

F.   Observable Determinants of Default and Selection on Observables

Significant interactions between interest rates and observable characteristics beg the related question of how efficiently the Lender assesses risk. We begin by exploring whether observable characteristics help predict default, *conditional on the Lender's summary statistic for risk*, by reporting results obtained from adding several additional observables to equation (2). Table 10 shows that the Lender's summary statistic for observable risk does not in fact completely summarize the role of observables, at least over the range of interest rates used in our experiment, as several readily observed variables help predict default. Significant predictors include credit scores and the number of prior transactions with the Lender. However, adding observables beyond the summary statistic generates little or no improvements in the overall explanatory power of the models (as measured by the adjusted R-squareds in Tables 4 and 10).

G.   External Validity and the Power of Repeated Transactions

We may find little evidence of adverse selection in the full sample because our borrowers have already revealed their types to the Lender; i.e., in the process of transacting, private information becomes public.[28] We explore this possibility indirectly, and within-sample, by exploring whether the offer rate effect varies with the number of prior loans the borrower has taken from the Lender. Table 11 shows that this is indeed the case: adding a prior loans main effect and interaction with r^o to equation (2) produces a negative and significant interaction term. The interaction effect is large; e.g., it eliminates 43% of adverse selection at the mean number of prior loans (4.3) in the full sample. Thus selection is indeed relatively more adverse for those borrowers

---

[28] We sought to include clients with no prior relationship with the Lender by extending 3,000 offers to names obtained from a mailing list; unfortunately, the list yielded only one borrower. A pilot follow-up list of 5,000 offers (without randomization of interest rates) yielded two borrowers. Hence, in order to conduct this experiment on "new" clients, we need to find an alternative channel for identifying the sample frame and contact information.

with whom the Lender is least familiar. The repayment burden effect (through the contract interest

rate) appears to decrease with prior transaction frequency as well. Moral hazard (as identified by

the dynamic repayment incentive), on the other hand, abates only a little. Note, however, that we

should not conclude that the relationship *caused* the information asymmetries to abate. Rather, if

those with more frequent transaction histories are simply less able or willing to *exploit* information

asymmetries, then we also would find negative coefficients on the interactions between interest rate

and prior borrowing frequency. In other words, we can not distinguish whether it is the frequent

borrowers' "type" or the relationship itself that drives the results. Think of implementing our

experiment on the most frequent borrowers, but years ago, *before their first loan with the Lender*.

If their "type" is driving today's behavior, then this initial experiment also would have found no

evidence of adverse selection or moral hazard. Repeated experiments, and panel data, would help

identify the causal effects of lending relationships.

H.   Magnitude Calculations Comparing Observables and Unobservable Effects

We now estimate the relative importance of private vs. public information in determining

default. Recall that the Lender uses the three-tiered risk category to determine interest rates, and

that the high risk clients can borrow at 11.75% per month and the low risk clients borrow at 7.75%

per month. We pose the question: how much of the reduced default observed between high and

low risk clients is due to the *unobservable* selection and incentive effects generated by lowering the

interest rate by 4.00% per month, and how much is due to the *observable* characteristics that

allowed the lender to classify the clients as such. High risk clients have much higher default rates:

e.g., a 22 percentage point higher proportion of months in arrears than low risk clients (27

percentage points for men, and 17 percentage points for women).[29] We take this *observable* difference and calculate how much of it appears to be due to information asymmetries with respect to interest rates. The remainder is due to the *observable* characteristics that led the lender to classify the clients into high and risk categories. The point estimates suggest that this interest rate reduction alone will produce a 5.6 percentage point drop in default among men, through reduced repayment burden (Table 5, Column 2), and a 3.2 percentage point drop among women due to reduced adverse selection (Table 5, Column 8).[30] Given a disparity between high and low risk default rates of 27 percentage points for men and 17 for women, information asymmetries with respect to the interest rate explain about 20% of differential default rate between the Lender's high and low risk clients (5.6%/27% for men; 3.2%/17% for women).

VII. Conclusion

We develop a new market field experiment methodology that disentangles adverse selection from moral hazard. The experiment is implemented in a South African consumer credit market, and yields evidence of moral hazard among males and adverse selection among females. This confirms the conventional wisdom on moral hazard in microfinance, but raises a new puzzle for theorists and practitioners: why do we only see adverse selection among women? The moral hazard finding, taken alone, could be driven by a lower cost of default for men than for women. This could occur if men had better outside options, or different social norms and values regarding default and/or smoothing bad shocks (by, e.g., drawing on family and community resources). However, the lower-cost-of-default story does not square well with our finding that only women

---

[29] In our sample, there is a 22 percentage point difference in default between high and low risks among borrowers who obtained rates within 100 basis points of the Lender's normal rate for their risk class. The unadjusted mean difference in our sample is 15.0 percentage points (Table 1a).

[30] The percentage point reduction is calculated by multiplying the coefficient on the interest rate variable (contract rate in the case of men, offer rate in the case of women) by the normal interest rate spread between low and high risk clients (4 percentage points). Estimates are generally smaller (or zero) if we calibrate on the full sample, although the result on the size of the dynamic repayment incentive ("Rate valid for one year * points below the normal rate") for the proportion-of-months-in-arrears measure suggests a 3.2 percentage point reduction in default (Table 4, Column 4).

exhibit adverse selection, since men facing relatively low costs of default should select adversely as well if screening is imperfect (and it is, as evidenced by the adverse selection among women). Hence the finding the men commit moral hazard and women engage in adverse selection poses a difficult puzzle that motivates further research, both empirical and theoretical.

In the full sample, our findings aggregate to robust evidence for moral hazard and weak evidence for adverse selection. We also find evidence suggesting that asymmetric information problems are mitigated if there are prior loan transactions—a lending "relationship", in banking and finance parlance—between the borrower and Lender. Where statistically significant, the effects of private information are economically important: our results indicate that adverse selection and moral hazard explain about 20% of default in our sample. Overall the findings provide unique empirical evidence of significant, specific information asymmetries in a consumer credit market, and thus help explain the prevalence of credit constraints even in a market that specializes in financing high-risk borrowers at very high rates.

These findings should translate into policy and practical implications with future research. The presence of empirically importance information asymmetries indicates the potential for welfare-enhancing investments in screening and monitoring technologies on the margin. Inframarginally, it suggests that loan production innovations ranging from credit scoring to group lending have helped drive credit market development by mitigating information problems. But more work is needed to identify the prevalence or absence of asymmetric information in different settings and markets. For example, we are currently planning additional experiments, in the lab and in the field, to test the robustness of our gender effects and explore underlying mechanisms driving those results.[31]

---

[31] The results also motivate additional data collection on, for example, the incidence of bad shocks by gender. In a brief phone survey, reaching 374 former borrowers of the Lender (including 61 defaulters), we found no differential incidence of adverse shocks.

Development finance, over the past thirty years, has seen a plethora of interventions intended to help overcome presumed information asymmetries; e.g., group liability rules (Armendariz de Aghion and Morduch 2005). Lenders can best design such interventions if they are equipped with better information about the nature and importance of any information asymmetries. However, optimal policy design requires the identification of additional parameters, including the relative effectiveness of different interventions at mitigating information problems, and the marginal (rationed) borrower's returns to borrowing. Our study has provided unique, micro-foundation evidence for specific information asymmetries. Combining this type of finding with further research on specific interventions and potential borrower returns will produce specific policy prescriptions.

Theory Appendix

Here is a more formal derivation of how our research design maps into theoretical models of private information, and thereby permits identification of unobservable selection and moral hazard effects.

Assume a Lender implementing our experiment is faced with loan applicants who have identical observable characteristics but may be heterogeneous with respect to unobservable information. These characteristics q are not observable to the Lender, but are known to the applicant. Let q ("riskiness") be continuous, and bounded below by zero: $Q = \{q, q \in [0, \infty)\}$, where a higher q negatively impacts but does not wholly determine the "success" of the borrower's project, which is defined discretely. The success/fail framework is more typically applied to a pure commercial credit market, but it also makes sense in a consumer credit setting. Here we can think of "success" as having sufficient funds to repay the consumer loan, whether these funds are from entrepreneurial activity, wage income, or other financing sources. Borrowers succeed with probability p(q, e) and fail with probability $1-p(q,e)$, where e is effort exerted by the agent. We allow effort to be continuous, $e \in [0,1]$, and assume it imposes a linear cost.

We assume that p is twice continuously differentiable in effort, and differentiable with respect to the common shock and the unobservable risk q. Next we impose the following standard assumptions on the probability structure:

Assumption 1.
$$\frac{\partial p(q,e)}{\partial q} < 0, \frac{\partial p(q,e)}{\partial e} > 0, \frac{\partial^2 p(q,e)}{\partial e^2} < 0, \frac{\partial^2 p(q,e)}{\partial e \partial q} < 0.$$
$$Inada\ conditions: \frac{\partial p(q,0)}{\partial e} \to \infty, \frac{\partial p(q,1)}{\partial e} = 0.$$

i.e., the probability of success decreases with riskiness, q; the probability of success increases with borrower's effort, e; and, the probability of success with respect to effort, e, decreases with riskiness, q (and vice versa).

Next assume for simplicity that the return to the borrower is $R(q)$ in the event of success and zero in the event of failure. We assume for the moment that returns are observable and verifiable, thereby abstracting from the possibility of *voluntary default* (Eaton and Gersovitz 1981; Ghosh and Ray 2001). Then default occurs if and only if the project doesn't succeed under the additional simplifying assumption that:

Assumption 2. $\forall q \in Q: \quad R(q) > (1 + r^c)B.$

Where B is the loan principal amount demanded and r^c is the interest rate on the loan contract. So a borrower repays in full if her project succeeds and repays nothing if her project fails.

We now show that the assumption on the relationship between risk (to the Lender) and returns (to the borrowers) is critical to identifying any selection effects of interest rates. Stiglitz and Weiss (1981) shows that *adverse selection* results if this relationship is positive. Formally,

Assumption 3a ("SW"). $\forall q, q' \in Q: \quad p(q,e)R(q) = p(q',e)R(q') = C(e).$

Where C is a constant. The equation states that expected returns to the borrower are constant— projects that yield high returns in the successful state have low probabilities of success. We show below that this condition will indeed produce adverse selection in our setting.

De Meza and Webb (1987) shows that *advantageous selection* results if risk and returns are negatively correlated. Formally:

Assumption 3b ("DW"). $\forall q, q' \in Q: \quad q > q' \Leftrightarrow p(q,e)R(q) < p(q',e)R(q')$.

This will hold, e.g., if borrowers differ only in the probability of project success, but not in project payoff conditional on success. We show below that in our setting 3b implies that raising the interest rate discourages *low* quality borrowers on the margin, thereby improving the average composition of the borrower pool via *advantageous selection*.

We solve for selection and moral hazard effects by focusing on the borrower's problem. Define the borrower's expected return (after the effort choice is made) as:

$$(1) \quad E(\pi) = p(q,e)[R(q) - (1 + r^c)B] - e$$

The Lender does not negotiate on price in either standard or experimental operations, so borrowers are price takers at r. We ignore the Lender's problem since in our setting the interest rate is not a choice variable (it was randomly assigned), and the variables the Lender does control (loan supply on the extensive and intensive margins) are orthogonal to the rate by construction. Therefore we can assume (without loss of generality) that applicants are approved by the Lender.[32]

Accordingly we return to the borrower's problem and begin by solving the model through backwards induction; i.e., conditional on the borrower deciding to apply, she decides upon the

---

[32] In practice, 84% of applicants were approved in the experiment. More generally, one can think of any rejected borrowers as being *observably* differentiated— and this model conditions on observable information.

repayment effort after learning the contract interest rate, r^c. Note the interest rate that the agent

takes into account is the contract rate, not the offer rate (in the case where they differ). Therefore,

holding riskiness, q, fixed, the agent solves

$$(2) \max_{e} \quad p(q,e)[R(q) - (1 + r^c)B] - e$$

Given our set of assumptions, the optimization program yields a unique interior solution for each

value of q and is characterized by the following first-order condition:

$$(3) \; \tilde{e} = e(q, r^c): \quad \frac{\partial p(q, \tilde{e})}{\partial e} = \frac{1}{R(q) - (1 + r^c)B}.$$

Proposition 1. The level of effort chosen is inferior to the first-best value (that is, when effort is

observable and verifiable). Moreover, note that $\partial \tilde{e} / \partial r^c < 0$. This is the *debt overhang* version of

moral hazard effect -- the higher the interest rate, the less the optimal effort since the agent only

receives a positive return in case of success (i.e., the return function is convex). Proof is at the end

of this Appendix. In this setup, a *voluntary default* model would yield qualitatively identical

results regarding the relationship between repayment and r^c (Eaton and Gersovitz (1981) or Ghosh,

Mookherjee et al (2000)).

The next step in solving the borrower's problem is to examine the decision to apply for the loan,

which is made using the offer rate, r^o. (Recall from Section III that borrowers are not aware that

there might be a distinct contract rate, r^c, when they are deciding whether to apply for the loan.)

Define the marginal applicant as the one who has expected returns of exactly zero. That is,

$$(4) \; \hat{q}: \quad p(\hat{q}, e(r^o))[R(\hat{q}) - (1 + r^o)B] - e(r^o) = 0.$$

Proposition 2. If the SW assumption (3a) holds, the agent applies for a loan if $q \geq \hat{q}$. Moreover,

an infinitesimal increase in r^o increases the marginal borrower's q, $\partial \hat{q} / \partial r^o > 0$. (There is no

effect through effort since it is endogenous, and the marginal effect is zero by the envelope theorem). Therefore when the offer rate increases the marginal applicant is riskier; i.e., the safer borrowers choose not to apply, creating a pool that is riskier on average. This is the classic adverse selection effect *a la* SW. If instead the DW assumption (3b) holds, the agent applies if $q \leq \tilde{q}$. In this case $\partial \tilde{q} / \partial r^o < 0,$; i.e., increasing the offer rate decreases the marginal applicant's riskiness, q, and the applicant pool becomes less risky on average. This is advantageous selection *a la* DW. (See the end of this Appendix for proofs.)

We can now tie our propositions regarding the selection effects of the offer interest rate, r^o, and the moral hazard effect of the contract interest rate, r^c, directly to an empirical outcome of interest, the probability of default. According to the model, the expected probability of default, once r^c is known and effort is chosen, can be expressed as:

(5) $E(\text{probability of default}) = \int_{\hat{q}}^{\infty} [1 - p(q, \tilde{e})] \frac{f(q)dq}{1 - F(\hat{q})}.$

Proposition 3. The marginal effect of r^o on the default probability captures the effect of selection. If the SW assumption holds, then:

(6) $\dfrac{\partial E(\text{probability of default})}{\partial r^o} = \dfrac{\partial \hat{q}}{\partial r^o} \dfrac{f(q)}{1 - F(q)} \int_{\hat{q}}^{\infty} [1 - p(q, e)] \dfrac{f(q)dq}{1 - F(\hat{q})} - \dfrac{\partial \hat{q}}{\partial r^o} \dfrac{f(q)}{1 - F(q)} [1 - p(\hat{q}, e)] > 0.$

The proof is a direct application of proposition 2. If instead the DW assumption holds then the effect of a marginal change in the offer rate on the estimated probability of default has a negative sign.

On the other hand the marginal effect of r^c will capture the moral hazard effect,

(7) $\dfrac{\partial E(\text{probability of default})}{\partial r^c} = -\displaystyle\int_{\hat{q}}^{\infty} \dfrac{\partial p(q,\tilde{e})}{\partial e}\,\dfrac{\partial \tilde{e}}{\partial r^c}\,\dfrac{f(q)dq}{1-F(\hat{q})} > 0.$

The result is again immediate since, by proposition 1, $\partial \tilde{e}/\partial r^c < 0$.

Incorporating the dynamic repayment incentive (D) is no different substantively than increasing the benefits of repayment, holding the costs constant. This additional repayment incentive may inspire more effort to ensure a successful outcome (mitigating *debt overhang*) or simply more incentive to choose to repay the first loan (discouraging *voluntary default*). Incorporating this into the above, we add D to formula 2 as a benefit which accrues to the borrower if they repay the first loan:

(8) $\displaystyle\max_{e}\quad p(q,e)[R(q) + D - (1 + r^c)B] - e$

(9) $\tilde{e} = e(q, D, r^c):\quad \dfrac{\partial p(q,\tilde{e})}{\partial e} = \dfrac{1}{R(q) + D - (1 + r^c)B}.$

(10) $\partial \tilde{e}/\partial D > 0.$

So D increases effort and hence reduces default.

Proof of Proposition 1.

To show that the effort level is lower than the first-best level of effort, begin by noting that in a first-best setting where effort is observable and verifiable to all parties the first-order condition reads:

$\dfrac{\partial p(q, e^*)}{\partial e} = \dfrac{1}{R(q)}$

The right-hand side of this first-order condition is smaller than the one for which effort is unobservable, making the optimal effort level larger due to decreasing returns in effort. To show the moral hazard effect, $d\tilde{e}/dr^c < 0$, totally differentiate the first-order condition to obtain:

$$\frac{d\tilde{e}}{dr^c} = \frac{B}{\frac{\partial^2 p}{\partial e^2}\left[R - (1+r^c)B\right]^2} < 0,$$

since the returns to effort are decreasing. *Q.E.D.*

Proof of Proposition 2.

Assume the SW assumption (3a) holds. Recall that the marginal borrower's return once the offer rate is announced is, by definition, zero. Since the expected returns are increasing in q only applicants with q's higher than marginal borrower's $\hat{q}$ are going to have nonnegative expected returns. Accordingly these borrowers, with $q > \hat{q}$, form the pool of applicants.

Totally differentiating the marginal applicant condition yields:

$$\frac{d\hat{q}}{dr^o} = \frac{pB}{-(1+r^o)B\frac{\partial p(\upsilon,\hat{q},\tilde{e})}{\partial q}} > 0,$$

since p is decreasing in unobservable risk q. If instead the DW assumption (3b) holds the steps are the same but the signs are the opposite. In particular expected returns are decreasing in q and so only safer projects than the marginal q apply. In this case the total differentiation produces:

$$\frac{d\tilde{q}}{dr^o} = \frac{pB}{\left[R - (1+r^o)B\right]\frac{\partial p(\tilde{q},\tilde{e})}{\partial q}} < 0. \ \textit{Q.E.D.}$$

Figure 1: Stylized Depiction of Experimental Design

	High Contract Rate	Low Contract Rate
High Offer Rate	Moral Hazard / Repayment Burden ⟷	Adverse Selection ↕
Low Offer Rate	N/A	

Figure 2: Operational Steps of Experiment

| 57,533 direct mail solicitations with randomly different offer interest rates sent out to former clients. | 5,028 clients go to branch and apply for loan. | Client is offered r^o (regardless of whether she brings in letter). | Loan officer makes credit and loan supply decisions based on "normal" interest rates, hence "blind" to experimental rates. 4,348 clients are approved. | Client offered loan at r^c (contract rate). Borrower may revise size and term. | Contract finalized and client told whether rate is good for one year (D=1) or just one loan (D=0). | Client given short survey and then picks up cash | Repayment behavior observed. |

Figure 3: Scatter plot of Contract versus Offer Interest Rates on 4 Month Loans
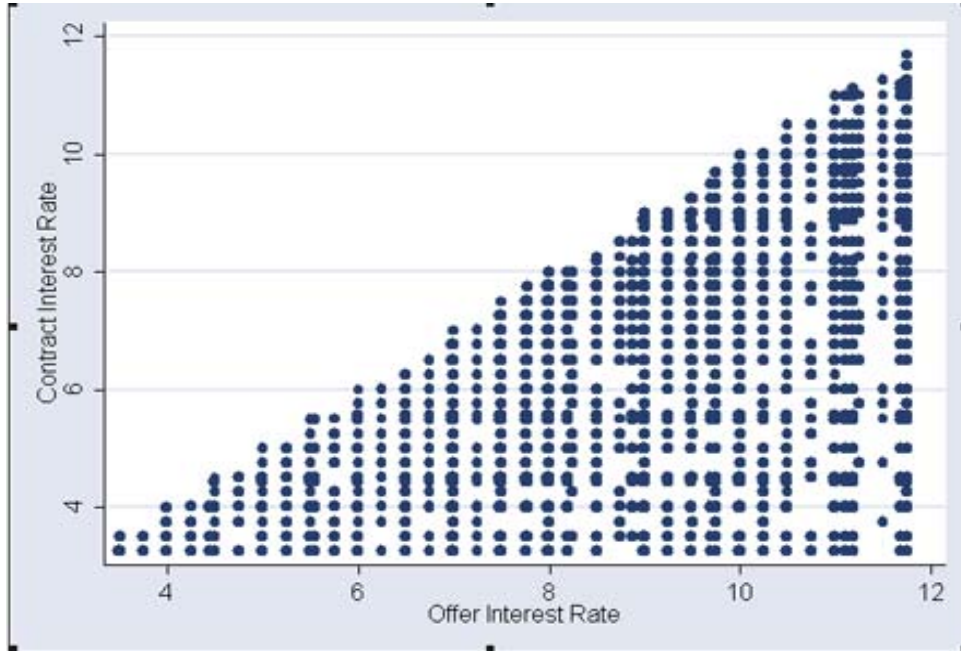
Plot includes only the 41% of borrowers that received a contract rate less than the offer rate.

Figures 4-7: Females
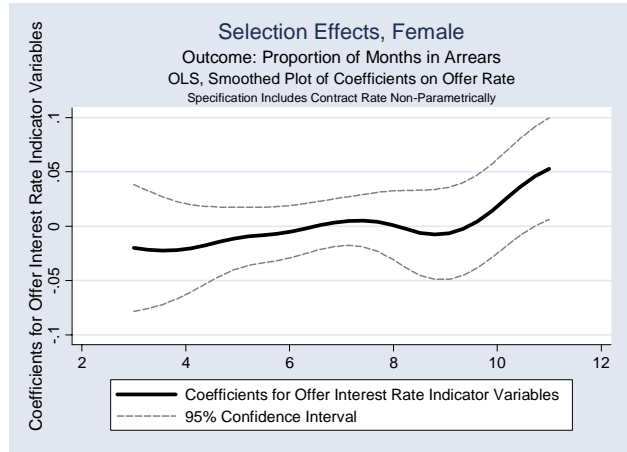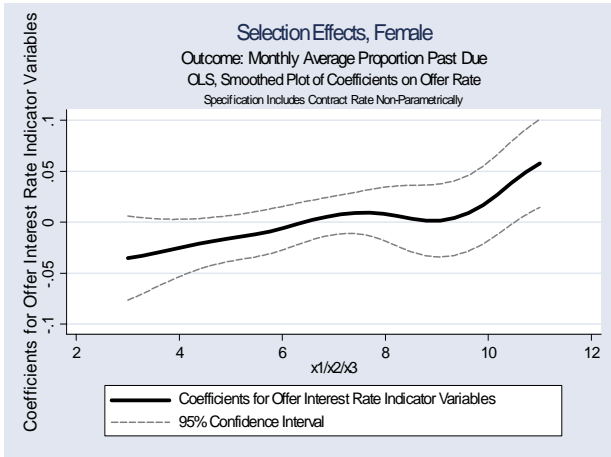
Selection Effects, Female
Outcome: Monthly Average Proportion Past Due
OLS, Smoothed Plot of Coefficients on Offer Rate
Specification Includes Contract Rate Non-Parametrically

Selection Effects, Female
Outcome: Proportion of Months in Arrears
OLS, Smoothed Plot of Coefficients on Offer Rate
Specification Includes Contract Rate Non-Parametrically

These are smoothed plots of the coefficients on the offer rate indicator variables (y-axis), for each individual offer rate (x-axis), for females. The dependent variable is default, measured as the average past due amount as a proportion of original principal (on the left graph) and as the proportion of months in arrears (on the right graph). Other independent variables include contract rate indicator variables (grouped by integer to preserve degrees of freedom), the lender-defined risk level, and the month of the offer. The upward slope is indicative of adverse selection with respect to the interest rate.
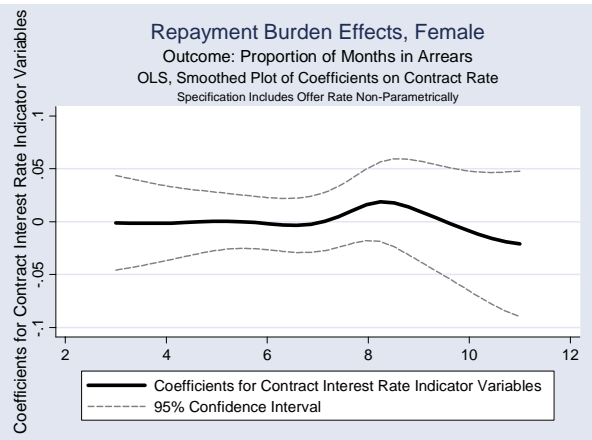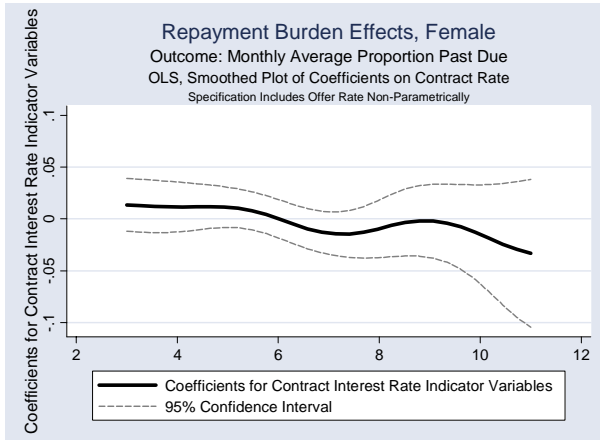
Repayment Burden Effects, Female
Outcome: Monthly Average Proportion Past Due
OLS, Smoothed Plot of Coefficients on Contract Rate
Specification Includes Offer Rate Non-Parametrically

Repayment Burden Effects, Female
Outcome: Proportion of Months in Arrears
OLS, Smoothed Plot of Coefficients on Contract Rate
Specification Includes Offer Rate Non-Parametrically

These are smoothed plots of the coefficients on the contract rate indicator variables (y-axis), for each individual contract rate (x-axis), for females. The dependent variable is default, measured as the average past due amount as a proportion of original principal (on the left graph) and as the proportion of months in arrears (on the right graph). Other independent variables include offer rate indicator variables (grouped by integer to preserve degrees of freedom), the lender-defined risk level, and the month of the offer. The flat slope is indicative of a lack of repayment burden effects with respect to the interest rate.

Figures 8-11: Males
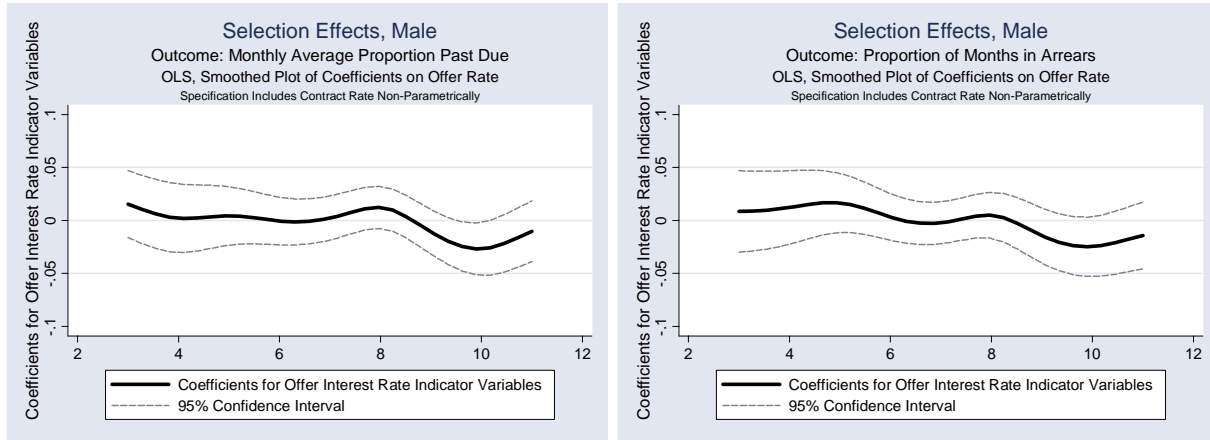
Selection Effects, Male
Outcome: Monthly Average Proportion Past Due
OLS, Smoothed Plot of Coefficients on Offer Rate
Specification Includes Contract Rate Non-Parametrically

Selection Effects, Male
Outcome: Proportion of Months in Arrears
OLS, Smoothed Plot of Coefficients on Offer Rate
Specification Includes Contract Rate Non-Parametrically

These are smoothed plots of the coefficients on the offer rate indicator variables (y-axis), for each individual offer rate (x-axis), for males. The dependent variable is default, measured as the average past due amount as a proportion of original principal (on the left graph) and as the proportion of months in arrears (on the right graph). Other independent variables include contract rate indicator variables (grouped by integer to preserve degrees of freedom), the lender-defined risk level, and the month of the offer. The flat slope is indicative of no selection effects with respect to the interest rate.

Repayment Burden Effects, Male
Outcome: Monthly Average Proportion Past Due
OLS, Smoothed Plot of Coefficients on Contract Rate
Specification Includes Offer Rate Non-Parametrically

Repayment Burden Effects, Male
Outcome: Proportion of Months in Arrears
OLS, Smoothed Plot of Coefficients on Contract Rate
Specification Includes Offer Rate Non-Parametrically

These are smoothed plots of the coefficients on the contract rate indicator variables (y-axis), for each individual contract rate (x-axis), for males. The dependent variable is default, measured as the average past due amount as a proportion of original principal (on the left graph) and as the proportion of months in arrears (on the right graph). Other independent variables include offer rate indicator variables (grouped by integer to preserve degrees of freedom), the lender-defined risk level, and the month of the offer. The upward slope is indicative of a repayment burden effect with respect to the interest rate.

41

REFERENCES

Armendariz de Aghion, B. and J. Morduch (2005). <u>The Economics of Microfinance</u>, MIT Press.
Ashraf, N., D. Karlan, et al. (2004). "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines." <u>Princeton University working paper</u>.
Ausubel, L. M. (1991). "The Failure of Competition in the Credit Card Market." <u>American Economic Review</u> **81**(1): 50-81.
Ausubel, L. M. (1999). "Adverse Selection in the Credit Card Market." <u>Mimeo, University of Maryland</u>.
Banerjee, A. and E. Duflo (2004). "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program." <u>M.I.T. Working Paper</u>.
Banerjee, A. and A. Newman (1993). "Occupational Choice and the Process of Development." <u>Journal of Political Economy</u> **101**: 274-298.
Bank of Sweden (2001). "Press Release: The 2001 Sveriges Riksbank (Bank of Sweden) Prize in Economic Sciences in Memory of Alfred Nobel, Advanced Information." <u>Available at http://www.nobel.se/economics/laureates/2001/ecoadv.pdf</u>.
Barber, B. and T. Odean (2001). "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment." <u>Quarterly Journal of Economics</u> **116**(1): 261-292.
Bernanke, B. and M. Gertler (1990). "Financial fragility and economic performance." <u>Quarterly Journal of Economics</u> **105**(1): 87-114.
Bertrand, M., D. Karlan, et al. (2005). "How Much Does Psychology Matter? Evidence from a Marketing Experiment in Consumer Credit." <u>mimeo</u>.
Calem, P. S. and L. J. Mester (1995). "Consumer Behavior and the Stickiness of Credit-Card Interest Rates." <u>The American Economic Review</u> **85**(5): 1327-1336.
Chiappori, P.-A. and B. Salanie (2000). "Testing for Asymmetric Information in Insurance Markets." <u>Journal of Political Economy</u> **108**(1): 56-78.
Chiappori, P.-A. and B. Salanie (2003). Testing Contract Theory: A Survey of Some Recent Work. <u>Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress</u>. M. Dewatripont, L. Hansen and P. Turnovsky. Cambridge, Cambridge University Press**:** 115-149.
Cressy, R. and O. Toivanen (2001). "Is there adverse selection in the credit market?" <u>Venture Capital</u> **3**(3): 215-238.
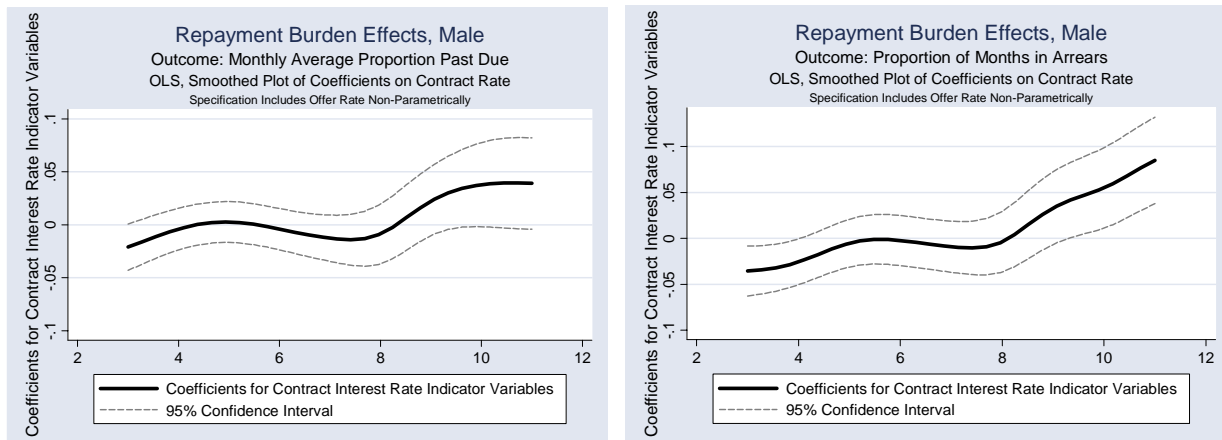Crook, J. (2002). "Adverse Selection and Search in the Bank Credit Card Market." <u>Mimeo, Credit Research Centre, University of Edinburgh</u>.
de Meza, D. and D. Webb (1987). "Too Much Investment: A Problem of Asymmetric Information." <u>The Quarterly Journal of Economics</u> **102**(2): 281-92.
de Meza, D. and D. Webb (2000). "Does credit rationing imply insufficient lending?" <u>Journal of Public Economics</u> **78**(3): 215-34.
de Meza, D. and D. Webb (2001). "Advantageous Selection in Insurance Markets." <u>RAND Journal of Economics</u> **32**(2): 249-62.
Department of Trade and Industry South Africa (2003). "Credit Law Review: Summary of findings of the Technical Committee."
Drake, L. M. and M. J. Holmes (1995). "Adverse selection and the market for consumer credit." <u>Applied Financial Economics</u> **5**: 161-167.
Eaton, J. and M. Gersovitz (1981). "Debt with Potential Repudiation: Theoretical and Empirical Analysis." <u>Review of Economic Studies</u> **48**(2): 289-309.

Edelberg, W. (2004). "Testing for Adverse Selection and Moral Hazard in Consumer Loan Markets." <u>Finance and Economics Discussion Paper Series, Board of Governors of the Federal Reserve System</u> **9**.

Elyasiani, E. and L. G. Goldberg (2004). "Relationship lending: a survey of the literature." <u>Journal of Economics and Business</u> **56**: 315-330.

Finkelstein, A. and K. McGarry (2003). "Private information and its effect on market equilibrium: New evidence from long-term care insurance." <u>Mimeo</u>.

Gale, W. G. (1990). "Federal Lending and the market for credit." <u>Journal of Public Economics</u> **42**(2): 177-193.

Ghosh, P., D. Mookherjee, et al. (2000). "Credit Rationing in Developing Countries: An Overview of the Theory." <u>Readings in the Theory of Economic Development</u> **Chapter 11**: 383-401.

Ghosh, P. and D. Ray (2001). "Information and Enforcement in Informal Credit Markets." <u>mimeo</u>.

Hubbard, R. G. (1998). "Capital Market Imperfections and Investment." <u>Journal of Economic Literature</u> **36**(1): 193-225.

Hurst, E. and A. Lusardi (2004). "Liquidity Constraints, Household Wealth and Entrepreneurship." <u>Journal of Political Economy</u> **112**(2): 319-347.

Karlan, D. (2004). "Social Connections and Group Banking." <u>Princeton University working paper</u>.

Karlan, D. and J. Zinman (2005). "Liquidity Constraints and Demand Elasticities: Evidence from a Randomized Field Experiment in South Africa." <u>Princeton University Mimeo</u>.

Kevane, M. and B. Wydick (2001). "Microenterprise Lending to Female Entrepreneurs: Sacrificing Economic Growth for Poverty Reduction." <u>World Development</u> **29**(7): 1225-1236.

Klonner, S. and A. Rai (2004). "Does Credit Rationing Reduce Defaults? Evidence from South Indian ROSCAs." <u>Working Paper</u>.

Microcredit Summit Campaign (2003). "State of the Microcredit Summit Campaign Report 2002." <u>http://www.microcreditsummit.org/pubs/reports/socr02_en.pdf</u>.

Mookherjee, D. and D. Ray (2004). "Contractual Structure and Wealth Accumulation." <u>American Economic Review</u> **92**(4): 818-49.

Morduch, J. (1994). "Poverty and Vulnerability." <u>American Economic Review (AEA Papers and Proceedings)</u> **84**: 221-225.

Pitt, M. and S. Khandker (1998). "The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter?" <u>The Journal of Political Economy</u> **106**(5): 958-996.

Porteous, D. (2003). "Is Cinderella Finally Coming to the Ball: SA Microfinance in broad perspective." <u>Micro Finance Regulatory Council working paper</u>.

Stiglitz, J. E. and A. Weiss (1981). "Credit Rationing in Markets with Imperfect Information." <u>The American Economic Review</u> **71**(3): 393-410.

Wasmer, E. and P. Weil (2004). "The Macroeconomics of Credit and Labor Market Imperfections." <u>The American Economic Review</u> **94**(4): 944-963.

Table 1a: Summary Statistics

	All	Borrowed	Female Borrowed	Male Borrowed	Did Not Borrow	High Risk	Medium Risk	Low Risk
						Lender-Defined Risk Category		
A. Full Sample								
# of months since last loan	10.3	5.9	6.0	5.8	10.6	12.7	2.8	2.8
	(6.9)	(5.8)	(5.8)	(5.8)	(6.8)	(6.1)	(1.7)	(1.6)
Size of last loan prior to project (Rand)	1116.4	1156.0	1161.4	1150.9	1113.1	1086.4	1176.5	1229.7
	(829.9)	(825.7)	(798.2)	(851.6)	(830.2)	(785.2)	(878.4)	(994.5)
# of prior loans with the lender	4.3	4.9	4.8	4.9	4.2	3.6	5.7	6.6
	(3.9)	(4.2)	(4.2)	(4.2)	(3.8)	(3.5)	(4.2)	(4.3)
Term of last loan prior to project								
1 or 2 months	1,656	132	54	78	1,524	1,407	93	156
	2.88%	3.04%	2.53%	3.52%	2.87%	3.26%	1.50%	1.92%
4 months	53,296	3,939	1,926	2,013	49,357	40,687	5,658	6,951
	92.64%	90.59%	90.30%	90.88%	92.80%	94.18%	91.17%	85.54%
6 months	2,030	223	123	100	1,807	887	369	774
	3.53%	5.13%	5.77%	4.51%	3.40%	2.05%	5.95%	9.52%
12 months	551	54	30	24	497	220	86	245
	0.96%	1.24%	1.41%	1.08%	0.93%	0.51%	1.39%	3.02%
Number of Observations	57,533	4,348	2,133	2,215	53,185	43,201	6,206	8,126
B. Randomized Variables								
Offer Interest Rate	7.88	7.18	7.16	7.22	7.94	8.10	7.20	5.73
	(2.42)	(2.30)	(2.32)	(2.29)	(2.42)	(2.48)	(1.85)	(1.36)
Contract Interest Rate	7.08	6.53	6.46	6.58	7.12	7.29	6.56	5.28
	(2.42)	(2.26)	(2.25)	(2.27)	(2.42)	(2.52)	(1.87)	(1.34)
Proportion Receiving Rate for One year (vs. one loan)	0.43	0.47	0.47	0.47	0.43	0.46	0.47	0.48
	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)
Proportion Receiving a Contract Rate < Offer Rate	0.41	0.40	0.40	0.40	0.41	0.41	0.39	0.39
	(0.49)	(0.49)	(0.49)	(0.49)	(0.49)	(0.49)	(0.49)	(0.49)
C. Default Measure								
Monthly Average Past Due Amount		152.56	131.10	173.21		180.13	224.49	57.40
		(359.28)	(337.39)	(378.09)		(404.86)	(408.52)	(181.67)
Monthly Avg Past Due Amount, Proportion of Principal		0.09	0.08	0.11		0.12	0.13	0.03
		(0.21)	(0.19)	(0.23)		(0.24)	(0.24)	(0.11)
Proportion of Months With Some Arrearage		0.22	0.20	0.24		0.25	0.32	0.10
		(0.29)	(0.28)	(0.30)		(0.31)	(0.31)	(0.19)
Account is in Collection (3+ months arrears)		0.12	0.10	0.14		0.14	0.17	0.04
		(0.32)	(0.30)	(0.33)		(0.35)	(0.38)	(0.19)
Number of Observations	57,533	4,348	2,133	2,215	53,185	2,090	941	1,317

Standard deviations are in parentheses

Money amounts in South African Rand, ~7.5 Rand = US $1 at the time of the experiment.

	Full Sample	Female	Male	Female Borrowed	Male Borrowed
A. Client Characteristics					
Female, proportion	0.48	1	0	1	0
	(0.50)	(0)	(0)	(0)	(0)
Married, proportion	0.44	0.37	0.50	0.39	0.52
	(0.50)	(0.48)	(0.50)	(0.49)	(0.50)
# of dependents	1.59	1.53	1.64	1.82	1.97
	(1.74)	(1.62)	(1.85)	(1.61)	(1.87)
Age	41.25	42.03	40.55	41.74	40.10
	(11.53)	(11.89)	(11.14)	(11.38)	(10.82)
Education (# of years, estimated from occupation)	6.78	7.23	6.36	7.45	6.53
	(3.32)	(3.45)	(3.14)	(3.51)	(3.19)
Monthly gross income at last loan (000's Rand)*	3.42	3.26	3.56	3.39	3.45
	(19.66)	(2.63)	(27.05)	(2.19)	(2.07)
Home bond, proportion	0.07	0.07	0.06	0.08	0.06
	(0.25)	(0.25)	(0.24)	(0.26)	(0.24)
External credit score	551.35	544.23	557.82	547.77	571.69
	(215.64)	(210.22)	(220.27)	(203.20)	(204.22)
No external credit score, proportion	0.12	0.11	0.12	0.11	0.10
	(0.32)	(0.32)	(0.33)	(0.31)	(0.30)
Months at Employer	93.82	90.42	96.92	93.34	96.86
	(88.01)	(82.55)	(92.59)	(82.33)	(88.53)
# of Observations	57533	27387	30146	2133	2215
B. Loan Characteristics					
Amount of last loan prior to experiment	1116.36	1122.87	1110.44	1161.37	1150.86
	(829.90)	(844.42)	(816.46)	(798.21)	(851.56)
Term of last loan prior to experiment	4.06	4.09	4.03	4.15	4.07
	(1.00)	(1.01)	(1.00)	(1.16)	(1.09)
# of prior loans with the lender	4.26	4.22	4.29	4.83	4.90
	(3.86)	(3.82)	(3.90)	(4.20)	(4.26)
# of months since the last loan	10.26	10.21	10.31	5.98	5.82
	(6.88)	(6.84)	(6.92)	(5.78)	(5.82)
Internal credit score when new borrower	29.66	32.59	26.99	32.97	27.40
	(8.75)	(8.53)	(8.06)	(8.38)	(8.22)
# of Observations	57533	27387	30146	2133	2215
C. Self-Reported Loan Usage					
School				24.2%	13.6%
Housing (mostly renovations)				12.6%	9.8%
Payoff other debt				10.9%	11.1%
Family/Event				5.7%	8.1%
Consumption				5.6%	7.1%
Transport				4.1%	7.6%
Funeral/Medical				3.8%	4.4%
Durable				2.3%	1.0%
Business/Other Investment				2.3%	2.7%
Misc/unreported				28.7%	34.6%
# of Observations				690	775

* Standard deviations are in parentheses. Gross income at time of last loan is missing for participants from pilot phase. Age, gender and other demographic information also missing for <10 observations. Number of observations reported is the total number, irrespective of missing data. Usage sample size is low relative to takeup due to reluctance of loan officers to administer survey (the Lender does not typically ask applicants about intended usage, and if anything emphasizes that it does not ask such questions). Reported "Consumption" uses are primarily food (39%) and clothing (23%); "Family/Events" are largely Christmas (45%) expenses; "School" is largely the fees required for children to attend; "Misc" is largely borrowers declining to specify (88%).

Table 2: Experimental Integrity Checks
OLS

					Sample Restricted to Applied = 1
			Rate Valid for One Year (versus One Loan)		
Dependent variable:	Contract Rate	Offer Rate	Loan)	Applied=1	Rejected = 1
	(1)	(2)	(3)	(4)	(5)
Female	0.009	0.028	-0.002		
	(0.022)	(0.021)	(0.004)		
Married	0.017	0.022	0.004		
	(0.022)	(0.021)	(0.004)		
External credit score	-0.000	-0.000	0.000		
	(0.000)	(0.000)	(0.000)		
No External credit score	-0.017	-0.006	0.016		
	(0.093)	(0.091)	(0.016)		
Internal credit score	-0.001	-0.002	0.000		
	(0.001)	(0.001)	(0.000)		
Log (Size of last loan prior to project)	-0.017	-0.003	-0.004		
	(0.017)	(0.017)	(0.003)		
Term of last loan prior to project	-0.010	-0.011	-0.001		
	(0.011)	(0.010)	(0.002)		
# of prior loans with the lender	0.003	0.003	0.001**		
	(0.003)	(0.003)	(0.001)		
Gross income	-0.001	-0.000	0.000		
	(0.001)	(0.000)	(0.000)		
Years at Employer	0.000	0.001	-0.000		
	(0.002)	(0.002)	(0.000)		
Mean education	0.002	-0.002	-0.000		
	(0.003)	(0.003)	(0.001)		
# of dependants	0.002	-0.005	0.000		
	(0.007)	(0.006)	(0.001)		
Age	-0.000	-0.001	-0.000*		
	(0.001)	(0.001)	(0.000)		
Home bond	0.053	0.028	0.011		
	(0.041)	(0.040)	(0.007)		
# of months since last loan	-0.001	-0.001	-0.001***		
	(0.002)	(0.002)	(0.000)		
Offer Interest Rate				-0.003***	
				(0.001)	
Contract Interest Rate				0.000	-0.001
				(0.001)	(0.002)
Dynamic Repayment Incentive					-0.014
					(0.012)
Constant	7.700***	8.369***	0.228***	0.081***	0.334***
	(0.297)	(0.292)	(0.051)	(0.005)	(0.075)
Observations	57339	57339	57339	57533	5028
Joint F-Test	0.87	0.96	0.01		
R-squared	0.10	0.14	0.37	0.04	0.09

 * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors in parentheses. Columns 1 through 3 test whether the randomized variables are correlated with information observable before the experiment launch. For column 3, if the dormancy variable is omitted the F-test is 0.21. Column 4 shows that the decision to borrow by the client was affected by the Offer Interest Rate, but not the Contract Interest Rate, hence verifying the internal controls of the experimental protocol. Column 5 shows that the decision by the branch manager to reject applicants was not predicted by the contract interest rate or the dynamic repayment incentive. Column 5 sample frame includes only those who applied for a loan. Regressions include controls for lender-defined risk category, month of offer letter and branch.

Table 3: Disentangling Selection on Unobservables from Repayment Burden Effects: Comparison of Means

		Moral Hazard / Repayment Burden Effects			Selection Effects		
		High Offer, High Contract	High Offer, Low Contract	t-stat: diff≠0	High Offer, Low Contract	Low Offer, Low Contract	t-stat: diff≠0
Full Sample							
	Average Monthly Proportion Past Due	0.105	0.102	0.23	0.102	0.082	1.90*
		(0.006)	(0.009)		(0.009)	(0.004)	
	Proportion of Months in Arrears	0.244	0.211	2.38**	0.211	0.202	0.72
		(0.008)	(0.011)		(0.011)	(0.006)	
	Account in Collection Status	0.139	0.123	0.99	0.123	0.101	1.50
		(0.009)	(0.013)		(0.013)	(0.007)	
	# of observations	1636	625		625	2087	
Female							
	Average Monthly Proportion Past Due	0.089	0.101	-0.85	0.101	0.067	2.42**
		(0.007)	(0.013)		(0.013)	(0.005)	
	Proportion of Months in Arrears	0.221	0.209	0.64	0.209	0.181	1.55
		(0.011)	(0.02)		(0.02)	(0.008)	
	Account in Collection Status	0.107	0.121	-0.65	0.121	0.082	1.88*
		(0.121)	(0.019)		(0.019)	(0.008)	
	# of observations	779	307		307	1047	
Male							
	Average Monthly Proportion Past Due	0.120	0.103	1.05	0.103	0.099	0.30
		(0.008)	(0.013)		(0.013)	(0.007)	
	Proportion of Months in Arrears	0.264	0.213	2.60***	0.213	0.223	-0.51
		(0.011)	(0.016)		(0.016)	(0.009)	
	Account in Collection Status	0.168	0.126	1.87*	0.126	0.120	0.26
		(0.013)	(0.019)		(0.019)	(0.010)	
	# of observations	857	318		318	1040	

In this stylized demonstration of the results, "high" is defined as above the median offer rate for that risk category. This is equal to 7.77% for high risk clients, 7.50% for medium risk clients and 6.00% for low risk clients. T-tests assume unequal variances across columns.

Table 4: Full Sample Results: Disentangling Selection on Unobservables from Moral Hazard

	OLS						Tobit		Probit
Dependent Variable:	*Monthly Average Proportion Past Due*		*Proportion of Months in Arrears*		*Account in Collection Status*		*Monthly Average Proportion Past Due*	*Proportion of Months in Arrears*	*Account in Collection Status*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Offer Rate (Selection)	0.004	0.004	0.002	0.002	0.007	0.007	0.004	0.002	0.007
	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.007)	(0.004)
Contract Rate (Repayment Burden)	-0.000	-0.002	0.007*	0.003	0.001	-0.001	0.002	0.010	0.001
	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.004)	(0.007)	(0.004)
Dynamic Repayment Incentive Dummy (Moral Hazard)	-0.011*	0.003	-0.016**	0.013	-0.019**	0.000	-0.021**	-0.032**	-0.019**
	(0.005)	(0.011)	(0.008)	(0.018)	(0.009)	(0.019)	(0.009)	(0.016)	(0.008)
Dynamic Repayment Incentive Size (Moral Hazard)		-0.004		-0.008**		-0.005			
		(0.003)		(0.004)		(0.004)			
Constant	0.079***	0.094***	0.139***	0.171***	0.069***	0.090***	-0.065**	-0.081	
	(0.014)	(0.019)	(0.025)	(0.027)	(0.024)	(0.028)	(0.028)	(0.052)	
Observations	4348	4348	4348	4,348	4348	4348	4348	4348	4348
Adjusted R-squared	0.04	0.04	0.11	0.11	0.03	0.03			
Mean of dependent variable	0.09	0.09	0.22	0.22	0.12	0.12	0.09	0.22	0.12
Prob(both Dynamic Incentive variables = 0)		0.08*		0.01***		0.05**			

* significant at 10%; ** significant at 5%; *** significant at 1%. Each column presents results from a single model estimated using the base specification (equation 2). All probit results are marginal effects. Robust standard errors in parentheses are corrected for clustering at the branch level. "Offer Rate" and "Contract Rate" are in monthly percentage point units (7.00% interest per month is coded as 7.00). "Dynamic Repayment Incentive" is an indicator variable equal to one if the contract interest rate is valid for one year (rather than just one loan) before reverting back to the normal (higher) interest rates. "Dynamic Repayment Incentive Size" interacts the above indicator variable with the difference between the Lender's normal rate for that individual's risk category and the experimentally assigned contract interest rate. All models include controls for lender-defined risk category and month of offer letter. Including controls for loan size and term does not affect the results. A positive coefficient on the Offer Rate variable indicates adverse selection, a positive coefficient on the Contract Rate variable indicates a reduced-form repayment burden effect, and a negative coefficient on the Dynamic Repayment Incentive variables indicate moral hazard that is alleviated by the dynamic pricing incentive.

Table 5: Disentangling Selection on Unobservables from Moral Hazard
By Gender

Dependent Variable:	Male						Female					
	OLS			Tobit		Probit	OLS			Tobit		Probit
	Monthly Average Proportion Past Due	Proportion of Months in Arrears	Account in Collection Status	Monthly Average Proportion Past Due	Proportion of Months in Arrears	Account in Collection Status	Monthly Average Proportion Past Due	Proportion of Months in Arrears	Account in Collection Status	Monthly Average Proportion Past Due	Proportion of Months in Arrears	Account in Collection Status
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Offer Rate	-0.002	-0.004	0.001	-0.005	-0.008	0.001	0.010***	0.008*	0.013**	0.013**	0.013	0.011***
	(0.004)	(0.005)	(0.007)	(0.006)	(0.010)	(0.007)	(0.003)	(0.005)	(0.005)	(0.006)	(0.009)	(0.004)
Contract Rate	0.005	0.014***	0.010	0.010	0.021**	0.009	-0.005	-0.001	-0.009	-0.006	-0.003	-0.007*
	(0.003)	(0.005)	(0.007)	(0.006)	(0.010)	(0.006)	(0.004)	(0.005)	(0.006)	(0.006)	(0.009)	(0.004)
Dynamic Repayment Incentive Indicator	-0.014	-0.025**	-0.020	-0.033**	-0.050**	-0.020	-0.007	-0.006	-0.017	-0.008	-0.009	-0.016
	(0.009)	(0.012)	(0.015)	(0.016)	(0.023)	(0.014)	(0.008)	(0.012)	(0.012)	(0.014)	(0.025)	(0.011)
Constant	0.108***	0.178***	0.092**	-0.008	0.002		0.050***	0.097***	0.043	-0.121***	-0.176***	
	(0.025)	(0.040)	(0.043)	(0.043)	(0.072)		(0.015)	(0.026)	(0.027)	(0.029)	(0.057)	
Observations	2215	2215	2215	2215	2215	2215	2133	2133	2133	2133	2133	2133
R-squared	0.05	0.12	0.04				0.05	0.10	0.04			

 * significant at 10%; ** significant at 5%; *** significant at 1%. All probit results are marginal effects. Robust standard errors in parentheses are corrected for clustering at the branch level. Results reported here are estimated using the base specification (equation 2) on samples split by gender. They include controls for lender-defined risk category and month of offer letter. Including controls for loan size and term does not affect the results.

Table 6: Disentangling Selection on Unobservables from Moral Hazard
OLS: Matching Specification

	Full sample			Male			Female			# of Fixed Effects for All Combinations of Other Random Variables (see notes)
Dependent Variable:	*Monthly Average Proportion Past Due*	*Proportion of Months in Arrears*	*Account in Collection Status*	*Monthly Average Proportion Past Due*	*Proportion of Months in Arrears*	*Account in Collection Status*	*Monthly Average Proportion Past Due*	*Proportion of Months in Arrears*	*Account in Collection Status*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
Offer Rate	0.0027	0.0004	0.0046	-0.0035	-0.0064	-0.0028	0.0092**	0.0067	0.0128**	167
	(0.0026)	(0.0048)	(0.0049)	(0.0040)	(0.0062)	(0.0073)	(0.0038)	(0.0052)	(0.0062)	
Contract Rate	-0.0001	0.0062*	0.0009	0.0052	0.0128***	0.0091	-0.0056	-0.0023	-0.0090	165
	(0.0025)	(0.0034)	(0.0047)	(0.0040)	(0.0049)	(0.0075)	(0.0035)	(0.0051)	(0.0056)	
Dynamic Repayment Incentive Indicator	-.0120**	-0.019**	-0.023**	-0.013	-0.022	-0.023	-.006	-0.007	-0.018*	620
	(0.006)	(0.010)	(0.010)	(0.010)	(0.015)	(0.018)	(0.008)	(0.014)	(0.011)	
# of Observations		4348			2,215			2133		

 * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors in parentheses are corrected for clustering at the branch level.

Each cell represents a separate OLS specification. For each cell, we report the coefficient on the independent variable named in the row heading obtaining from including a full set of indicator variables for all possible combinations of the other independent random variables and lender-defined risk category. For example, the top left cell reports the coefficient on the Offer Rate after including fixed effects for every combination of a set of indicator variables for the Contract Rate, an indicator for the Dynamic Repayment Incentive, and two indicator variables for the lender-defined risk category. All specifications include controls for month of offer letter.

Table 7: Disentangling Selection on Unobservables from Moral Hazard

Instrumental Variables

Sample:	Full	Full	Full	Male	Male	Male	Female	Female	Female
Dependent Variable:	*Average Monthly Proportion Past Due*	*Proportion of Months in Arrears*	*Account in Collection Status*	*Average Monthly Proportion Past Due*	*Proportion of Months in Arrears*	*Account in Collection Status*	*Average Monthly Proportion Past Due*	*Proportion of Months in Arrears*	*Account in Collection Status*
	(1)	(2)	(3)	(4)	(6)	(7)	(8)	(10)	(11)
Panel A: IV									
Total Interest Cost, 000's Rands, at Offer Interest Rate	0.35*	0.29	0.66*	-0.03	-0.00	0.41	0.74***	0.65*	1.00**
	(0.20)	(0.28)	(0.35)	(0.29)	(0.42)	(0.51)	(0.27)	(0.35)	(0.41)
Total Interest Cost, 000's Rands, at Contract Interest Rate	-0.00	0.10*	0.01	0.07	0.19**	0.14	-0.09	-0.02	-0.14
	(0.04)	(0.06)	(0.08)	(0.05)	(0.08)	(0.11)	(0.06)	(0.09)	(0.10)
Constant	0.09***	0.19***	0.09***	0.10***	0.21***	0.09**	0.08***	0.21***	0.12***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
Observations	4348	4348	4348	2215	2215	2215	2133	2133	2133
Panel B: OLS									
Total Interest Cost, 000's Rands, at Offer Interest Rate	-0.12**	0.01	-0.08	-0.15*	0.06	0.06	-0.08	-0.02	-0.19
	(0.06)	(0.10)	(0.10)	(0.08)	(0.14)	(0.14)	(0.06)	(0.12)	(0.13)
Total Interest Cost, 000's Rands, at Contract Interest Rate	-0.00	0.01	0.01	0.00	0.01	0.01	-0.01	0.01	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
Constant	0.13***	0.25***	0.15***	0.13***	0.27***	0.17***	0.12***	0.26***	0.15***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	4348	4348	4348	2215	2215	2215	2133	2133	2133
	0.05	0.10	0.03	0.05	0.11	0.03	0.05	0.10	0.03

 * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors in parentheses are corrected for clustering at the branch level. The two endogenous variables (total interest cost evaluated separately at the offer and contract rate) are instrumented with the offer rate and contract rate (please see equations 6-8 in the text). Regressions include controls for lender-defined risk category and month of offer letter. Including controls for loan size and term does not affect the results.

Table 8: Disentangling Selection on Unobservables from Moral Hazard, by Gender & Demographics

OLS, Dependent Variable: Monthly Average Percentage Past Due

Demographic Control Variable:	Married		Dependents in Household		Educated		Age		Log(Monthly Gross Income)		Tenure at Employment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Experimental Variables												
Offer Rate	0.023	-0.252	0.089	0.176	0.079	-0.213	0.282	-0.341	2.700	-0.716	0.122	0.082
	(0.435)	(0.537)	(0.432)	(0.530)	(0.402)	(0.410)	(1.162)	(1.325)	(2.338)	(4.218)	(0.456)	(0.495)
Contract Rate	0.415	0.716	0.482	0.660	0.260	0.557	0.269	0.652	-0.968	1.852	0.404	0.442
	(0.393)	(0.508)	(0.446)	(0.527)	(0.414)	(0.440)	(1.098)	(1.283)	(2.707)	(4.600)	(0.465)	(0.523)
Dynamic Repayment Incentive Indicator	-1.158	-0.706	-1.098	-1.920	-0.878	-1.081	-1.280	0.934	7.378	3.457	-1.165	-0.691
	(1.160)	(1.510)	(1.237)	(1.434)	(1.028)	(1.252)	(2.678)	(4.049)	(8.692)	(12.814)	(1.145)	(1.465)
Female	-2.985	-3.095	-2.558	-1.021	-2.215	-2.652	-1.887	-5.296	-2.821	-30.918	-2.667	-2.298
	(1.939)	(2.585)	(1.980)	(3.110)	(1.886)	(2.631)	(1.914)	(7.409)	(1.926)	(28.386)	(1.875)	(3.073)
Demographic Variable (see column heading)	-1.838	-2.040	-0.036	0.427	-1.761	-2.487	-0.172	-0.223	-0.001	-2.020	-0.015	-0.013
	(1.952)	(2.854)	(0.536)	(0.741)	(2.432)	(3.909)	(0.105)	(0.157)	(1.669)	(3.019)	(0.012)	(0.018)
*Female * Experimental Variables*												
Female * Offer Rate	0.887*	1.369**	0.834*	0.637	0.902*	1.534**	0.763*	1.951	0.890**	6.945	0.807*	0.891
	(0.456)	(0.632)	(0.460)	(0.661)	(0.480)	(0.604)	(0.455)	(1.916)	(0.445)	(6.117)	(0.447)	(0.749)
Female * Contract Rate	-1.042**	-1.575**	-1.029**	-1.440**	-1.138**	-1.783***	-0.977**	-1.782	-1.040**	-6.318	-0.967**	-1.047
	(0.476)	(0.718)	(0.497)	(0.678)	(0.482)	(0.640)	(0.486)	(1.979)	(0.474)	(7.129)	(0.479)	(0.748)
Female * Dynamic Repayment Incentive	0.813	-0.037	0.896	2.732	1.077	1.554	0.701	-3.491	0.603	8.148	0.730	-0.290
	(1.350)	(2.143)	(1.343)	(2.052)	(1.351)	(1.903)	(1.336)	(5.867)	(1.353)	(14.026)	(1.328)	(2.363)
*Demographic Control Variable * Experimental Variables*												
Demographic Variable * Offer Rate	-0.135	0.415	-0.046	-0.084	-0.400	0.626	-0.008	0.008	-0.343	0.079	-0.002	-0.001
	(0.540)	(0.796)	(0.122)	(0.164)	(0.625)	(0.853)	(0.026)	(0.030)	(0.289)	(0.522)	(0.003)	(0.003)
Demographic Variable * Contract Rate	0.195	-0.397	-0.009	-0.124	0.748	-0.279	0.006	-0.003	0.183	-0.166	0.001	0.001
	(0.511)	(0.788)	(0.141)	(0.177)	(0.583)	(0.776)	(0.026)	(0.031)	(0.325)	(0.561)	(0.003)	(0.003)
Demographic Variable * Dynamic Repayment Incentive	-0.577	-1.442	-0.224	0.162	-1.577	-1.017	-0.002	-0.056	-1.077	-0.592	-0.002	-0.007
	(1.211)	(1.897)	(0.353)	(0.431)	(1.307)	(2.104)	(0.061)	(0.092)	(1.042)	(1.530)	(0.006)	(0.009)
*Female * Demographic Control Variable*												
Female * Demographic Variable		0.305		-1.217		1.155		0.083		3.457		-0.004
		(3.234)		(1.138)		(4.167)		(0.167)		(3.460)		(0.023)
*Female * Demographic Control Variable * Experimental Variables*												
Female * Demographic Variable * Offer Rate		-1.079		0.111		-1.755		-0.029		-0.748		-0.001
		(0.951)		(0.274)		(1.080)		(0.044)		(0.758)		(0.006)
Female * Demographic Variable * Contract Rate		1.181		0.277		1.777		0.020		0.654		0.001
		(1.033)		(0.292)		(1.196)		(0.045)		(0.882)		(0.005)
Female * Demographic Variable * Dynamic Repayment Incentive		1.797		-0.968		-1.049		0.102		-0.937		0.011
		(2.652)		(0.654)		(2.713)		(0.125)		(1.672)		(0.015)
Constant	10.161***	10.236***	8.917***	8.252***	9.608***	9.821***	14.984***	17.066**	9.240	25.704	10.281***	10.122***
	(2.476)	(2.791)	(2.542)	(2.986)	(2.240)	(2.546)	(5.136)	(7.222)	(13.856)	(25.009)	(2.642)	(3.133)
Observations	4317	4317	4317	4317	4348	4348	4348	4348	4348	4348	4348	4348
R-squared	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.06	0.05	0.05	0.06	0.06

* significant at 10%; ** significant at 5%; *** significant at 1%. Each column presents results from a single OLS regression on a version of equation (2). Robust standard errors in parentheses are corrected for clustering at the branch level. "Educated" is a binary indicator for the top 25% in years of education, predicted by the client's occupation. Regressions include controls for lender-defined risk category and month of offer letter. Including controls for loan size and term does not affect the results. The dependent variable here is defined in percentage point terms, not proportions, and hence equals 100x the variable used in other tables.

Table 9: Selection on Observable Information
Probit, Dependent Variable: "Applied for Loan"

	Sample:	All (1)	All (2)	Female (3)	Male (4)
Offer Rate		-0.0029***	-0.004***	-0.006***	-0.002
		(0.0004)	(0.001)	(0.002)	(0.002)
Predicted Past Due Percentage			-158.064**	-268.010**	-17.186
			(76.012)	(122.971)	(112.913)
Offer Rate* Predicted Past Due Percentage			10.071	23.443*	-4.948
			(8.852)	(13.105)	(13.786)
Observations		52985	52985	25221	27764

 * significant at 10%; ** significant at 5%; *** significant at 1%. Each column presents results from a probit version of equation (2). All results are marginal effects. Robust standard errors in parentheses are corrected for clustering at the branch level. Predicted past due percentage is the predicted value for default from Table 10 OLS specifications. 4548 observations omitted because they were either missing income data for the non-applicants, received an upward-sloped yield curve on their offer letter (see Karlan and Zinman 2005), or had their letter returned by the postal system due to a bad address. Regressions include controls for lender-defined risk category and month of offer letter.

Table 10: Including Observable Determinants of Default

OLS

Dependent Variable:	Monthly Average Proportion Past Due		Proportion of Months in Arrears		Account in Collection Status	
	(1)	(2)	(3)	(4)	(5)	(6)
Offer Rate	-0.001		-0.003		0.003	
	(0.003)		(0.005)		(0.006)	
Contract Rate	0.005		0.014***		0.010	
	(0.003)		(0.005)		(0.007)	
Dynamic Repayment Incentive Indicator	-0.017*		-0.024**		-0.022	
	(0.010)		(0.012)		(0.016)	
Female * Offer Rate	0.007*		0.008		0.007	
	(0.004)		(0.006)		(0.007)	
Female * Contract Rate	-0.009**		-0.015**		-0.017**	
	(0.005)		(0.007)		(0.008)	
Female * Dynamic Repayment Incentive	0.008		0.014		0.003	
	(0.013)		(0.018)		(0.021)	
Female	-0.015	-0.021***	-0.005	-0.035***	0.033	-0.029**
	(0.019)	(0.007)	(0.026)	(0.010)	(0.027)	(0.012)
Log(loan size)	-0.026***	-0.026***	0.013*	0.013*	0.004	0.004
	(0.005)	(0.005)	(0.007)	(0.007)	(0.008)	(0.008)
Age	0.000	0.000	0.002	0.001	0.002	0.002
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Age squared	-0.000	-0.000	-0.000	-0.000	-0.000*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Years at Employer	-0.001	-0.001	-0.001**	-0.001**	-0.002*	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Gross Income	0.003	0.003	-0.007*	-0.007*	-0.006	-0.005
	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)
Education (predicted by occupation)	-0.001	-0.001	-0.001	-0.002	-0.002	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
# of Dependents	-0.001	-0.001	0.000	0.000	-0.006*	-0.006**
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
External Credit Score	-0.000***	-0.000***	-0.000***	-0.000***	-0.000*	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No External Credit Score	-0.097***	-0.100***	-0.244***	-0.251***	-0.075*	-0.082*
	(0.035)	(0.034)	(0.049)	(0.049)	(0.045)	(0.044)
Internal Credit Score at First-Time Application	-0.001*	-0.001*	-0.001*	-0.001**	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Married	0.002	0.003	0.005	0.005	0.014	0.015
	(0.007)	(0.007)	(0.009)	(0.009)	(0.012)	(0.012)
Home Bond	0.010	0.009	0.014	0.012	0.041*	0.038*
	(0.014)	(0.014)	(0.021)	(0.022)	(0.023)	(0.022)
# of prior loans with the lender	-0.003***	-0.003***	-0.005***	-0.005***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
# of months since last loan	0.004***	0.004***	0.004**	0.004**	0.005***	0.005***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.466***	0.488***	0.412***	0.486***	0.277***	0.368***
	(0.069)	(0.068)	(0.087)	(0.080)	(0.100)	(0.089)
Observations	4348	4348	4348	4348	4348	4348
R-squared	0.0886	0.0862	0.1570	0.1520	0.0711	0.0660
Adjusted r-squared	0.0808	0.0796	0.1497	0.1459	0.0631	0.0593

 * significant at 10%; ** significant at 5%; *** significant at 1%. Each column presents results from a single OLS regression on a version of equation (2). Robust standard errors in parentheses are corrected for clustering at the branch level. Control dummies are included for Lender's 14-category risk level (the Lender uses 3 categories for pricing and term, and 11 categories for credit approval and loan size decisions, internal tracking & reporting).

Table 11: Do Information Asymmetries Diminish for Clients with More Frequent Borrowing History?

OLS, Dependent Variable: Monthly Average Proportion Past Due

Sample:	All			Female			Male		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Offer Rate	0.008**	0.004	0.004	0.014***	0.010***	0.010***	0.001	-0.002	-0.002
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)
Contract Rate	0.000	0.004	0.000	-0.005	-0.001	-0.005	0.005	0.008*	0.005
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)
Dynamic Repayment Incentive Indicator	-0.011*	-0.011*	-0.013	-0.007	-0.007	-0.006	-0.015	-0.015	-0.023
	(0.006)	(0.006)	(0.010)	(0.008)	(0.008)	(0.011)	(0.010)	(0.010)	(0.017)
# of prior loans with the lender	0.001	0.000		0.003	0.001		-0.001	-0.001	
	(0.002)	(0.001)		(0.002)	(0.002)		(0.003)	(0.003)	
Offer Rate*# of prior loans	-0.001***			-0.001***			-0.001		
	(0.000)			(0.000)			(0.000)		
Contract Rate*# of prior loans		-0.001***			-0.001***			-0.001	
		(0.000)			(0.000)			(0.000)	
Rate Valid for One Year*# of prior loans			0.001			-0.000			0.002
			(0.001)			(0.001)			(0.002)
Constant	0.078***	0.083***	0.105***	0.040**	0.046***	0.070***	0.119***	0.121***	0.142***
	(0.018)	(0.017)	(0.014)	(0.017)	(0.016)	(0.015)	(0.034)	(0.032)	(0.025)
Observations	4317	4317	4317	2119	2119	2119	2198	2198	2198
R-squared	0.05	0.05	0.05	0.06	0.06	0.05	0.06	0.06	0.06

 * significant at 10%; ** significant at 5%; *** significant at 1%. Each column presents results from a single OLS regression on a version of equation (2). Robust standard errors in parentheses are corrected for clustering at the branch level. Regressions include controls for lender-defined risk category and month of offer letter. Including controls for loan size and term does not affect the results.

Appendix Table 1. Frequency of Monthly Offer and Contract Interest Rates

| | Low Risk Clients | | | | Medium Risk Clients | | | | High Risk Clients | | | |
| | Offer Interest Rate | | Contract Interest Rate | | Offer Interest Rate | | Contract Interest Rate | | Offer Interest Rate | | Contract Interest Rate | |
	Freq.	Percent	Freq.	Percent	Freq.	Percent	Freq.	Percent	Freq.	Percent	Freq.	Percent
3.25%	144	1.77%	304	3.74%	94	1.51%	172	2.77%	586	1.36%	1,017	2.35%
3.49%	281	3.46%	347	4.27%	110	1.77%	135	2.18%	756	1.75%	934	2.16%
3.50%	267	3.29%	393	4.84%	116	1.87%	163	2.63%	540	1.25%	931	2.16%
3.75%	32	0.39%	42	0.52%	18	0.29%	26	0.42%	53	0.12%	80	0.19%
3.99%	367	4.52%	580	7.14%	104	1.68%	229	3.69%	754	1.75%	1,400	3.24%
4.00%	199	2.45%	341	4.20%	99	1.60%	144	2.32%	525	1.22%	845	1.96%
4.25%	40	0.49%	61	0.75%	22	0.35%	29	0.47%	59	0.14%	69	0.16%
4.44%	208	2.56%	380	4.68%	79	1.27%	214	3.45%	494	1.14%	1,220	2.82%
4.49%	399	4.91%	330	4.06%	139	2.24%	136	2.19%	775	1.79%	866	2.00%
4.50%	176	2.17%	288	3.54%	99	1.60%	149	2.40%	591	1.37%	826	1.91%
4.75%	45	0.55%	39	0.48%	22	0.35%	29	0.47%	60	0.14%	77	0.18%
4.99%	202	2.49%	378	4.65%	117	1.89%	211	3.40%	713	1.65%	1,347	3.12%
5.00%	283	3.48%	332	4.09%	119	1.92%	168	2.71%	550	1.27%	809	1.87%
5.25%	45	0.55%	49	0.60%	19	0.31%	26	0.42%	67	0.16%	77	0.18%
5.49%	338	4.16%	387	4.76%	149	2.40%	239	3.85%	712	1.65%	1,330	3.08%
5.50%	426	5.24%	415	5.11%	97	1.56%	144	2.32%	604	1.40%	761	1.76%
5.55%	288	3.54%	267	3.29%	81	1.31%	120	1.93%	513	1.19%	660	1.53%
5.75%	46	0.57%	56	0.69%	20	0.32%	27	0.44%	74	0.17%	92	0.21%
5.99%	495	6.09%	409	5.03%	213	3.43%	259	4.17%	712	1.65%	1,175	2.72%
6.00%	402	4.95%	315	3.88%	118	1.90%	141	2.27%	586	1.36%	766	1.77%
6.25%	49	0.60%	51	0.63%	24	0.39%	25	0.40%	74	0.17%	80	0.19%
6.50%	388	4.77%	377	4.64%	125	2.01%	201	3.24%	611	1.41%	1,286	2.98%
6.75%	422	5.19%	335	4.12%	148	2.38%	198	3.19%	569	1.32%	903	2.09%
6.99%	464	5.71%	308	3.79%	231	3.72%	192	3.09%	775	1.79%	903	2.09%
7.00%	435	5.35%	292	3.59%	201	3.24%	194	3.13%	855	1.98%	881	2.04%
7.25%	399	4.91%	273	3.36%	200	3.22%	205	3.30%	834	1.93%	1,028	2.38%
7.49%	575	7.08%	347	4.27%	260	4.19%	212	3.42%	1,015	2.35%	977	2.26%
7.50%	357	4.39%	229	2.82%	195	3.14%	166	2.67%	849	1.97%	825	1.91%
7.75%	354	4.36%	201	2.47%	181	2.92%	162	2.61%	909	2.10%	1,033	2.39%
7.77%	-	-	-	-	200	3.22%	138	2.22%	825	1.91%	719	1.66%
7.99%	-	-	-	-	224	3.61%	159	2.56%	1,029	2.38%	933	2.16%
8.00%	-	-	-	-	168	2.71%	160	2.58%	891	2.06%	830	1.92%
8.19%	-	-	-	-	235	3.79%	167	2.69%	1,024	2.37%	829	1.92%
8.25%	-	-	-	-	25	0.40%	28	0.45%	74	0.17%	79	0.18%
8.50%	-	-	-	-	215	3.46%	164	2.64%	830	1.92%	984	2.28%
8.75%	-	-	-	-	35	0.56%	23	0.37%	82	0.19%	77	0.18%
8.88%	-	-	-	-	221	3.56%	153	2.47%	805	1.86%	851	1.97%
8.99%	-	-	-	-	263	4.24%	174	2.80%	1,044	2.42%	814	1.88%
9.00%	-	-	-	-	214	3.45%	128	2.06%	877	2.03%	756	1.75%
9.25%	-	-	-	-	218	3.51%	145	2.34%	890	2.06%	867	2.01%
9.49%	-	-	-	-	300	4.83%	170	2.74%	1,162	2.69%	879	2.03%
9.50%	-	-	-	-	37	0.60%	28	0.45%	89	0.21%	82	0.19%
9.69%	-	-	-	-	234	3.77%	137	2.21%	1,201	2.78%	892	2.06%
9.75%	-	-	-	-	217	3.50%	116	1.87%	889	2.06%	727	1.68%
9.99%	-	-	-	-	-	-	-	-	1,242	2.87%	887	2.05%
10.00%	-	-	-	-	-	-	-	-	1,253	2.90%	876	2.03%
10.25%	-	-	-	-	-	-	-	-	1,276	2.95%	892	2.06%
10.49%	-	-	-	-	-	-	-	-	1,494	3.46%	964	2.23%
10.50%	-	-	-	-	-	-	-	-	1,282	2.97%	833	1.93%
10.75%	-	-	-	-	-	-	-	-	93	0.22%	73	0.17%
10.99%	-	-	-	-	-	-	-	-	1,390	3.22%	899	2.08%
11.00%	-	-	-	-	-	-	-	-	1,385	3.21%	857	1.98%
11.11%	-	-	-	-	-	-	-	-	1,345	3.11%	800	1.85%
11.19%	-	-	-	-	-	-	-	-	1,498	3.47%	867	2.01%
11.25%	-	-	-	-	-	-	-	-	104	0.24%	77	0.18%
11.50%	-	-	-	-	-	-	-	-	99	0.23%	72	0.17%
11.69%	-	-	-	-	-	-	-	-	1,431	3.31%	834	1.93%
11.75%	-	-	-	-	-	-	-	-	1,382	3.20%	753	1.74%
Total	8,126	100%	8,126	100%	6,206	100%	6,206	100%	43,201	100%	43,201	100%

Appendix Table 2: Table 4 Results using a Common Support
Sample Restricted to Those Offered ≤ 7.75%, the Ceiling for Low Risk Clients

	OLS			Tobit		Probit
Dependent Variable:	*Monthly Average Proportion Past Due*	*Proportion of Months in Arrears*	*Account in Collection Status*	*Monthly Average Proportion Past Due*	*Proportion of Months in Arrears*	*Account in Collection Status*
	(1)	(2)	(3)	(4)	(5)	(6)
Offer Rate	0.007	0.005	0.012	0.007	0.004	0.011
	(0.005)	(0.006)	(0.008)	(0.009)	(0.014)	(0.007)
Contract Rate	-0.006	0.001	-0.003	-0.006	0.002	-0.003
	(0.005)	(0.006)	(0.008)	(0.009)	(0.014)	(0.007)
Dynamic Repayment Incentive Indicator	-0.010*	-0.021**	-0.017**	-0.023**	-0.040**	-0.016**
	(0.005)	(0.009)	(0.007)	(0.011)	(0.020)	(0.007)
Constant	0.087***	0.149***	0.067***	-0.035	-0.058	
	(0.016)	(0.029)	(0.023)	(0.033)	(0.064)	
Observations	2715	2715	2715	2715	2715	2,715
R-squared	0.06	0.12	0.04			

 * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors in parentheses are corrected for clustering at the branch level. All models include controls for lender-defined risk category and month of offer letter. Including controls for loan size and term does not affect the results.