Artículos

Gender Issues in Microfinance and Repayment Performance: The Case of a Nicaraguan Microfinance Institution[1]

Silvia L. Saravia-Matus[*] and Jimmy A. Saravia-Matus[**]

Conventional wisdom and increasing empirical evidence in microfinance hold that women are better risks than men. In the present work, a logit model controlling for a range of borrower and loan characteristics was carried out to assess the validity of this statement by comparing repayment rates. The study includes a sample of loans disbursed by a Nicaraguan microfinance institution during the years 2003-2004, a period characterized by high oil prices. A dichotomous dependent variable is created, taking the value of 1 if the credit turned out to be of the best quality i.e. an "A" credit by Nicaraguan regulations, and 0 otherwise. The dependent variable is regressed on variables summarizing the characteristics of the borrower and the loan to investigate the impact of gender on repayment performance. The results provide significant evidence that female client's repayment performance is in fact better than that of male's at the conventional levels of statistical significance. However, the results also show that the perceived difference in gender risk is lower than what popular wisdom would suggest when borrower characteristics and other exogenous economic variables are taken into account. With the present sample, we conclude that other characteristics of the borrower as well as changes in the economic environment can have a similar or larger impact on risk than gender when it comes to repayment performance.

Keywords: Microfinance / Credit Repayment Performance / Gender / Nicaragua

1. Introduction

Microfinance in the context of Less Developed Countries (LDCs) is about making small, short term loans to the self-employed poor who generally do not have credit-bureau records or formal wage jobs (Schreiner, 2004a, p. 1). Usually, the clients are small merchants or manufacturers who do not keep formal financial statements and do not have access to conventional banking.

The process by which a micro-credit is approved or rejected in poor countries differs from homologous procedures in rich countries in several aspects (Schreiner, 2004a, p. 1). The traditional process of pre-disbursement of a micro-credit in LDCs starts when a borrower submits a loan application at one of the branches of the micro-lender. A loan officer then checks the application against an initial screen. If the loan-officer considers that the applicant could be a good candidate, he or she makes a visit to the potential borrower at his work place to take note of the characteristics of the applicant's business on the field. This generally includes checking the applicant's cash, inventory, and other information so as to make a rudimentary approximation of the financial statements of the business. The loan officer also looks for *qualitative data* and takes note of the characteristics of the borrower. These characteristics range from the occupation of the borrower, gender, marital status, etc. Moreover, he or she asks the potential borrower's neighbours to give reference about the applicant's creditworthiness. The loan officer then *subjectively evaluates* whether the applicant is a good candidate to be submitted to a credit committee evaluation. Finally, in the credit committee the loan officer determines in conjunction with the branch manager whether to approve or reject the application.

In this mainly judgemental process loan officers and branch managers often have to rely on conventional wisdom for making good decisions, and in this regard one of the most widely held views among lenders is that women are better risks than men when it comes to disbursing micro-loans. Can this often strongly held view be justified through econometric analysis? If so, how much 'safer' are female clients? Does the perceived difference in risk persist when controlling for other factors such as the client's occupation, age, marital status among others?

The motivation of this paper is to investigate these questions which address key issues in the allocation of loans among the self-employed poor in less developed countries; using a Nicaraguan microfinance institution dataset. The results of our case study indicate that female client's repayment performance is in fact better than male's at the 1% level of significance. However, we also find that the gender gap in risk is not particularly a large one and that other significant factors involved can have a similar or larger impact when it comes to repayment performance. The aim of this paper is thus to acquaint loan decision makers of all potentially relevant aspects which should be considered in order to increase repayment rates and improved overall efficiency in the microfinance system. The econometric tools employed are not without limitations and these are also explored.

This article is organised as follows: in section 2 we perform a brief review of the relevant academic literature. In Section 3 we describe our methodology, data and variable selection process. The main results of the study, their interpretation and constraints are discussed in section 4. Concluding remarks are presented in section 5.

2. Background and Previous Work

Statistical credit scoring has been commonly used for assessing microfinance performance and it uses information about the performance of past loans, obtained from a database, in order to predict the performance of new loans. Hence, a key assumption of the model is that the future will be similar to the past. The model is built using quantitative measures and characteristics of past disbursements so that when one introduces the characteristic of a new application in the credit scoring formula one obtains a score that gives the analyst an objective measure of the probability of default of the case at hand. It is important to note that the formula weights different characteristics of the borrower, the lender and the loan and then produces an objective estimate of the probability of risk that a borrower will not repay (Schreiner, 2000, p. 3). This contrasts with the traditional judgemental scoring method that relies exclusively on qualitative information held in the heads of the loan officers and branch managers.

Since micro-lenders only keep records of approved loans on their databases the estimated parameters of the scoring model are subject to sample selection bias when the models are applied to all candidates (Schreiner, 2004b, p. 15). This happens because the models are fitted to distinguish good loans from bad ones among approved applicants only and it does not include information about rejected applicants (Roszbach, 2004, p. 946). The figure below (based on Schreiner, 2005, p. 35) is a good illustration of the way in which it is proposed that statistical credit scoring should be used in microfinance.
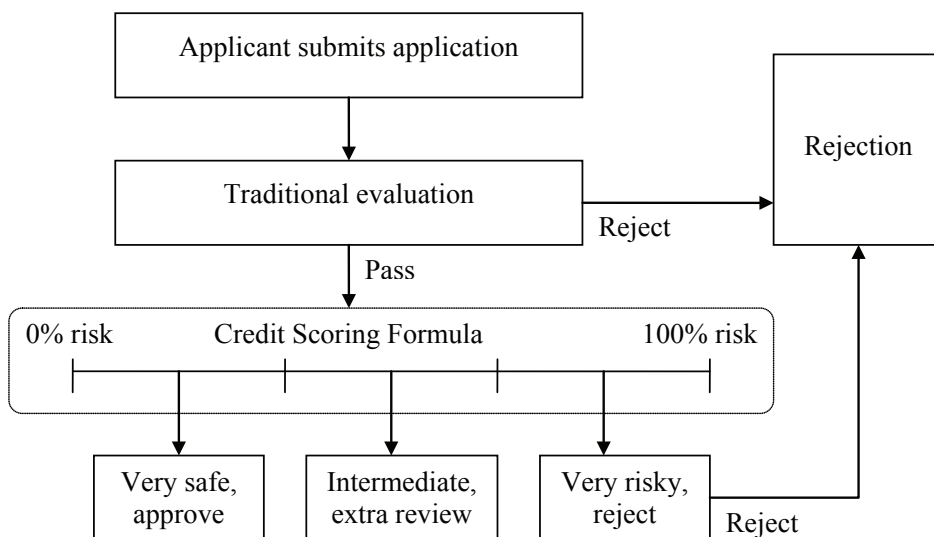
Figure 1. The proposed use of statistical credit scoring in microfinance

Once an application passes the traditional screen, the statistical scoring formula is applied: if the formula shows that the applicant's probability of default is below some predetermined lower threshold, the application is considered to be "very safe" and it is approved without further enquiry. If the applicant's probability

of default is "intermediate", then the credit committee should do an extra review of the application in order to make sure the applicant is a good risk or take actions in order to reduce the exposure to risk. Finally, if the formula shows that the applicant is "very risky", then the application is rejected even though it had passed the traditional screen. As can be seen in Figure 1, statistical credit scoring in microfinance does not offer to save time and money by eliminating the loan officer's "traditional" individualized evaluation. Instead, the proposed use of credit scoring in microfinance is as a final check to detect some high risk cases that may have sneaked through the standard judgemental screens (Schreiner, 2005, p. 6). Schreiner (2005, pp. 10-14) proposes that credit scoring can bring the following benefits to microfinance institutions[2]:

- Although the method does not save time in the initial evaluation of loans, it does save time for loan officers in collections after disbursement. This happens because loan officers spend less time visiting clients that are in arrears thanks to improvements in portfolio quality. As a result there is an efficiency gain which leaves loan officers more time to look for new clients.

- Another benefit of statistical credit scoring is that it can be used as a teaching tool. Once the formula is built one can obtain vital information about which are the key risk factors when analysing applications. This information can then be used to train novice loan officers.

- Finally, statistical credit scoring can induce managers to adopt a culture of explicit data decision making. Managers can use their databases to build statistical credit scoring models with data up to a certain point in time and then apply it to cases that come later (in their databases) with the objective of checking the predictive power of the scorecards. Managers can then adjust rejection thresholds on their scoring models and then see the results of potential scoring policies before they are implemented.

Thanks to these and other benefits, statistical credit scoring promises to reduce the costs of making loans to the self-employed poor in LDCs. This would be desirable from a social point of view because assuming that the microfinance institutions are socially oriented and/or that there is a competitive credit market, these efficiencies can be passed on to borrowers in the form of a lower cost of capital. However, how well can this technique work in the context of LDCs? Table 1 summarises the relevant previous work on statistical credit scoring in microfinance that has been published. The first two columns show the name of the author, details about the publication of the article and the country in which the study was performed. The third and fourth columns describe the data from which the scorecard was estimated. Finally, the last two columns summarise the research approach, that is, how the data was gathered and analysed, and the results and conclusions reached on each of the past works. As can be seen from the table, the microfinance credit scoring empirical studies in LDCs are relatively recent. We can compare the usefulness of these models as statistical scorecards on three main aspects: data collection technique, sample size, and tests of predictive power.

---

2 *Please note that this list is not exhaustive and it is included for illustrative purposes.*

The two main approaches to data collection are surveys and micro-lenders' databases. The works of Reinke (1998) and Schreiner (2004a) rely on information that the microfinance institutions already collect, while Viganò (1998), Sharma and Zeller (1997) and Zeller (1998) use information from surveys. Caire and Kossmann (2002, p. 8) report that to build a useful statistical model when using a micro-lendr's database one requires payment performance data for approximately 15,000 loans. From table 1, we can see that Schreiner (2004a) has a sample size large enough (39,956 loans) to comply with this requirement while the other scorecard built from a database (Reinke, 1998) uses 1,641 observations only. Even though in models derived from surveys the sample size is smaller, it is worth noting that the information gathered is more detailed and generally includes more variables than the data gathered from micro-lenders' databases. However, this comes at a higher data collection cost for both the borrower and the lender. Regarding this last point, Hand and Henley (1997, p. 528) point out that there is a trade-off on the number of variables to be included in the scorecard: if the data set is large enough over-fitting problems may not occur and one might try to use as many variables as possible, however too many questions will likely deter applicants, who might go elsewhere.

It is important to note that only Schreiner (2004a) checks the predictive power of his model by using 'out of sample data' in order to check how his scorecard would have performed in practice. While Viganò checks the predictive power of her model by using the same data from which it was estimated the other three studies do not perform tests of predictive power. Apparently, the reason for this is that these latter models aim to detect characteristics of the borrowers that are linked with risk and not to develop scorecards destined to be used in practice (Schreiner, 2004a, p. 5).

The methods used to analyse the data in previous studies include factor analysis, tobit, probit and logit models. According to Hand and Henley (1997, p. 535) there is no overall "best" method, and what is best depends on the details of the problem. Finally, regarding the conclusions of the different papers in the literature, it is worth noting that all authors suggest, to a certain extent, that credit scoring techniques can actually be used in microfinance.

Several previous studies in the microfinance literature have investigated the relationship between the characteristics of borrowers and their loans and repayment performance. Among the most known papers are: Schreiner (2004a), Reinke (1998), Viganò (1998), Zeller (1998) and Sharma and Zeller (1997). Zeller (1998) and Sharma and Zeller (1997) investigate the performance of credit groups and therefore these two articles are not directly relevant to the present study. On the other hand, while Viganò (1998) investigates the performance of individual clients, the author does not report in her article the estimated impact of gender or other factors in repayment performance.

This leaves us with two papers which are directly relevant to the present work: Schreiner (2004a) and Reinke (1998). These two papers do present findings regarding the impact of the borrower's gender and other characteristics on repayment performance; interestingly the results in these two articles differ. While Reinke (1998, pp. 53-54) finds that women are significantly better risks than men, Schreiner (2004a, p. 11), on the other hand, finds that, at least for his sample, there is no statistically significant difference in risk between men and women. Moreover,

Schreiner indicates that his results show that if other characteristics of the loan are taken into account, many of which are related to gender (such as occupation) "much of the gender gap in risk vanishes". Therefore, it is possible that the difference between these two results may be due to the correct specification of the model. In this article, we aim to contribute to the microfinance literature by presenting a "correctly" specified model in the sense that it passes a basic misspecification test, namely the Ramsey RESET test.

Table 1. Summary of previous published studies of 'statistical credit scoring' in microfinance

Authors/ publication	Country in study	Sample size	Sample details	Research approach (instruments)	Results and conclusions
Viganò (1998)	Burkina Faso	118 individual customers	Sample was obtained from the population of all individual customers of a bank in Burkina Faso in 1993. The data were not complete for all observations requiring that the original sample be reduced to 100.	Quantitative and qualitative information was collected through questionnaires and direct interviews with loan managers. The questionnaire targeted population topologies and loan characteristics. Selection of variables was determined by the number of observations for which they were available. Factor analysis was applied to the final set of variables. The characteristics were condensed into 11 factors.	In spite of limited available information the results showed that credit scoring techniques could be used in microfinance. The original sample classification is satisfactory for defaults as it correctly classifies 92% of observations.
Sharma and Zeller (1997)	Bangladesh	From 1,725 loans recorded, a subset of 868 loans was used	Loan transactions collected from 128 borrower groups belonging to three credit programs in Bangladesh. Data was collected during 1994.	Survey conducted by the International Food Policy Research Institute (IFPRI) of 128 borrower groups. Groups were randomly selected from 41 villages across Bangladesh. A formal questionnaire was administered to the chairperson of each group. Using tobit analysis, specific tests are performed on hypothesised determinants of group risk.	The Study identifies factors that affect group repayment rates. The analysis indicates that factors such as asset and enterprise diversity within groups significantly affect repayment rates. The conclusion is that if basic principles of prudential banking are adhered to, repayment rates can be good even in poor and remote communities.

Reinke (1998)	South Africa	1,641 borrowers	The data were copied from the lending institution's associated Informal Business Training Trust records in April 1995. The data comprised all existing borrowers.	Database of a South African micro-lender. The variables in the data-set were chosen to capture the social and business characteristics of the clientele and also because of data availability criteria. Probit estimation was carried out to determine the relative importance of the variables for repayment probabilities.	The article argues that efficient institutional design can make microfinance lending financially sustainable. Women are shown to be more conscientious with their repayments. The results suggest that the average repayment may be improved if women were to participate in greater numbers.
Zeller (1998)	Madagascar	168 group loans	The loans in the dataset were obtained by a random sample of 146 groups from six different lending programmes in 1992. Because of missing values for some of the variables the number of cases used in the regression analysis is 141.	The data was gathered by performing a survey. Questionnaires targeted households, credit groups and communities. The model was estimated by using the tobit maximum likelihood technique. The regression model incorporates variables that reflect the process of group formation, characteristics of the group as well as their programme and community. The variables in the model were hypothesised to influence repayment rate directly or indirectly.	This is an empirical analysis to test a theoretical model to explain the performance of credit groups. Groups that are located in communities with a high exposure to covariance risks were fount to have a significantly lower repayment rate than those with lower exposure. Groups with higher levels of social cohesion were found to have a better repayment rate.
Schreiner (2004)	Bolivia	39,956 loans in 1988-96 10,555 loans in 1997	Loans repaid in 1988-1996 are used to create scorecard. Loans repaid in 1997 are used to test scorecard.	Database of a Bolivian micro-lender. The credit scoring formula or scorecard was derived from a logit model. The scorecard relies on characteristics that most microlenders already collect or are easy to collect.	Scorecard is tested at different risk thresholds in order to determine how well it separates "good" loans from "bad" ones. The scorecard separates "good" from "bad" loans. On average it assigns higher risk to "bad" than to "good" loans. The scoring model for arrears developed in this paper suggests that statistical credit scoring can help microfinance.

3. Methodology, Variable Selection and Data Collection

While performing our literature survey, it was evident that the logit model had significant advantages in ease of calculation and hypothesis testing procedures (Gujarati, 1995, pp. 567-568) and also that it was easier to perform the RESET test (manually) for the logit in comparison with e.g. a tobit model. For this reason, we decided to do a logit model for our scorecard.

The process of variable selection was based on the theoretical and empirical models of Reinke (1998) and Schreiner (2004b). We have thus relied for the most part on the variables that had shown to be linked with risk of default and also on variables that the microfinance literature suggests that are relevant but that have not been tested (because the data were not available)[3]. Consequently, the petition of data from the Nicaraguan microfinance institution included the socio-economic characteristics of the individual borrower, the lender and the loan. It should be mentioned that it was not possible to secure full access to the database of the lender in order to investigate if there were more variables available that could be used. The management of the Nicaraguan bank kindly agreed to provide information about the variables requested, but they would not consent to give a complete list of all the items contained in their database. Table 2 gives a list of the variables used in the model.

Table 2. List of variables for the study

Number	Dependent Variable	Type of variable
	Whether or not the loan had a spell of arrears of more than 15 days	binary
	Independent Variables	
	Characteristics of the borrower:	
1	Gender	binary
2	Sector of the borrower's business (e.g. merchant, manufacturer)	discrete
3	Age	continuous
4	Marital status	binary
5	Whether the client is new or old	binary
6	Whether the client had a spell of arrears of more than 15 days in the previous loan	binary
	Characteristics of the lender:	
7	Branch where the loan was disbursed	discrete
	Characteristics of the loan:	
8	Repayment period	continuous
9	Type of Guarantee	discrete
10	Interest rate (in percentages)	discrete
11	Loan size (in U.S. dollars)	continuous
12	Foreign currency	binary

---

3   *For a comprehensive list of the relevant variables see Schreiner (2004b, pp. 23-24).*

The sample used consists of 48,602[4] loans disbursed by a Nicaraguan microfinance institution between the 2nd of January 2003 and the 30th of June 2004. Of this sample, 3,114 loans went "bad" (6.41%), where bad is defined as a loan with at least one spell of arrears of more than 15 days. We have chosen to use this definition of "bad" with two motivations in mind. First, according to Nicaraguan bank regulatory bodies, microfinance "A" loans are the ones with no arrears or with a spell of arrears of up to 15 days at a given point in time (Rosales, 2002, p. 3); second, to make the results of this article comparable to those in previous work which use this definition such as Schreiner (2004a). Note that loans classified "B" and below are costly because they require extra collection efforts. This implies that loan officers lose time visiting clients in arrears instead of looking for new clients. In addition, regulatory bodies require banks to weight their assets on a risk basis. Consequently, in order to comply with solvency ratios, banks and microfinance institutions with bad quality portfolios must keep larger amounts of capital. As in previous studies in the microfinance literature, it is important to note that since the lender only keeps data on accepted applicants and information from credit reference agencies (credit-bureaux) is not available, in the present article we model risk conditional on the applicant having passed the microfinance institution's standard evaluation. Therefore, in the present work, when we talk about 'risk' or the 'impact on risk' we refer to 'post evaluation risk'.

Consequently, in the model specification we have focused on capturing the probability that a loan will go "bad" which can be expressed as a function of the characteristics of (a) the individual borrower (b) his/her loan (c) the lender and (d) the economic environment.

Hence, the logit model is the following:

$$\text{Prob}(arrears>15_i = 1 \mid \mathbf{x}_i) = F(\mathbf{x}_i ß)$$

The dependent variable "arrears>15" is a binary variable taking the value of 1 if the loan had a spell of arrears of 15 days or more during its life and 0 otherwise, F is the logistic cumulative distribution function, $\mathbf{x}_i$ is a vector of values for the ith observation of the independent variables and ß is a vector of parameters.

Next, we summarise our a priori expectations concerning the impact that the characteristics of the borrowers, the lender and the loan should have on the dependent variable. Note that the dependent variable is whether or not the loan had a spell of arrears of 15 days or more.

- **Business sector:** as will be seen below, we set up a complete range of dummy variables for the sector in which the borrower works. The base category is borrowers that have a high school diploma but no other qualification and that engage in a variety of service sector activities. At this point it is not clear on a priori grounds if the dummy variables will be significantly riskier than the base category or vice versa. At the moment

---

4   The initial sample we obtained from the microfinance institution contained 48,625 loans; however, we dropped 23 observations in which the age of the clients was below 16 years. Our results are robust to including these observations in the database.

we only have references from the literature indicating that the merchant sector should be relatively safe, and also that carpenters have been found to be amongst the most risky clients (Schreiner, 2004a, p. 12).

- **Gender of borrower:** different opinions in the literature can be found as to the risk implied by this variable. Reinke (1998, p. 54) reports that women are "definitively more reliable than males". However, Schreiner (2004a, p. 11) argues that if other characteristics are taken into account, such as the sector of the firm, gender is much less predictive.

- **Age:** we expect that older people are more stable, careful and responsible. Moreover, since the loans in our study have very short maturity, we believe that the chance of ill health may not be so important. Therefore we expect on a priori grounds that risk decreases with age.

- **Marital status:** we expect that married people are more constant and dependable than divorced people and people having free unions. Therefore we expect on a priori grounds that risk is lower for loans issued to married individuals.

- **Repayment history of the borrower:** there are two variables in the model that account for the repayment history of the borrower. The first one indicates if the borrower is a new client or not. It is sensible to expect that if the client is new the risk of the loan will be greater than otherwise. This is because in the case of microfinance in Nicaragua there are no credit bureaux that can supply information about the credit record of a given borrower. The second variable indicates whether an old client had a spell of arrears of 15 or more in the previous loan. It can be expected that if the client had costly arrears in the past it is more likely that he or she will also have costly arrears in the next loan.

- **Branch of the microfinance institution:** dummy variables for each of the branches of the Nicaraguan financial institution are created; the reference category is the central office. On a priori grounds it can be expected that the branches should be riskier than the central office since the central office concentrates on a few larger loans that receive special review before being disbursed.

- **Length of repayment period:** clearly, one can expect that the longer the time to maturity, the greater the chance that a given borrower will fall into a spell of arrears.

- **Type of guarantee:** there are three types of guarantee for the Nicaraguan microfinance institution: mortgage (M), joint liability (JL) and guarantee on the personal belongings (PB) of the borrower. On a priori grounds, it can be expected that a loan in which the guarantee consists exclusively of PB, has been considered to be less risky by the credit committee than a loan where the guarantee is "JL plus PB". Also, one can expect that a loan requiring only JL as guarantee has been categorized by the traditional evaluation as being less risky than another requiring PB as guarantee. Note that plain JL is the base or reference category. We form no a priori expectations about the variable "mortgage" because this guarantee is usually required for larger loans, different from the above.

- **Interest rate:** regarding this variable, it is sensible to expect on a priori grounds that the higher the interest rate the higher should be the risk involved in the loan. The reason is that, because of the institutions' credit policy, the credit committee will charge a higher interest rate to cases that are regarded as being more risky.
- **Loan size:** Schreiner (2004a, p. 12) reports that the link between risk and the amount disbursed is positive but weak. According to Schreiner, if this variable is found to be insignificant at the conventional levels, then this could indicate that the lender is successfully rationing borrowers suspected as high risks.
- **Foreign currency:** during the period of our study, the Nicaraguan currency, the "córdoba", was pegged to the US dollar. Moreover, all lending contracts required loans to be repaid in their equivalent to US dollars. Hence we do not expect any extra exposure to currency risk for borrowers with loans in US dollars above the risk that other borrowers already incur. However, this variable is introduced in the model because we expect on a priori grounds that this group of borrowers may have a higher level of risk than other borrowers because loans in dollars are usually requested by entrepreneurs who travel abroad to do their business (e.g. to Panama), or that have foreign suppliers.

In addition to the above variables that are commonly used in the literature for the assessment of repayment performance, we also included a dummy variable named 'oilprice' which takes the value of 1 if the loan expired after the 30th of June 2004. During the second half of 2004 there were important increases in the international price of oil. For a country like Nicaragua, which produces no oil, it is plausible that this change had adverse consequences for the economy of the country, which in turn may have affected the repayment performance of the clients of the micro-finance institution.

An inspection of Nicaraguan newspaper editions for the time period under study provides anecdotic evidence to support the hypothesis that the price of oil is having a negative impact on the economic environment in Nicaragua. For instance, a journal article published on November 15th 2004 in *El Nuevo Diario*, the journal with the second highest level of circulation in the country, reports that the authorities of INE [National Institute of Energy] point out that "the price of oil, which is between 46 and 47 dollars per barrel at the moment is substantially above the historic average of 27 to 30 dollars per barrel" and that "this price level is exerting great pressure on the finances of the country since oil expenses normally consume 30% of the national export income; it is believed that this year these expenditures will rise way above this ceiling" (Sánchez Rizo, 2004).

Figure 2 shows the development of the spot price of Brent Oil in U.S. dollars per barrel for the period in study. In the period from May 2004 to May 2005 the price of oil was higher than the previous period. Moreover, the volatility of oil prices was much greater, which may have contributed to business uncertainty.

Figure 3 was designed to investigate the impact that the adverse business conditions may have had on borrower repayment performance and to help determine

when the unfavourable conditions probably started affecting loan repayment performance. To generate the figure, we first grouped all loans according to the month in which they were disbursed. Then we calculated monthly percentages of "bad" loans. Finally, we broke down the "bad" loan percentages according to the time to maturity in the contract of the loans.
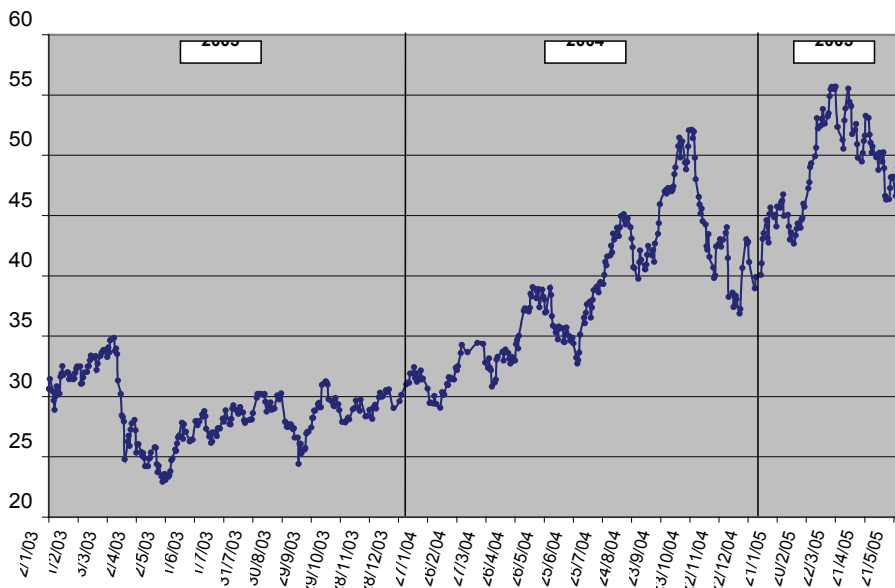

Figure 2. Spot Brent Oil prices in U.S. dollars for the period in study[5]
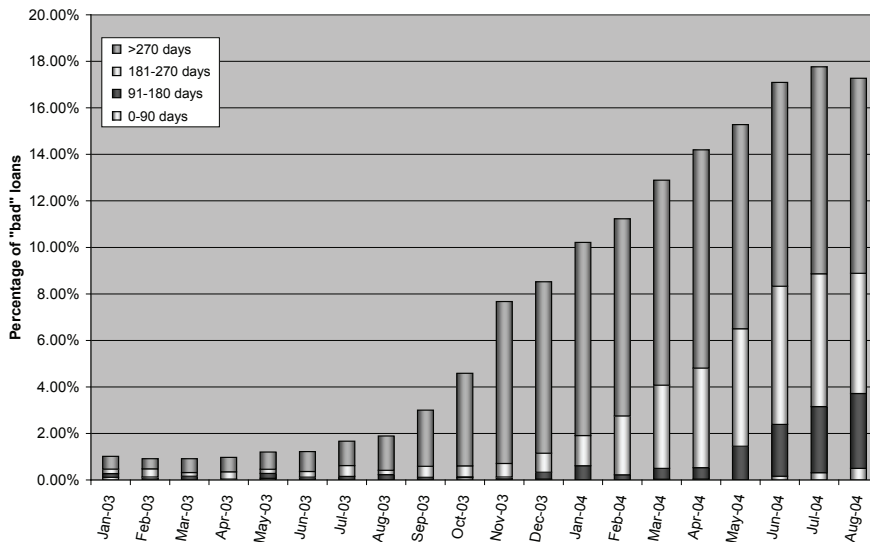

Figure 3. Disbursement date and behaviour of arrears in different maturity groups

From Figure 3 one can notice a marked increase in the percentage of "bad" outcomes in the loans with a life period extending beyond the second half of 2004. It is interesting to observe that there is a strong increase in the percentage of "bad" loans among the contracts with maturities greater than 270 days that were issued between July-03 and December-03, from 1.05% in July-03 to 7.37% in Dec-03. Also, for the loans issued between Jan-04 and June-04 there is a sharp increase in the percentage of bad loans in the contracts with maturities between 181-270 days, from 1.30% in January to 5.94% in June. There is also a steady increase in the percentage of bad loans for loans with maturities between 91-180 days issued in the period between April-04 and June-04, from 0.49% per month to 2.23% per month. Finally, from June-04 to August-04 there is an increase in the percentage of bad loans for loans with maturities of less than 90 days, from 0.16% in June-04 to 0.50% in August-04.

These observations are consistent with the hypothesis that the worsening of the economic conditions started affecting loan repayment performance at some point during the months of June and July 2004 almost coinciding with the start of a more rapid rise in the price of oil. Naturally, the selection of this date is of necessity an approximation.

Table 3 presents descriptive statistics for the observations in the sample data, according to the selected dependent and independent variables.

Table 3. Descriptive statistics for all observations in the sample data

Variable	Definition	Mean	Std. Dev.	Min	Max
arrears>15	1 if client had arrears > 15 days in the present loan, 0 otherwise	0.0641	0.2449	0	1
artisan	1 if artisan, 0 otherwise	0.0104	0.1013	0	1
carpenter	1 if carpenter, 0 otherwise	0.0021	0.0462	0	1
merchant	1 if trader, 0 otherwise	0.8811	0.3237	0	1
seamstress	1 if seamstress, 0 otherwise	0.0102	0.1006	0	1
baker	1 if baker, 0 otherwise	0.0028	0.0530	0	1
othermanuf	1 if other manufacturer, 0 otherwise	0.0335	0.1800	0	1
grocery	1 if grocery, 0 otherwise	0.0033	0.0575	0	1
transport	1 if transport, 0 otherwise	0.0113	0.1057	0	1
professionals	1 if professional, 0 otherwise	0.0116	0.1073	0	1
skillwrk	1 if skilled worker, 0 otherwise	0.0038	0.0617	0	1
unskillwork	1 if unskilled worker, 0 otherwise	0.0289	0.1676	0	1
male	1 if male, 0 otherwise	0.2909	0.4542	0	1
married	1 if married, 0 otherwise	0.5099	0.4999	0	1
lnsize	Loan size in U.S. dollars, continuous variable	504.27	876.98	19.28	30,000.00
age	Age of borrower at disbursement date in years	41.80	10.77	16.51	94.30
repayperiod	Repayment period in days, continuous variable	208	103	5	394
intrate	Monthly Interest rate (percentage)	0.0490	0.0059	0.0085	0.0550
mortgage	1 if the guarantee includes mortgage, 0 otherwise	0.0080	0.0889	0	1
jointliability	1 if the guarantee includes joint liability, 0 otherwise	0.7120	0.4529	0	1

persbelong	1 if the guarantee includes client's personal belongings, 0 otherwise	0.9832	0.1284	0	1
currency	1 if currency is USA dollars, 0 otherwise	0.0115	0.1067	0	1
new	1 if client is new, 0 otherwise	0.2364	0.4249	0	1
prevarrears	1 if client had arrears > 15 days in previous loan, 0 otherwise	0.0026	0.0509	0	1
branch1	1 if branch is branch 1, 0 otherwise	0.0757	0.2645	0	1
branch2	1 if branch is branch 2, 0 otherwise	0.0581	0.2339	0	1
branch3	1 if branch is branch 3, 0 otherwise	0.1265	0.3324	0	1
branch4	1 if branch is branch 4, 0 otherwise	0.1276	0.3336	0	1
branch5	1 if branch is branch 5, 0 otherwise	0.0548	0.2277	0	1
branch6	1 if branch is branch 6, 0 otherwise	0.1018	0.3025	0	1
branch7	1 if branch is branch 7, 0 otherwise	0.0852	0.2792	0	1
branch8	1 if branch is branch 8, 0 otherwise	0.0625	0.2421	0	1
branch9	1 if branch is branch 9, 0 otherwise	0.0832	0.2762	0	1
branch10	1 if branch is branch 10, 0 Otherwise	0.0727	0.2597	0	1
branch11	1 if branch is branch 11, 0 otherwise	0.0606	0.2387	0	1
branch12	1 if branch is branch 12, 0 otherwise	0.0323	0.1769	0	1
branch13	1 if branch is branch 13, 0 otherwise	0.0437	0.2044	0	1
branch14	1 if branch is branch 14, 0 otherwise	0.0071	0.0837	0	1
branch15	1 if branch is branch 15, 0 otherwise	0.0062	0.0786	0	1
branch16	1 if branch is branch 16, 0 otherwise	0.0011	0.0336	0	1
centraloffice	1 if loan is disbursed in the central office, 0 otherwise	0.0008	0.0276	0	1
oilprice	1 if loan matures after 30/06/04, 0 otherwise	0.4133	0.4924	0	1

4. Results

4.1. Logistic Regression

Table 4 presents the results of the estimated logit model derived from the data described above. As we mentioned before, the model is a logit estimated by using the maximum likelihood method. The dependent variable is "arrears>15"; this is a binary variable taking the value of 1 if the loan had a spell of arrears of 15 days or more during its life and 0 otherwise. The results were obtained using the STATA 8 software. Note that in the header of table 4, "Log pseudo-likelihood" corresponds to the value of the log likelihood at convergence. The number of observations is 48,602. Wald chi2(41) = 2946.9 is the value of a Wald chi-squared for the test of the null hypothesis that all of the coefficients associated with independent variables are simultaneously zero. The p-value is indicated by Prob > chi2 = 0 and hence the null hypothesis that all coefficients are simultaneously zero is rejected at all conventional significance levels. Pseudo R2 is the measure of fit known as McFadden's R^2, also known as the "likelihood-ratio index". It compares a model with just the intercept to a model with all parameters (Long & Freese, 2003).

Table 4. Results

Logit estimates		Number of obs	=	48602	
Log pseudolikelihood = -8940.2846		Wald chi2(41)	=	2946.9	
		Prob > chi2	=	0	
		Pseudo R2	=	0.2272	
arrears>15	Definition	Coef.	Robust Std. Err.	z	
artisan	1 if artisan, 0 otherwise	-0.1493811	0.2012283	-0.74	
carpenter	1 if carpenter, 0 otherwise	0.2387077	0.2934317	0.81	
merchant	1 if trader, 0 otherwise	-0.1966845	0.0961874	-2.04	**
seamstress	1 if seamstress, 0 otherwise	-0.4161685	0.2036856	-2.04	**
baker	1 if baker, 0 otherwise	-0.5465168	0.3676158	-1.49	
othermanuf	1 if other manufacturer, 0 otherwise	0.0125404	0.1351197	0.09	
grocery	1 if grocery, 0 otherwise	-1.601503	0.4353523	-3.68	*
transport	1 if transport, 0 otherwise	-0.5600704	0.1881701	-2.98	*
professionals	1 if professional, 0 otherwise	-0.123267	0.1607943	-0.77	
skillwrk	1 if skilled worker, 0 otherwise	-0.5172441	0.2466334	-2.10	**
male	1 if male, 0 otherwise	0.1622637	0.0447026	3.63	*
married	1 if married, 0 otherwise	-0.1958985	0.0407113	-4.81	*
lnsize	Loan size in U.S. dollars, continuous variable	-0.0001317	0.0000481	-2.74	*
lnsize2	Square of lnsize	7.88E-09	2.56E-09	3.07	*
age	Age of borrower in years, continuous variable	-0.0145667	0.0018986	-7.67	*
repayperiod	Repayment period in days, continuous variable	0.0148225	0.0016662	8.90	*
repayperiod2	Square of repayperiod	-0.0000152	3.16E-06	-4.81	*
intrate	Monthly interest rate (percentage)	13.22084	3.905531	3.39	*
mortgage	1 if the guarantee includes mortgage, 0 otherwise	-0.1003047	0.2674882	-0.37	
jointliability	1 if the guarantee includes joint liability, 0 otherwise	0.1750788	0.05842	3.00	*
persbelong	1 if the guarantee includes client's personal belongings, 0 otherwise	0.1947813	0.1325769	1.47	
currency	1 if currency is USA dollars, 0 otherwise	0.2490385	0.1672549	1.49	
new	1 if customer is new, 0 otherwise	0.4876172	0.0444901	10.96	*
prevarrears	1 if customer had arrears > 15 days in previous loan, 0 otherwise	1.561937	0.2666845	5.86	*
branch1	1 if branch is branch 1, 0 otherwise	-0.4547961	0.5335187	-0.85	
branch2	1 if branch is branch 2, 0 otherwise	0.2559626	0.5318795	0.48	
branch3	1 if branch is branch 3, 0 otherwise	-0.4387723	0.5325582	-0.82	
branch4	1 if branch is branch 4, 0 otherwise	-0.3061833	0.5321612	-0.58	
branch5	1 if branch is branch 5, 0 otherwise	-0.396935	0.5313132	-0.75	
branch6	1 if branch is branch 6, 0 otherwise	-0.5892935	0.5336401	-1.10	
branch7	1 if branch is branch 7, 0 otherwise	-0.441166	0.5329375	-0.83	
branch8	1 if branch is branch 8, 0 otherwise	-0.4676275	0.5327621	-0.88	
branch9	1 if branch is branch 9, 0 otherwise	-0.7110977	0.5362554	-1.33	
branch10	1 if branch is branch 10, 0 Otherwise	-0.3674053	0.5343074	-0.69	
branch11	1 if branch is branch 11, 0 otherwise	-0.2687152	0.5333652	-0.50	
branch12	1 if branch is branch 12, 0 otherwise	0.1748783	0.533798	0.33	
branch13	1 if branch is branch 13, 0 otherwise	-0.165788	0.5340773	-0.31	
branch14	1 if branch is branch 14, 0 otherwise	-0.8984887	0.5788471	-1.55	
branch15	1 if branch is branch 15, 0 otherwise	0.0944552	0.5495912	0.17	
branch16	1 if branch is branch 16, 0 otherwise	-0.8449208	0.644088	-1.31	
oilprice	1 if loan matures after 30/06/04, 0 otherwise	2.23579	0.065081	34.35	*
_cons	Constant	-6.809966	0.5812038	-11.72	*

* significant at the 1% level (two sided test)

** significant at the 5% level (two sided test)

*** significant at the 10% level (two sided test)

The left column lists the variables in the model, with the dependent variable situated at the top. A definition of each of the variables is found in the second column. The column "Coef." contains the ML estimates. Column "Robust Std. Err." is the standard error of the estimates. Since robust to heteroskedasticity 'White standard errors' were found to be very close to the "traditional" standard errors, heteroscedasticity is not likely to be a problem. However, it is probably safer to allow for heteroskedasticity, hence we decided to use the White standard errors. The resulting 'z-test', equal to the estimate divided by its standard error, is labelled z. With the estimated values of the latent variable on which the logit is based, we can use the logistic cumulative distribution function to obtain predicted probabilities (Long & Freese, 2003, pp. 110-113). The latter probabilities are the basis of the discussion of our results in the following sub-section.

A RESET test was performed on this model in order to detect for omitted variable and/or functional form misspecification. The following table shows that the test statistic of the RESET test was insignificant at any conventional level of significance.

Table 5. Test Statistic of the RESET

Dependent Variable = arrears>15		Robust		Level of
Variable for test	Coef.	Std. Err.	z	significance
xbeta2	0.0017809	0.0151395	0.12	Not significant at usual levels

The result of the test indicates that we cannot reject the hypothesis that the model is correctly specified. Thus, the model passes the RESET test. Naturally, this does not imply that the model can be considered "flawless". It means that the model passes a minimum quality check so that it is reasonable to attempt to identify the links between the characteristics of the borrower and repayment risk. Although the presented scalar measure of fit is still low (which might be interpreted as indicating that the model does not have an excellent fit), it can be shown that the model classifies the observations into "goods" and "bads" in such a way which is acceptable and useful for microfinance institutions (see Table 6), especially, if we evaluate the model under the assumption that the future will be similar to the past and in the context of the methodology that has been proposed (i.e. the use of scoring models in microfinance) (further details are available in the Appendix).

Table 6. Power to predict at different risk thresholds for the model on table 5

Measure	Formula	Risk Threshold							
		All-bad							All-good
		0%	5%	10%	15%	20%	25%	30%	100%
Share of "Goods" approved	Goods Approved / (Goods Approved + Goods Rejected)	0.00	0.51	0.63	0.76	0.88	0.94	0.98	1.00
Share of "Bads" approved	Bads Approved / (Bads Rejected + Bads Approved)	0.00	0.07	0.15	0.34	0.58	0.76	0.89	1.00

From Table 6, we can see that if a lower threshold of 10% would have been used together with the estimated model (Table 5); 63% of all "goods" and also 15% of "bad" loans would have been approved without further review. If "very risky" cases were defined as those with predicted risk at or above the 30% threshold, then the percentage of rejected loans which would have turned out to be "goods" would have been substantial. This analysis can be improved and refined by varying the thresholds according to the definition and/or perception of "low", "intermediate/medium" and "high" loan risks of the institution as well as in consonance to current macroeconomic environment, objectives and needs.

4.2. Discussion of Results

The distinguishing feature of the interpretation of linear regression models is that one only needs to specify which variable is changing and by how much while specifying that all other variables are being held constant. In contrast, logit models are non-linear and as a consequence the effect (on the dependent variable) of a given change in an independent variable will depend on the value of that variable at the start of its change and the value of the other independent variables in the model (Long & Freese, 2003, pp. 95-97).

In view of this characteristic of logit models, in this section we decided to follow a two step strategy in order to discuss our results. First, using the results in table 4 we calculated the probability that an artificial "representative borrower" may fall in a spell of arrears of 15 days or more, and second we studied the estimated impact of changing each of the variables one at a time on repayment performance. Moreover, in order to better illustrate our results, we then made extensive use of tables and figures. We also compared our results with those from the established literature when possible.

In order to select the characteristics of our representative borrower we used the descriptive statistics of table 3 and chose the characteristics with higher means for the case of dummy variables (because these are therefore the most common characteristics), and the average of continuous variables. Proceeding in this way, our representative client is a married female merchant aged 42, she is a previously existing client of branch 4 with no arrears in her previous loan; her loan, in national currency, is for the equivalent of US $504.27 to be repaid in 208 days with a monthly interest rate of 4.9%. The guarantee consists of joint liability plus some of the client's personal belongings. Since we have introduced the dummy variable "oilprice" in our model to account for changes in the economic conditions, we can also compare how our representative borrower would perform in a "normal economic environment"[6] versus a scenario with adverse business conditions.

According to the results in table 4, the representative borrower has a predicted risk of 0.86% of incurring arrears greater than 15 days under normal economic conditions, while the predicted risk for the same client under adverse

---

6   In this section we refer to the economic conditions prevailing before the period of oil price instability as "normal economic conditions", perhaps a more precise term would be "conditions prevailing before oil market instability".

economic conditions is 7.54%. Hence, holding all other factors equal, there is an estimated impact of 6.68% in repayment performance due to changes in the economic environment. In other words, we can expect that on average 9 out of 1000 such representative clients would incur spells of arrears of 15 days or more under normal conditions, while under adverse conditions, we can expect that on average 75 out of 1000 representative clients would incur such arrears.

Gender of borrower: now we are ready to study the impact of gender on repayment performance. The variable "male" is positive and significant at the 1% level. This indicates that on average males are more risky borrowers than women, other things equal, as Reinke (1998, p. 54) finds in his work. However, it is interesting to observe that the variable has a somewhat small impact[7] on risk. Our model suggests that if the representative client was male instead the predicted risk would increase by 0.15% under normal conditions, and by 1.21% under adverse economic conditions, all other factors equal. This seems to be in accordance with Schreiner (2004a, p. 11) who argues that the gender of the borrower is less predictive when other characteristics are taken into account (such as the sector of the business).

Business sector: the following table shows the estimated average effect on repayment performance from changing the business sector characteristic of the representative client from "merchant" to one of the other business sectors, under both "normal" and "adverse" economic conditions while holding all other things equal. Note that the results in table 3 indicate that (at least for our sample) the business sectors "artisan", "carpenter', "baker", "othermanuf" and "professionals" are not statistically different from the base category and thus are not reported in table 7 below. We interpret the difference in risk between "merchant" and the business sectors omitted in the table as similar to that between "merchant" and the base category.

Table 7. Impact of business sector on repayment performance

From	To	Normal conditions	Adverse conditions
merchant	grocery	-0.65%	-5.58%
merchant	seamstress	-0.17%	-1.40%
merchant	skilled worker	-0.24%	-1.95%
merchant	transport	-0.26%	-2.18%
merchant	unskilled workers (base category)	0.19%	1.49%

As can be seen from table 7, the impacts of changing the sector of the business can be small under normal conditions, but under adverse economic conditions the impact can be relatively large.

---

7   In this section, we define "impact of a change of characteristic on repayment performance" as the difference between the estimated risk of the representative client once one of the characteristics has been changed and the original estimated risk of the representative client. For example, for gender under adverse conditions, the impact on risk for changing from a female to a male client ceteris paribus equals 1.21%, the new risk if the client is male is 8.76%, and the original risk (female client) is 7.54%.

Age: the model suggests that, on average, risk decreases with the age of the borrower, all other factors equal. This coefficient is significant at the 1% level. Table 8 below summarises the effect on risk from changes in the age of the representative borrower. The percentages show the estimated impact on risk when we change the age of the representative client from 42 years to the ages indicated in the table.

Table 8. Impact of age on repayment performance

| Age from	To	Normal conditions	Adverse conditions
42 years	18 years	0.36%	2.83%
42 years	26 years	0.22%	1.80%
42 years	34 years	0.11%	0.85%
42 years	42 years	0.00%	0.00%
42 years	50 years	-0.09%	-0.77%
42 years	58 years	-0.18%	-1.47%
42 years	66 years	-0.25%	-2.10%

Under normal conditions, if the representative borrower was 18 years old instead of 42, the model suggests that, other things equal, the predicted risk of this client would be 1.22% rather than 0.86% (a 0.36% impact); while under adverse conditions this change would mean that the predicted risk of the client would be 10.37% instead of 7.54% (a 2.83% impact), other factors equal. The other values in the table can be interpreted in a similar fashion: add the values on the "normal conditions" column to 0.86% to obtain total risk for a change in the age of the borrower under normal conditions, all other things equal. Similarly add the values in the "adverse conditions" column to 7.54% to obtain the total risk for a change in the age of the borrower under adverse economic conditions, all other factors equal.

Marital status: the coefficient of the variable "married" is significant at the 1% level and hence the model suggests that on average married individuals are less risky than single, divorced, individuals having free unions, etc, *ceteris paribus*. The model indicates that if the representative client was not married instead, the predicted risk would increase by 0.19% under normal economic conditions and by 1.49% under adverse conditions, other factors equal. Perhaps this may be an indication that in the context of LDCs married individuals are more likely to be concerned about acting in accordance with the terms of a contract.

Repayment history of the borrower: the variable "new" has a significant coefficient at the 1% level and as one could expect the model indicates that on average new clients are riskier than old clients, other factors equal. Under normal economic conditions, the model suggests that if the representative client was new rather than old the predicted risk would raise by 0.54%, and that under adverse conditions the impact on risk would be an increase of 4.18%, all other things equal. The coefficient of the variable "prevarrears" is also significant at the 1% level and it indicates that on average individuals with a previous history of arrears are more likely to continue to incur in arrears in future loans, other things equal. The impact of this variable on the risk of the representative client under normal conditions is 3.13% and under adverse conditions 20.46%.

Length of repayment period: the variables "repayperiod" and "repayperiod2" are both significant at the 1% level. Clearly, these variables must be interpreted together since it is not possible to vary one while keeping the other fixed. Figure 4 below shows the estimated impact of changing the repayment period of the representative client from 208 days to the number of days indicated on the horizontal axis under both normal and adverse economic conditions while holding all other things equal.

Note that the impact of changing the repayment period from 208 days to shorter periods is increasingly negative as we reduce the number of days in the period. This means that on average the risk of the representative borrower is lower at shorter loan repayment periods. The reason for this may be that the shorter the repayment periods are, the less chance that a given borrower will fall in a spell of arrears. On the other hand, the impact on risk of rising the repayment period above 208 days is positive, but note however that the impact on risk increases at a decreasing rate (see regression results). One possible reason for this phenomenon may be that there are two effects at work: a "time effect" and a "loan officer effect". Hence, although one can expect that the longer the time to maturity the greater the chance that a given borrower will fall in a spell of arrears, it may be the case that the loan officers are doing a good job at issuing "longer term"[8] loans to the safest clients, and that this fact somewhat counters the effect of repayment period on the risk of default.
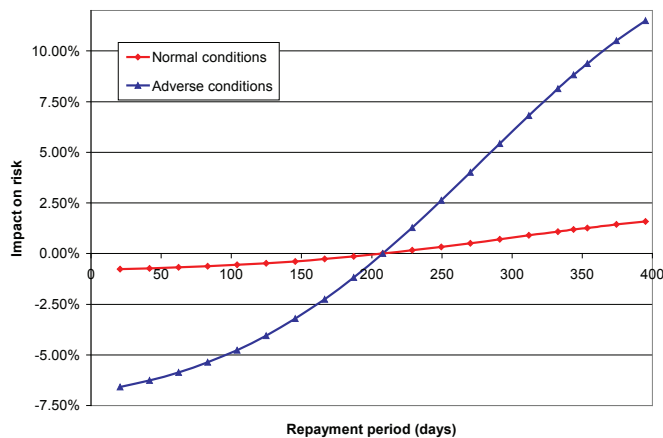
Figure 4. Impact of repayment period on predicted risk of arrears

Type of guarantee: the dummy variable "jointliability" (joint liability included in the guarantee) has a coefficient significant at the 1% level. Thus, the model indicates that, on average, loans having this type of guarantee are riskier, other things equal. This is a sensible result since loan officers are likely to require better guarantees for cases they may assess as riskier. The impact on risk from changing the attribute of the representative client from having 'jointliability' guarantee to not having one is -0.14% under normal conditions, and -1.13% under adverse conditions. Note

---

8   *In microfinance, loans tend to have very short term maturities. In our sample the average loan had a repayment period of only 208 days.*

that the coefficients of the variables "mortgage" and "personal belongings" are not significantly different from zero at any conventional level, this suggests that these variables may represent guarantees which are required as a matter of routine: the mortgages for housing loans and the personal belongings for ordinary micro-loans.

Interest rate: the variable "intrate" is significant at the 1% level of significance and hence the model suggests that on average a higher interest rate indicates that the loan is riskier than otherwise, other factors equal. Note that this variable may capture subjective observations by both the loan officer and branch manager who are likely to require a higher interest rate for cases that they regard to be more risky. Table 9 below summarises the effect on risk from changes in 'intrate'.

Table 9. Impact of monthly interest rate on repayment performance

Interest rate from	To	Normal conditions	Adverse conditions
4.80%	1.00%	-0.35%	-2.90%
4.80%	2.00%	-0.27%	-2.28%
4.80%	3.00%	-0.19%	-1.58%
4.80%	4.00%	-0.10%	-0.79%
4.80%	5.00%	0.01%	0.09%
4.80%	5.50%	0.07%	0.57%

Loan size: The variables "lnsize" and "lnsize2" are both significant at the 1% level. Figure 5 below summarises the estimated impact of changing the loan size of the representative client from $510 to the amount indicated on the horizontal axis under both normal and adverse economic conditions while holding all other things equal. Note that the estimated impact of changing the loan size from $510 to smaller amounts is positive. This means that, on average, the risk of the representative borrower is greater at smaller loan sizes, all other factors equal. On the other hand, the estimated impact of raising the loan size above $510 is negative. Hence, on average, the risk of the representative client is smaller at larger loan sizes *ceteris paribus*.
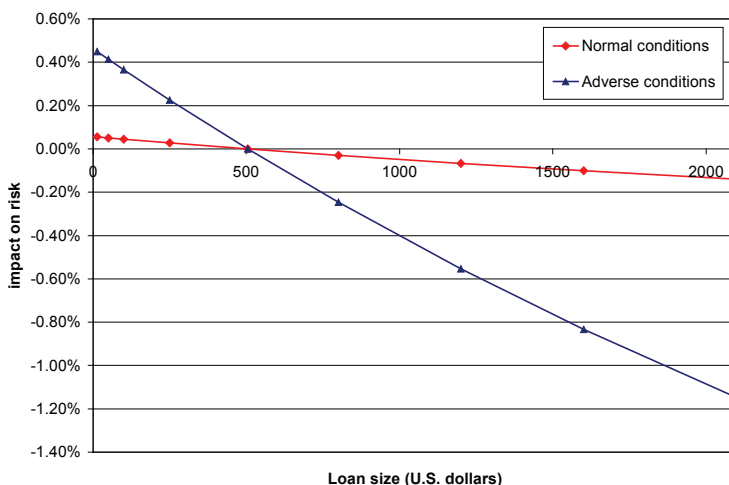
Figure 5. Impact of loan size on predicted risk of arrears

This result contrasts with that in Schreiner (2004a, p. 12). where the link between risk and the amount disbursed is found to be positive but weak in the sense that the variable's impact on risk is small. In our sample, the link between risk and the amount disbursed appears to be negative but weak.

Branch of the microfinance institution: our results in table 4 suggest that none of the lender's branches is riskier than the base category (central office) at any conventional level of significance. We interpret this result as an indication that the effect on repayment performance from changing the branch of the representative client from "branch 4" to one of the other branches is not statistically different from zero at the usual levels of significance, *ceteris paribus*.

Currency: The variable "currency" is insignificant at any conventional level of significance. Hence, our model suggests that for the period under study, the loans disbursed to the group of borrowers requesting their credit in U.S. dollars were not riskier than those disbursed to clients with credit in local currency, other things equal.

6. Conclusions

The results of the present study provide evidence that female client's repayment performance is better than that of male's at the 1% level of significance. According to the logistic regression estimated in the present paper a representative male client had a 0.15% higher probability of incurring in a spell of arrears of more than 15 days than a representative female client during the period January 2003 to June 2004 before the increase of international oil prices impacted the Nicaraguan economy, and a 1.21% higher probability of incurring in arrears greater than 15 days during the subsequent period.

However, it is important to note that the model also shows that gender has a relatively small impact on repayment risk (as defined in this specific context). In fact, the effect of gender on risk seems to be much lower than what popular wisdom would suggest specially when compared to that of other variables. For instance, the results suggest a predicted negative impact of 6.68% in the repayment performance of a representative (female) client due to the changes in the economic environment (oil price crisis). The model also indicates that an increase in the repayment period of the representative (female) client from 208 days to 395 days would have meant a 1.59% higher probability of incurring in a spell of arrears of more than 15 days during the period January 2003 to June 2004 (before the increase of international oil prices), and a 11.49% higher probability of incurring in arrears greater than 15 days during the ensuing period.

In sum, we conclude that other characteristics of the borrower as well as changes in the economic environment can have a similar or larger impact on risk than gender when it comes to repayment performance. Since in microfinance the repayment period of the loans tends to be very short, including economic conditions variables seems to be an important possibility to consider.

Appendix. A RESET test for the logit model

In this appendix we describe the steps we followed in order to perform the RESET test for the model in table 5 that includes characteristics of the borrower, the lender, the loan and a variable to allow for changes the economic circumstances.

1. We first state our null and alternative hypotheses:

    H_0: the model is correctly specified
    H_1: H_0 is false.

2. From the original model shown in table 4, we obtain the estimated latent continuous variable ($\hat{y}_i^* = \mathbf{x}_i \hat{\beta}$).

3. We then rerun the model introducing the square of the estimated latent continuous variable as an additional regressor. We call this additional variable "xbeta2" as can be seen at the bottom of table A1.

4. Since the resulting z-test for the additional regressor "xbeta2" is statistically insignificant at any conventional level of significance; we cannot reject the hypothesis that the model is correctly specified.

Table A1. Rerun of model with the addition of the square of the estimated latent continuous variable as an additional regressor

Logit estimates		Number of obs	= 48602		
		Wald chi2(42)	= 2985.81		
Log pseudolikelihood = -8940.2773		Prob > chi2	= 0		
		Pseudo R2	= 0.2272		
arrears>15	Definition	Coef.	Robust Std. Err.	z	
artisan	1 if artisan, 0 otherwise	-0.1498344	0.2014544	-0.74	
carpenter	1 if carpenter, 0 otherwise	0.2403979	0.2939691	0.82	
merchant	1 if trader, 0 otherwise	-0.197247	0.0965535	-2.04	**
seamstress	1 if seamstress, 0 otherwise	-0.4181718	0.2045315	-2.04	**
baker	1 if baker, 0 otherwise	-0.5493202	0.3688338	-1.49	
othermanuf	1 if other manufacturer, 0 otherwise	0.0138067	0.1347788	0.10	
grocery	1 if grocery, 0 otherwise	-1.614048	0.4463006	-3.62	*
transport	1 if transport, 0 otherwise	-0.5627059	0.1907459	-2.95	*
professionals	1 if professional, 0 otherwise	-0.1236178	0.1609056	-0.77	
skillwrk	1 if skilled worker, 0 otherwise	-0.5199389	0.2480817	-2.10	**
Male	1 if male, 0 otherwise	0.1633859	0.0461094	3.54	*
Married	1 if married, 0 otherwise	-0.1971569	0.0419855	-4.70	*
lnsize	Loan size in U.S. dollars, continuous variable	-0.0001326	0.0000481	-2.75	*
lnsize2	Square of lnsize	7.93E-09	2.61E-09	3.04	*
age	Age of borrower in years, continuous variable	-0.0146661	0.0021034	-6.97	*
repayperiod	Repayment period in days, continuous variable	0.0150499	0.0024188	6.22	*
repayperiod2	Square of repayperiod	-0.0000155	3.99E-06	-3.90	*
intrate	Monthly interest rate (percentage)	13.30484	4.009336	3.32	*
mortgage	1 if the guarantee includes mortgage, 0 otherwise	-0.1008657	0.2676945	-0.38	
jointliability	1 if the guarantee includes joint liability, 0 otherwise	0.1766317	0.0593417	2.98	*
persbelong	1 if the guarantee includes client's personal belongings, 0 otherwise	0.1959886	0.1327984	1.48	
currency	1 if currency is USA dollars, 0 otherwise	0.2502312	0.1673363	1.50	
new	1 if customer is new, 0 otherwise	0.4909269	0.0526477	9.32	*
prevarrears	1 if customer had arrears > 15 days in previous loan, 0 otherwise	1.570582	0.2737796	5.74	*
branch1	1 if branch is branch 1, 0 otherwise	-0.458097	0.5342824	-0.86	
branch2	1 if branch is branch 2, 0 otherwise	0.2574912	0.5320671	0.48	
branch3	1 if branch is branch 3, 0 otherwise	-0.4419576	0.5334755	-0.83	
branch4	1 if branch is branch 4, 0 otherwise	-0.3083845	0.5326621	-0.58	
branch5	1 if branch is branch 5, 0 otherwise	-0.3995918	0.5318993	-0.75	
branch6	1 if branch is branch 6, 0 otherwise	-0.5935695	0.5350369	-1.11	
branch7	1 if branch is branch 7, 0 otherwise	-0.4442774	0.5337649	-0.83	
branch8	1 if branch is branch 8, 0 otherwise	-0.4707834	0.5335974	-0.88	
branch9	1 if branch is branch 9, 0 otherwise	-0.7166945	0.5387715	-1.33	
branch10	1 if branch is branch 10, 0 Otherwise	-0.3701425	0.5350009	-0.69	
branch11	1 if branch is branch 11, 0 otherwise	-0.2705657	0.5337203	-0.51	
branch12	1 if branch is branch 12, 0 otherwise	0.1760482	0.5338904	0.33	
branch13	1 if branch is branch 13, 0 otherwise	-0.1670832	0.5342952	-0.31	
branch14	1 if branch is branch 14, 0 otherwise	-0.9048754	0.5815085	-1.56	
branch15	1 if branch is branch 15, 0 otherwise	0.0941953	0.5497091	0.17	
branch16	1 if branch is branch 16, 0 otherwise	-0.8503106	0.6458081	-1.32	
oilprice	1 if loan matures after 30/06/04, 0 otherwise	2.259686	0.2104019	10.74	*
xbeta2	Variable for RESET test	0.0017809	0.0151395	0.12	
_cons	Constant	-6.876992	0.7950128	-8.65	*

* significant at the 1% level (two sided test)

** significant at the 5% level (two sided test)

*** significant at the 10% level (two sided test)

References

Caire, D. & Kossman, R. S. (2002). Credit Scoring: Is It Right for Your Bank?. *Bannock Consulting*. Available from: http://www.microfinance.com.

Energy Information Administration (EIA). (n.d.). *A statistical agency of the U.S. Department of Energy*. Available from: http://www.eia.doe.gov/emeu/international/Crude1.xls.

Gujarati, D. (1995). *Basic Econometrics*. (3rd Ed.). McGraw-Hill Inc.

Hand, D. J. & Henley, W. E. (1997). Statistical Classification Methods in Consumer Credit Scoring: A Review. *Journal of the Royal Statistical Association*, Series A, *160*, Part 3, pp. 523–541.

Long, J. S. & Freese, J. (2003). *Regression Models for Categorical Dependent Variables using Stata*, Revised edition, Stata press.

Reinke, J. (1998). How to Lend like Mad and Make a Profit: A Micro-credit Paradigm versus the Start-Up Fund in South Africa, *Journal of Development Studies*, *34*(3), 44–61.

Rosales, R. (2002). *Tendencias recientes en la regulación del financiamiento a la microempresa*. Paper presented at the V foro interamericano sobre la microempresa, Rio de Janeiro, Brazil. Available from: http://www.iadb.org/foromic/Vforo/downloads/rosales.doc. Original in Spanish.

Roszbach, K. (2004). Bank lending policy, credit scoring, and the survival of loans. *The review of economics and statistics*, *86*(4), 946-958.

Sánchez Rizo, E. (2004, November 15). Crudo mantiene presión a economía nacional. *El Nuevo Diario*. Available from: http://www.elnuevodiario.com.ni/archivo/2004/noviembre/15-noviembre-2004/economia/economia-20041115-03.html.

Schreiner, M. (2000). Credit Scoring for Microfinance: Can It Work?. *Journal of Microfinance*, *2*(2), 105-118.

Schreiner, M. (2004a). *Scoring Arrears at a Microlender in Bolivia*. Manuscript, Center for Social Development Washington University in St. Louis. Available from: http://www.microfinance.com.

Schreiner, M. (2004b). *Benefits and Pitfalls of Statistical Credit Scoring for Microfinance*. Center for Social Development Washington University in St. Louis. Available from: http://www.microfinance.com.

Schreiner, M. (2005). *Can Scoring Help Attract Profit-Minded Investors to Microcredit?*. Financial Sector Development Conference, June 23 and 24, Frankfurt, Germany. Center for Social Development Washington University in St. Louis. Available from: http://www.microfinance.com.

Sharma, M. & Zeller, M. (1997). Repayment Performance in Group-Based Credit Programs in Bangladesh: An Empirical Analysis. *World Development*, *25*(10), 1731–1742.

Viganò, L. (1998). Credit scoring to predict loan repayment performance: an application to rural customers in Burkina Faso. In M. S. Kimenyi et al. (Eds.). *Strategic Issues in Microfinance*. London: Ashgate Publishing Ltd.

Zeller, M. (1998). Determinants of Repayment Performance in Credit Groups: The Role of Program Design, Intra-Group Risk Pooling, and Social Cohesion. *Economic Development and Cultural Change*, *46*(3), 5.