Identifying Information Asymmetries: New Methods and Evidence from a Randomized Field Experiment

January 2004

Dean Karlan*
Princeton University

Jonathan Zinman*
Federal Reserve Bank of New York

This paper estimates the prevalence of asymmetric information in a consumer credit market using a field experiment of our design. A major South African lender issued 60,000 direct mail offers where the interest rate was randomized along two different dimensions — an initial "offer rate" on the direct mail solicitation, and a weakly lesser "contract rate" the applicant received after responding to the solicitation and agreeing to the initial offer rate. These two dimensions of random variation in interest rates, combined with the large sample (including 6,200 accepted offers) and complete knowledge of the Lender's information set, will enable us to identify the prevalence and impacts of specific types of private information. Specifically, our setup distinguishes adverse selection from moral hazard/repayment burden effects on repayment and profitability, and thereby generates unique empirical evidence on the sources and importance (or lack thereof) of asymmetric information.

* dkarlan@princeton.edu, jonathan.zinman@ny.frb.org. Any views expressed herein are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of New York nor the Federal Reserve System.

Stiglitz and Weiss (1981) launched a cottage industry of theoretical papers on the role of asymmetric information in credit markets. Nevertheless, empirical tests of these theories remain relatively rare and inconclusive. Chiappori and Salanie (forthcoming) find this to be true for contract theory in general, and for tests of adverse selection and moral hazard in particular. Two types of problems have hindered the development of such empirical work. First, measuring observables has proven difficult. The econometrician typically falls short of obtaining precise data on what the lender (or insurer, or buyer) can and does observe, and she subsequently struggles to make the accurate distinction between public and private information that is necessary to test theories of asymmetric information. Second, identifying distinct effects of adverse selection and moral hazard is difficult even when precise data on underwriting criteria and clean variation in terms are available, as a single interest rate (or insurance policy variable) might produce both selection and incentive effects that are of independent interest. Hence, even random variation in interest rates (as in Ausubel, 1999) is not sufficient to decompose any reduced-form effect of the interest rate into its structural components.

We address these two problems by working with a South African lender of high-interest, unsecured term credit to generate a unique dataset designed specifically to identify the prevalence and impacts of private information. The Lender's clients comprise employed, but primarily poor, South African consumers who do not have access to the formal banking sector. We are privy to the Lender's credit scoring model and full database of actual and potential customers. Consequently, we are able to construct accurate measures of both the Lender's information set and the outcomes of interest (e.g., loan take-up, loan repayments, and profitability). Moreover, the Lender implemented a "market field experiment" of our design whereby the interest rates offered in solicitations and contracted upon in actual transactions varied randomly and differentially. This experiment was pilot tested in July 2003 and fully implemented between October and December 2003. The unique research design allows us to disentangle selection from incentive effects on repayment.

Specifically, we randomized the interest rate on lending contracts at two points. First, the Lender solicited 60,000 former clients through direct mail. The interest rate on these offers varied randomly and widely, from 3.25 percent per *month* to 11.75 percent per month. The Lender typically lends at interest rates between 7.75 percent and 11.75 percent per month depending on the term of the loan (four, six, or twelve months) and the prior repayment history of the client.

The second randomization, which is critical to our identification of adverse selection, occurred after a client accepted the terms of the loan solicitation. Loan applications were taken and assessed via the Lender's branch network, and per its normal underwriting procedures. In particular, the decision to grant credit, and the maximum loan amount and term offered to creditworthy clients, were determined based on the interest rate that would have been assigned in the absence of the field experiment. Then, some (but not all) clients were randomly assigned a contract interest rate that was lower

than the one presented on their mailer, while the loan size and term were held fixed. This assignment was done "double blind," so that neither the client nor the loan officer knew whether the client would receive a lower rate until after the initial loan terms had been agreed upon. This second randomization allows us to cleanly distinguish adverse selection from repayment burden, 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.*

Formally, our research design is constructed to identify the selection and incentive effects of interest rates by randomly assigning solicited and contracted rates to potential borrowers ("borrowers"), holding everything else constant. Our basic model takes the form:

$$(1)\quad Y_i = f(X_i, B_b, r_i^o, r_i^a, L_i, T_i, G_i, P_i),$$

where i indexes borrowers. Y is an outcome of interest, namely a measure of repayment or demand, as detailed below. X is a vector of observable borrower characteristics. B is branch fixed effect. r^o is the rate offered on the "pre-qualified" mail solicitation; $r^a =< r^o$ is the rate actually contracted upon loan approval; and L and T are the maximum loan size and term (4, 6, or 12 months) offered and contracted upon, respectively.[1] L and T are bounded by the Lender's standard underwriting criteria and interest rates, and these bounds are set "double blind" with respect to r^a; i.e., r^a is not revealed to either the loan underwriter or the potential client until the underwriter has decided whether grant credit, and if yes on what terms. G is a binary variable equal to 1 if r_i^a is less than the Lender's standard rate for someone with i's risk profile. P is a binary variable equal to 1 if r_i^a will apply to all of i's loans with the Lender within the next year, conditional on repayment (and equal to zero if r_i^a applies only to the initial loan).

In testing for effects of asymmetric information, the primary outcomes of interest are related to repayment and profit, as measured by delinquency (late payment), chargeoffs, refinancings, loan revenue, and operating profit (loan revenue minus administration and collection costs). We will exploit the random variation in r^o and r^a to identify any effects of adverse selection and repayment burden on these outcomes. Specifically, r^o identifies the selection effect, and r^a the reduced-form effect of repayment burden. Below we briefly sketch the intuition for how this works, following the relevant references. The paper will present a more formal derivation of the testable predictions.

r^o actually identifies the reduced-form combination of two underlying selection parameters of interest. One is the Stiglitz-Weiss (1981) adverse selection effect, whereby repayment probabilities decrease with interest rates under limited liability as higher rates induce unobservably less risky borrowers to drop out of the applicant pool. In the absence of compelling empirical evidence on the existence of this effect in comparable markets, our priors on its importance in the setting under consideration here should be agnostic. This uncertainty is reinforced by the institutional features of the Lender's market. On one hand, several aspects of contracting environment appear consistent with

[1] An upcoming experiment will randomly assign L, the maximum loan size offered.

the presence of Stiglitz-Weiss adverse selection (rationing; and a very limited menu of contracts, including an inability to take collateral); on the other hand, the applicant pool may have already revealed their types in previous transactions with the Lender.[2] The second potential selection component is a "lemons" effect a la Ausubel (1991), where given 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. Our prior is that the lemons effect is unimportant for the Lender, since empirical evidence suggests strongly that lenders in this market simply do not make price concessions, even for good customers. Nevertheless we will take the possibility of a lemons effect seriously by examining whether the reduced-form selection effect varies by outside debtholding status (observed via credit bureau data) at the time of application.

r^a identifies the reduced-form impact of repayment burden via a combination of several underlying structural parameters of interest. The *incentive effects* of repayment burden work through the (potential) 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 reduce effort (conditional on project choice), by producing "debt overhang" that reduces borrower returns in successful states (Ghosh et al., 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.

Of course, if we find a significant reduced-form effect of repayment burden, then distinguishing among the structural channels will prove interesting. We will undertake such analysis using at least three complementary methods. First, we have randomly assigned repeat contracting opportunities at preferential, experimental rates, conditional on previous repayment performance. In other words, a random subset of project borrowers (those with G*P = 1) will be granted the opportunity to borrow multiple times at lower-than-standard rates, provided that they remain in good standing on their previous loan.[3] This provides an additional, marginal incentive to repay. Second, we have collected data on utilization of loan proceeds and will use this to estimate whether project choice varies with the contracted interest rate. Third, finding statistically and economically significant reduced-form effects of repayment burden would motivate additional data collection on income risk.

---

[2] This motivates our ongoing effort to extend the experiment to potential borrowers who have *not* previously transacted with the Lender.

[3] This is in keeping with Lender's standard underwriting criteria, which automatically rejects any potential borrower that fails to pay back a previously granted loan.

Our approach to estimating the extent and nature of asymmetric information is thus most similar in intent to Edelberg (2003), and in methodology to Ausubel (1999). Edelberg estimates a structural model that attempts to disentangle the effects of adverse selection and moral hazard in collateralized U.S. consumer credit markets, and 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 repayment burden separately, arguing that any such effect must be trivially small over the range of interest rates (800 basis points) contracted on in his data. In contrast, our experiment generates variation in both the quantity (12,000 basis points) and quality (different rates offered versus contracted) required to separately identify any effects of adverse selection and repayment burden.

The takeup phase of this experiment concluded in December 2003, yielding a solicitation hit rate of over 10%. This means that the experiment has produced a sample of over 6,000 loans that can be used to estimate the prevalence of asymmetric information as described above. Since 81% of these loans are four months in duration, we will have most of the repayment data needed to conduct our analysis by May 2004.