A Scoring Model of the Risk of Costly Arrears at a Microfinance Lender in Bolivia

Mark Schreiner

Oct. 1999

Microfinance Risk Management
6070 Chippewa St. #1W, St. Louis, MO 63109-3060, U.S.A.
Telephone: (314) 481-9788, http://www.microfinance.com
and
Center for Social Development
Washington University in St. Louis
Campus Box 1196, One Brookings Drive, St. Louis, MO 63130-4899, U.S.A.

Abstract

Can scoring models help microfinance lenders in poor countries as much as they have helped credit-card lenders in rich countries? I model the probability that loans from a microlender in Bolivia had arrears of 15 days or more. Although arrears in microfinance depend on many factors difficult to include in statistical models, I find that inexpensive data does indeed have some predictive power. In microfinance, computer models will not replace loan officers, but they can flag the highest risks and act as a cross-check on human judgement.

Acknowledgments

I am grateful for support from the anonymous microfinance lender in Bolivia and from a grant from the Division of Asset Building and Community Development of the Ford Foundation. I am also thankful for comments from John Adams, Jonathan Conning, Claudio Gonzalez-Vega, Guillermo Rabiela, Loïc Sadoulet, Laura Viganò, and from participants in the Third Annual Seminar on New Development Finance.

1. Introduction

Microfinance lenders supply poor people with loans that are small, short, and unsecured. Few potential borrowers have standard collateral, records in a credit bureau, or formal wage jobs. Thus, most lenders do not know low-cost ways to judge their risk. If lenders price small loans to cover the higher per-dollar costs, then they may be accused of usury, but if they charge less, then they may go in the red.

Thus progress in microfinance has been marked by new ways to cut the cost to judge the risk of small loans to poor people. For example, group lenders tap the knowledge of risk held as a sunk cost by neighbors of a potential borrower. Likewise, individual lenders reduce the need to predict risk before disbursement through frequent repayments, very small initial loans, and token collateral.

Credit-card lenders in rich countries make massive numbers of small, short, unsecured microloans each year at very low costs because they judge risk with statistical scoring models (Hand and Henley, 1997; Mester, 1997; Lewis, 1990). Why can't microlenders in poor countries do the same?

As far as I know, no microlender scores. In essence, scoring uses a few simple-to-observe objective personal traits to compare a potential borrower with past borrowers. The share of similar past borrowers who were "bad" in some sense is an estimate of the likelihood that a potential borrower will also turn out to be bad.

Scoring may help microlenders to judge risk, but it is unlikely to replace human loan officers as it did in credit cards. The most important factors in credit-card models—employment and credit record—are often unavailable in poor countries because credit bureaux are absent or because potential borrowers are self-employed.

In this paper, I test whether a simple scoring model can predict the risk of costly arrears—spells of at least 15 days—for borrowers from a microlender in Bolivia. The model pinpoints those traits that influence risk. In out-of-sample tests, it predicts better than naïve models but worse than credit-card models. Thus scoring could help to cut the costs of microfinance, not as a replacement for human judgement, but as a cross-check on it and as a filter for very good or for very bad risks.

Section 2 below gives the background for the model. Section 3 reports how traits affect arrears, and Section 4 tests predictive power. Section 5 concludes the paper.

2. Scoring for a Bolivian microlender

In this section, I discuss the market for microfinance in Bolivia, past work on scoring for microfinance, and the data and model.

2.1 Microfinance in Bolivia

Bolivia is the showcase of microfinance in Latin America. In spite of its sparse population and poverty, microfinance has a high rate of penetration (Author, 2000). Most Latin American countries have, at most, one microfinance lender with more than 10,000 borrowers; Bolivia has a dozen. Three lenders converted from unregulated not-for-profits to regulated for-profits, and several more hope to follow. Most borrowers are near the poverty line but are not among the poorest.

Profits in microfinance attracted Bolivian banks and Chilean consumer-finance companies, and by 1996, the market started to saturate. Arrears skyrocketed, and this prompted keen interest in scoring. Arrears worsened in part because consumer-finance companies tolerated high arrears and weakened the culture of repayment for all borrowers, and in part because microlenders, in the battle for market share, made loans to people already in debt elsewhere. The crisis in Brazil in 1999 also hurt repayment from the women traders who make up the bulk of microfinance portfolios.

2.2 Past work

Many models link arrears to traits of the microlender, borrower, and/or loan (Reinke, 1998; Zeller, 1998; Sharma and Zeller, 1997; Aguilera-Alfred and Gonzalez-Vega, 1993). These models, however, are not very useful for scoring for three reasons. First, they are often not robust due to small samples. Second, some use traits that most lenders do not already collect or that cost a lot to collect. Third and most important, they do not check predictive power. An out-of-sample test is needed to convince loan officers and credit managers that a computer can help to predict risk, a task whose challenge they know all too well. Most past work aims to detect traits that influence risk, not to help lenders to score potential borrowers.

Viganò (1993) is the only true scoring model for microfinance. It links default to 53 traits at a rural development bank in Burkina Faso. With a small sample ($n = 100$), prediction was checked with the jack-knife. Unfortunately, the small sample also required that the 53 traits be condensed in 13 factors, obscuring individual effects. The model has the common technical drawbacks of discriminant analysis (Eisenbeis, 1981).

The model here is an improvement in three ways. First, the sample is big. The model uses 39,956 loans repaid in 1988-96. Second, I focus less on the statistical significance of the estimated coefficients and more on the power to predict arrears for 10,555 loans repaid in 1997. Statistical significance may not imply predictive power (Hand, 1994; Greene, 1993). Third, I use only traits that most lenders already collect.

2.3 Data and model

The Bolivian microlender makes loans to urban individuals in trade and manufacturing. It evaluates risk almost exclusively with the personal judgement of loan officers; few loans are collateralized, and a credit committee discusses only very big or unusual loans. From birth in August 1988 until the end of 1996, 1,987 loans out of 39,956 (5 percent) had a costly spell of arrears, defined as a spell of at least 15 days. Such long spells are costly because they require extra enforcement efforts by the lender. In the first nine months of 1997, arrears were 8.6 percent (913 of 10,555 loans).

The data set has the following variables for all loans disbursed and paid:

- Date of disbursement;
- Amount disbursed;
- Type of guarantee;
- Branch;
- Loan officer;
- Sex of the borrower;
- Sector of the firm;
- Number of spells of arrears; and
- Length of the longest spell of arrears.

This is a short list. Most scoring models would also use the age, marital status, education, and length of residence of the borrower; ownership of a phone, house, or car; and measures of the size and financial strength of the household and firm. Thus I do a conservative test: if a truncated model works, then a full model might work even better.

The econometric model uses knowledge of the traits of past borrowers at the time of disbursement and of their subsequent repayment performance to infer future repayment risk for potential borrowers whose traits are known and who have passed the standard evaluation. The dichotomous dependent variable is unity for loans with a spell of costly arrears and zero otherwise. I use a logit model derived from a structural random-utility model (Greene, 1993). Logit avoids the weaknesses of discriminant analysis, the most common scoring technique (Reichert, Cho, and Wagner, 1983). In particular, logit directly estimates the likelihood that a given loan will go bad.

Independent variables exogenous at the moment of disbursement are derived from the data set as suggested by theory or experience. Of course, the terms of the loan contract—such as amount disbursed and guarantee—depend endogenously on the evaluation of risk by the lender. I model risk, however, conditional on having passed the lender's standard evaluation; at this point, loan terms are exogenous. Thus the model applies only to applicants already accepted. Lenders would like to model pre-evaluation risk, but they are also interested in post-evaluation risk. The lender does not have data on rejected applicants, so I cannot model pre-evaluation risk.

3. Estimated effects of individual traits

Microlenders want to predict the probability of arrears, and they also want to know which traits influence that probability. This section discusses the influence of traits, and the next section discusses predictive power.

I estimate the logit model with the 39,956 loans repaid by the end of 1996. The Chi-square statistic for the model as a whole was significant at $p = 0.0001$, and 56 of 109 estimated coefficients were significant at $p = 0.10$.

3.1 Experience as a borrower

I measure experience as the number of previous loans and as months since the first loan. Table 1 shows estimated logit coefficients and estimated changes in risk as experience changes. Positive coefficients mark increased risk, and negative coefficients mark decreased risk. Table 1 also shows means, standard errors, and p-values.

3.1.1 Number of previous loans

Looking at precisely estimated effects, the chance of costly arrears decreases with the number of past loans. For example, bad arrears are 5 percentage points less likely for an eighth-time borrower than for a first-time borrower. Given normal evaluation, borrowers who have had more loans are better risks at disbursement.

3.1.2 Months since the first loan

Experience in months since the first loan differs from experience as previous loans because, for example, a borrower could get three one-month loans or three one-year loans. I expect the effects of time to be non-linear and to fade, so I define a set of dummy variables that bound increasingly long stretches of time (Table 1).

Although not all effects are precise, the pattern suggests that risk increases with time as a borrower. The effect is big; a borrower whose first debt was 54-147 months ago is 3.3 percentage points more likely to go bad than a new borrower.

This probably reflects regression to the mean. Debtors tend to ask for their first loan during uncommonly good times when their ability to repay is at a peak. If the first loan is repaid on time, then the lender tends to press for bigger and longer loans, whether or not borrowers are still as able to repay as for the first loan. As more time passes, however, the chances increase that something will happen to worsen risk.

3.2 Arrears in the most recent loan

Past arrears should predict future arrears well. Microlenders cannot check records with a credit bureaux, but they do know the repayment performance of their own borrowers. I measure past arrears as days in the longest spell and as number of spells (Table 2). To avoid collinearity with the set of dummies for previous loans, I count first-time borrowers as if they had had no past arrears, and I also lump zero and one spells in a single dummy. Spells of arrears were common, but most were short.

3.2.1 Length of spells

Except for spells of 5-7 days, the estimated effects are precise and big; compared with no arrears, one day of arrears decreases risk by 0.024, and 31 or more days increases risk by 0.016. Why would a short spell be better than no spell? After all, common sense suggests that the effect would grow with the length of past arrears.

I do count first-time borrowers as having had no arrears in their non-existent previous loan, but this does not explain the puzzle. The pattern remains even in a model with only dummies for the length of arrears and for first-time borrowers.

Most likely, length of arrears picks up the effect of some omitted variable, or perhaps the data is in error. But the effect might be real; some arrears are due to shocks that are not the fault of the borrower, and perhaps borrowers who have had some arrears but who worked to get back on track in just a few days are, on average, better risks than those who have not yet fallen into arrears but who might not be so quick to repay once they do.

3.2.2 Number of spells

The number of spells has a big, precise effect (Table 2). Compared to 0-1 spells, risk increases for 2-4 spells and then starts to decrease. This may reflect traders who make frequent installments but who are often a day or two late, not from negligence but because they wait to combine the trip to the branch with other errands. For them, the number of spells of arrears reveals little about the risk of long arrears.

3.3 Sex

The folk wisdom in microfinance is that women are better risks than men. The Bolivian lender made most of its loans to women (Table 3), and they were indeed better risks by 0.2 percentage points, although the estimate is not very precise.

3.4 Sector

Traders received 53 percent of loans, and they were better risks than manufacturers by 0.04. Changing sectors between loans increased risk by 0.005, but the estimate is imprecise, and very few borrowers switched sectors.

3.5 Amount disbursed

The effect of the level of the amount disbursed is precise but small. In dollars as of the end of 1998, each $100 disbursed raised risk by 0.02 percentage points (Table 3).

A $100 increase in the amount disbursed between two loans had no discernable effect, but a $100 decrease did decrease risk by 0.1 percentage points. It seems the lender successfully rations borrowers suspected as bad risks.

The effect of the amount disbursed is small. Furthermore, the lender has little scope to affect arrears via loan size because the average loan is already small ($680) and because the average increases ($140) and of decreases ($25) are even smaller.

3.6 Guarantees

Of the four types of guarantees, the only one with a big, precise effect is "none" (Table 3). Perhaps only borrowers judged as very low risks in the normal evaluation can borrow without a guarantee. Changes in the guarantee do not affect risk.

3.7 Branches

All branches are not equal (Table 4). Compared with "other" (the central office and four small branches), the safest branch decreased risk by 0.013. The few borrowers who switched branches were less risky by 0.008.

Of course, the model omits some key branch-level variables such as the nature of the local neighborhood. Still, the model detects risky branches better than simple measures of arrears. The branch effect matters because branch performance is susceptible to policy, for example through bonuses or training.

3.8 Loan officers

Most microfinance lenders base their normal evaluation on the subjective judgement of loan officers. Of course, officers differ in their ability to smell bad risks, and they may take time to learn the ropes and to sharpen their sixth sense.

It turns out that risk increases as loan officers age (Table 4). The effects are big and precise; the move from 0-6 months to 148 months increases risk by 3.2 percentage points. Although loan officers learn to work smarter with time, the amount of work to do also grows as their portfolios expand. Also, the quality of new borrowers may degrade as officers mine the neighborhoods where they work deeper and deeper.

Beyond experience, loan officers differ in their ability to sense bad risks (Table 5). Compared with "other" officers (those with less than 300 loans paid off) the safest officer decreased risk by 0.048, and the riskiest officer increased risk by 0.021. Loan officers are not interchangeable parts; microfinance rests on personal relationships, so the person who an officer is important. This matters because lender policy probably has more influence over officers than over borrowers.

The 12 percent of borrowers who changed officers—usually because the officer quit—were 0.005 more risky (Table 5). Thus decreased turnover may decrease arrears.

3.9 Date of disbursement

To control for seasonal or one-shot changes in the market or lender policy, I include sets of dummies for the year and month of disbursement. Loans disbursed in the months before Christmas when business is heaviest are more risky. Compared to 1988-91, risk increased in 1992-93 before falling in 1994-96.

In sum, risk depends on sex, sector, past arrears, the experience of the borrower and of the loan officer, and the specific loan officer and branch. Seasonality and changes in policy and the market also matter. Even if a lender does not score individual borrowers, these results could help to guide adjustments to normal operations.

4. Predictive power

Scoring uses what is known from the past to guess what will take place in the future. In this section, I check how well the model built on data from 1988-96 classifies loans repaid in the first nine months of 1997.

By most measures, the model does indeed have some predictive power. Still, it is less powerful than most scoring models for credit cards. This reflects the challenge of microfinance to judge risk without reference to credit bureaux or formal wage jobs. Risk is correlated with inexpensive-to-observe traits, and lenders can use this to reduce arrears, but the link is too weak for statistics to replace loan officers completely.

For example, 5 percent of loans were bad in 1988-96, but 8.6 percent were bad in 1997. A naïve model would predict 5 percent for 1997, but the scoring model predicted 6.4 percent. One-third of increased arrears were due to changes in factors in the model.

Unlike this naïve model, scoring also estimates the risk of each loan. For example, if the Bolivian lender had used the model in 1997 and had set a rejection threshold at 0.10, then the share of bad loans would have decreased from 8.6 to 6.9 percent. With a threshold of 0.05, the share of bads would have fallen to 4.8 percent.

As the threshold approaches zero, fewer bad loans sneak through but more good loans get rejected. Scoring gives estimates of risk, but, given the estimates, lenders must choose how to balance risk against the cost to reduce it and against other goals.

If estimated risk exceeds a threshold, then a loan is classified as *bad*; otherwise, it is *good*. Classification has four outcomes. A *true positive* is a known good predicted as good. Likewise, a *true negative* is a known bad predicted as bad. A *false positive* is a known bad predicted as good, and a *false negative* is a known good predicted as bad.

For thresholds from 0 to 0.30 and for 1, the outcomes for the test of the model with 1997 data for the Bolivian lender are in Table 7. In the test sample, 913 (8.6 percent) loans were known bads, and 9,642 (91.4 percent) were known goods. As the threshold rises, true positives increase and false negatives decrease; however, true negatives also decrease, and false positives increase. Lenders choose a threshold based on the trade-offs among the four outcomes, their goals, and the consequences of mistaken and correct guesses.

The all-bad naïve model sets the threshold so low that all loans are classified as bad. The all-good naïve model sets the threshold so high that all loans are classified as good. Although the all-bad model is a straw man, the all-good model is not; it is precisely the model that the Bolivian lender now uses once it has approved a borrower through its normal evaluation.

4.1. Sample separation

The most basic test of a scoring model is whether it separates goods from bads. The cumulative distributions of estimated risk for known goods and known bads

(Figure 1) show that the model does separate to some extent. Known goods (mean 0.062, median 0.042) are always left of known bads (mean 0.098, median 0.077).

4.2 True rates

To what extent does the model separate goods from bads? The proper measure of the sharpness of separation depends on the goals of the lender (Hand, 1994; Kennedy, 1998). *True rates* are best if a lender wants to optimize the share of known goods (bads) are predicted as goods (bads). The *true positive rate* is the share of known goods predicted as goods, defined as true positives / (true positives + false negatives) (Table 7). Likewise, the *true negative rate* is the share of known bads predicted as bads, defined as true negatives / (true negatives plus false positives). Table 7 also shows the *total true rate*, all trues divided by the size of the test sample.

For the true positive rate, the all-good model (100 percent) beats the scoring model at all thresholds. On the other hand, the true negative rate of the scoring model beats the all-good model (0 percent) at all thresholds.

An all-bad model would have a true positive rate of zero and a true negative rate of 0.086. The scoring model beats these for all thresholds below 0.22.

The total true rate for the scoring model never beats the highest naïve true total rate (0.914). If the Bolivian lender only wanted to predict the most cases right, it would predict that all loans would be good. In practice, however, the loss from a false positive (a disbursed loan with costly arrears) exceeds the gain from a true positive (a good disbursed loan). Likewise, the loss avoided due to a true negative (a rejected loan that would have had costly arrears) exceeds the gain missed due to a false negative (a rejected loan that would have been good). Lenders do not weigh all outcomes the same and thus would prefer scoring to a naïve model.

Figure 2 shows the trade-off between the true positive rate and the true negative rate. The diagonal is a naïve model that predicts a varying share as bad. Scoring has more power as its curve bends farther away from the diagonal; a perfect model would trace the upper border and then the right border (Hand and Henley, 1997).

The scoring model is near the upper border for true negative rates above 0.8 and near the right border for true positive rates above 0.8. This suggests that the model would work well as a super-pass or super-fail filter. The lender could use it to approve very good risks without further ado and to flag very bad risks for more review.

4.3 Predictive values

If a lender wants to optimize the share of predicted goods (bads) that are known goods (bads), then the best measure is *predictive value*. The *positive predictive value* is the share of predicted goods that are known goods, defined as true positives / (true positives + false positives) (Table 7). Likewise, the *negative predictive value* is the share of predicted bads that are known bads, defined as true negatives / (true negatives plus false negatives). Total predictive value is the same as the total true rate.

Unlike true rates, predictive values depend on the shares of goods and bads in the sample. If the shapes of the distributions of estimated risk for known goods and bads were unchanged but if their sample shares were changed, then the predictive values would change but the true rates would not.

As an example of the difference between true rates and predictive values, suppose the Bolivian lender used the model in 1997 with a threshold of 0.10 to accept or reject borrowers who had already passed its normal evaluation. This would produce 7,791 true positives, 335 true negatives, 578 false positives, and 1,851 false negatives (Table 7). The lender would reject 21 percent of the borrowers who would otherwise have been accepted. Of these rejected cases, 15 percent (negative predictive value, 335/[335+1,851]) would have been bad. Scoring removes 37 percent of the bads (true negative rate, 335/[335+578]). The lender would accept 79 percent of borrowers, of whom 93 percent (positive predictive value, 7,791/[7,791+578]) were good. The model keeps 81 percent of the goods (true positive rate, 7,791/[7,791+1,851]).

The all-good naïve model has a positive predictive value of 0.914, a negative predictive value of zero, and a total predictive value of 0.914. Scoring does worse on total predictive value but better on both negative and positive predictive value.

The all-bad naïve model has a positive predictive value of zero, a negative predictive value of 0.086, and a total predictive value of 0.086. For all thresholds, the scoring model is better.

In sum, the scoring model predicts risk well. It separates goods from bads imperfectly, but it does tend to assign higher risk to bads than to goods. If the lender puts some weight on true negatives and does not weigh all four outcomes the same, then scoring beats the all-good naïve model currently used once a borrower is approved.

5. Conclusion

Both credit-card lenders in rich countries and microfinance lenders in poor countries make massive numbers of small, short, unsecured loans. Unlike credit-card lenders, however, microfinance lenders do not use statistical models.

Can scoring help microfinance? A scoring model of costly arrears at a lender in Bolivia suggests that it can. The model pinpoints traits that influence risk and, more important, it predicts risk better than naïve models. Still, scoring for microfinance is less powerful than scoring for credit cards, so computers and the knowledge of a few quantitative traits will probably not replace loan officers and their knowledge of qualitative character.

How should scoring be used? As usual, the math is the easy part. The difficult work is to collect the data and then to use the estimates of risk. The model here is not enough to accept or to reject applicants without a standard evaluation; risk is linked with variables in the model, but it still depends strongly on omitted variables. Also, the model starts from the premise that an applicant has already passed the normal evaluation. The model is not pre-evaluation but rather post-evaluation.

The the model is probably best-used as a super-fail filter that flags cases that have passed the normal evaluation but that yet still have very high estimated risk and thus deserve a more careful review. This will channel more attention to borderline cases where extra effort may have greater rewards.

Even lenders who do not score each borrower can still use the results from scoring models to inform policy changes. For example, the Bolivian lender might try to attract more traders because they are safer than manufacturers, they are safer. Likewise, the lender might refer to a credit committee all loans to borrowers who had a spell of arrears in their most recent loan of more than 15 days. Finally, the model isolates the pure effects of individual branches and loan officers. Incentives based on these estimates are more fair than incentives based on gross measures of arrears because, regardless of the branch or loan officer, different portfolios of borrowers have different traits that differentially predispose them to arrears.

References

Aguilera-Alfred, N. and C. Gonzalez-Vega, 1993. A Multinomial Logit Analysis of Loan Targeting and Repayment at the Agricultural Development Bank of the Dominican Republic. Agric. Fin. Rev. 53, 55-64.

Eisenbeis, R.A., 1981. Credit-Scoring Applications. In: E.I. Altman et al. (Eds.), Application of Classification Techniques in Business, Banking, and Finance, JAI.

Greene, W.H., 1993. Econometric Analysis: Second Edition, MacMillan.

Hand, D.J., 1994. Assessing Classification Rules. J. Appl. Stat. 21 (3), 3-16.

Hand, D.J. and W.E. Henley, 1997. Statistical Classification Methods in Consumer Credit Scoring: A Review, J. Royal Stat. Assoc. Series A. 160 (3), 523-41.

Kennedy, P., 1998. A Guide to Econometrics, Fourth Edition, MIT.

Lewis, E.M., 1990. An Introduction to Credit Scoring, Athena.

Mester, L.J., 1997. What's the Point of Credit Scoring? Business Review, Sept./Oct., 3-16, Federal Reserve Bank of Philadelphia.

Reichert, A.K., Cho, C. and G.M. Wagner, 1983. An Examination of Conceptual Issues Involved in Developing Credit-Scoring Models. J. Bus. Econ. Stat., 1 (2), 101-14.

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

Sharma, M. and M. Zeller, 1997. Repayment Performance in Group-Based Credit Programs in Bangladesh: An Empirical Analysis. World Dev. 25 (10), 1731-42.

Viganò, L., 1993. A Credit-Scoring Model for Development Banks: An African Case Study. Savings and Development, 17 (4), 441-82.

Zeller, M., 1998, Determinants of Repayment Performance in Credit Groups: The Role of Program Design, Intra-Group Risk Pooling, and Social Cohesion. Econ. Dev. Cult. Change, 46 (3), 599-620.

Table 1: Estimated effects of the number of previous loans and months since the first loan

Independent variable		Mean	Logit coefficients			Change in prob. of arrears		
			Estimate	s.e.	p-value	Estimate	s.e.	p-value
Intercept		1.000	-4.404	0.265	0.01			
Previous loans	0	0.460						
	1	0.247	-0.396	0.590	0.50	-0.012	0.018	0.50
	2	0.131	-0.752	0.598	0.21	-0.023	0.018	0.21
	3	0.070	-0.935	0.605	0.12	-0.028	0.018	0.12
	4	0.039	-1.056	0.619	0.09	-0.032	0.019	0.09
	5	0.022	-1.304	0.654	0.05	-0.040	0.020	0.05
	6	0.013	-1.128	0.689	0.10	-0.034	0.021	0.10
	7	0.008	-1.773	0.846	0.04	-0.054	0.026	0.04
	8	0.005	-0.866	0.799	0.28	-0.026	0.024	0.28
	9 or more	0.006	-0.820	0.803	0.31	-0.025	0.024	0.31
Months since first loan	0-6	0.466						
	7-19	0.170	0.498	0.592	0.40	0.015	0.018	0.40
	20-53	0.233	0.706	0.594	0.23	0.021	0.018	0.23
	54-147	0.125	1.089	0.602	0.07	0.033	0.018	0.07
	148 or more	0.007	1.027	0.643	0.11	0.031	0.020	0.11

Table 2: Estimated effects of past arrears

Independent variable		Mean	Logit coefficients			Change in prob. of arrears		
			Estimate	s.e.	p-value	Estimate	s.e.	p-value
Longest spell of arrears	0	0.674						
in days in previous loan	1	0.127	-0.795	0.143	0.01	-0.024	0.004	0.01
	2	0.054	-0.591	0.175	0.01	-0.018	0.005	0.01
	3	0.034	-0.581	0.191	0.01	-0.018	0.006	0.01
	4	0.028	-0.441	0.203	0.03	-0.013	0.006	0.03
	5	0.012	-0.249	0.233	0.28	-0.008	0.007	0.28
	6	0.009	0.114	0.236	0.63	0.003	0.007	0.63
	7	0.016	0.043	0.226	0.85	0.001	0.007	0.85
	8	0.007	0.563	0.246	0.02	0.017	0.008	0.02
	9	0.004	0.533	0.269	0.05	0.016	0.008	0.05
	10-14	0.014	0.406	0.205	0.05	0.012	0.006	0.05
	15-23	0.009	0.923	0.209	0.01	0.028	0.006	0.01
	24-30	0.003	0.657	0.304	0.03	0.020	0.009	0.03
	31 or more	0.007	0.524	0.246	0.03	0.016	0.008	0.03
Number of spells of arrears	0-1	0.761						
in previous loan	2	0.062	0.288	0.164	0.08	0.009	0.005	0.08
	3	0.044	0.345	0.175	0.05	0.011	0.005	0.05
	4	0.032	0.504	0.182	0.01	0.015	0.006	0.01
	5 or 6	0.041	0.249	0.177	0.16	0.008	0.005	0.16
	7 or more	0.059	0.188	0.170	0.27	0.006	0.005	0.27

Table 3: Estimated effects of sex, sector, amount disbursed, and guarantee

			Logit coefficients			Change in prob. of arrears		
Independent variable		Mean	Estimate	s.e.	p-value	Estimate	s.e.	p-value
Sex	Male	0.422						
	Female	0.578	-0.051	0.054	0.34	-0.002	0.002	0.35
Sector	Manufacturing	0.473						
	Trade	0.527	-1.329	0.075	0.01	-0.040	0.002	0.01
Changed sector		0.006	0.157	0.243	0.52	0.005	0.007	0.52
Amount disbursed	Level	676	0.000075	0.000034	0.03	0.0000023	0.0000010	0.03
	Increase	140	-0.000009	0.000066	0.89	-0.0000003	0.0000020	0.89
	Decrease	25	-0.000405	0.000154	0.01	-0.0000123	0.0000047	0.01
Guarantee	Other	0.029						
	Personal	0.475	0.063	0.106	0.55	0.002	0.003	0.55
	None	0.248	-0.306	0.118	0.01	-0.009	0.004	0.01
	Multiple	0.248	-0.132	0.124	0.29	-0.004	0.004	0.29
Changed guarantee		0.100	0.024	0.080	0.76	0.001	0.002	0.76

Table 4: Estimated effects of the branch and the experience of the loan officer

Independent variable		Mean	Logit coefficients			Change in prob. of arrears		
			Estimate	s.e.	p-value	Estimate	s.e.	p-value
Branch	Other	0.438						
	1	0.114	-0.433	0.362	0.23	-0.013	0.011	0.23
	2	0.072	-0.392	0.250	0.12	-0.012	0.008	0.12
	3	0.161	-0.317	0.208	0.13	-0.010	0.006	0.13
	4	0.044	-0.247	0.231	0.28	-0.008	0.007	0.29
	5	0.053	-0.228	0.348	0.51	-0.007	0.011	0.51
	6	0.078	-0.100	0.173	0.56	-0.003	0.005	0.56
	7	0.040	0.006	0.289	0.98	0.000	0.009	0.98
Changed branch		0.024	-0.252	0.152	0.10	-0.008	0.005	0.10
Experience of loan officer	0-6	0.062						
in months	7-19	0.204	0.202	0.118	0.09	0.006	0.004	0.09
	20-53	0.322	0.302	0.125	0.02	0.009	0.004	0.02
	54-147	0.335	0.652	0.146	0.01	0.020	0.004	0.01
	148 or more	0.078	1.058	0.203	0.01	0.032	0.006	0.01

Table 5: Estimated effects of the loan officer

		Logit coefficients			Change in prob. of arrears		
Loan officer	Mean	Estimate	s.e.	p-value	Estimate	s.e.	p-value
Other	0.116						
1	0.008	-1.564	0.339	0.01	-0.048	0.010	0.01
2	0.067	-1.240	0.228	0.01	-0.038	0.007	0.01
3	0.019	-1.229	0.245	0.01	-0.037	0.007	0.01
4	0.009	-1.226	0.370	0.01	-0.037	0.011	0.01
5	0.037	-1.078	0.231	0.01	-0.033	0.007	0.01
6	0.025	-0.820	0.200	0.01	-0.025	0.006	0.01
7	0.038	-0.800	0.223	0.01	-0.024	0.007	0.01
8	0.045	-0.788	0.241	0.01	-0.024	0.007	0.01
9	0.059	-0.768	0.199	0.01	-0.023	0.006	0.01
10	0.048	-0.647	0.227	0.01	-0.020	0.007	0.01
11	0.016	-0.631	0.308	0.04	-0.019	0.009	0.04
12	0.015	-0.594	0.353	0.09	-0.018	0.011	0.09
13	0.017	-0.547	0.331	0.10	-0.017	0.010	0.10
14	0.014	-0.537	0.388	0.17	-0.016	0.012	0.17
15	0.031	-0.497	0.219	0.02	-0.015	0.007	0.02
16	0.027	-0.460	0.224	0.04	-0.014	0.007	0.04
17	0.035	-0.443	0.185	0.02	-0.013	0.006	0.02
18	0.024	-0.394	0.180	0.03	-0.012	0.005	0.03
19	0.010	-0.231	0.228	0.31	-0.007	0.007	0.31
20	0.016	-0.205	0.199	0.30	-0.006	0.006	0.30
21	0.019	-0.167	0.294	0.57	-0.005	0.009	0.57
22	0.031	-0.118	0.313	0.71	-0.004	0.010	0.71
23	0.019	-0.062	0.264	0.81	-0.002	0.008	0.81
24	0.011	-0.021	0.357	0.95	-0.001	0.011	0.95
25	0.008	-0.011	0.331	0.97	-0.000	0.010	0.97
26	0.016	0.058	0.380	0.88	0.002	0.012	0.88
27	0.022	0.071	0.354	0.84	0.002	0.011	0.84
28	0.016	0.081	0.297	0.79	0.002	0.009	0.79
29	0.015	0.136	0.245	0.58	0.004	0.007	0.58
30	0.010	0.152	0.246	0.54	0.005	0.008	0.54
31	0.010	0.165	0.241	0.49	0.005	0.007	0.49
32	0.035	0.219	0.367	0.55	0.007	0.011	0.55
33	0.010	0.236	0.231	0.31	0.007	0.007	0.31
34	0.009	0.261	0.406	0.52	0.008	0.012	0.52
35	0.041	0.300	0.354	0.40	0.009	0.011	0.40
36	0.016	0.306	0.397	0.44	0.009	0.012	0.44
37	0.014	0.690	0.217	0.01	0.021	0.007	0.01
38	0.011	0.697	0.216	0.01	0.021	0.007	0.01
Changed officer	0.116	0.160	0.080	0.04	0.005	0.002	0.05

Table 6: Estimated effects of the month and year of disbursement

			Logit coefficients			Change in prob. of arrears		
Independent variable		Mean	Estimate	s.e.	p-value	Estimate	s.e.	p-value
Year of disbursement	1988-1991	0.083						
	1992	0.086	1.330	0.193	0.01	0.040	0.006	0.01
	1993	0.131	2.244	0.194	0.01	0.068	0.006	0.01
	1994	0.198	1.948	0.209	0.01	0.059	0.006	0.01
	1995	0.353	1.833	0.216	0.01	0.056	0.006	0.01
	1996	0.150	1.647	0.234	0.01	0.050	0.007	0.01
Month of disbursement	January	0.056						
	February	0.064	0.211	0.139	0.13	0.006	0.004	0.13
	March	0.088	0.147	0.135	0.28	0.004	0.004	0.28
	April	0.091	0.063	0.137	0.64	0.002	0.004	0.65
	May	0.102	0.096	0.136	0.48	0.003	0.004	0.48
	June	0.096	0.209	0.135	0.12	0.006	0.004	0.12
	July	0.081	0.199	0.141	0.16	0.006	0.004	0.16
	August	0.081	0.213	0.141	0.13	0.006	0.004	0.13
	September	0.087	0.289	0.139	0.04	0.009	0.004	0.04
	October	0.086	0.260	0.139	0.06	0.008	0.004	0.06
	November	0.089	0.301	0.137	0.03	0.009	0.004	0.03
	December	0.079	0.326	0.141	0.02	0.010	0.004	0.02

Table 7: Power to predict out-of-sample

		Threshold							
		All-bad							All-good
Measure	Formula	0.00	0.05	0.10	0.15	0.20	0.25	0.30	1.00
True positives	TP	0	5,343	7,791	8,976	9,330	9,491	9,561	9,642
True negatives	TN	913	646	335	173	98	52	27	0
False positives	FP	0	267	578	740	815	861	886	913
False negatives	FN	9,642	4,299	1,851	666	312	151	81	0
True positive rate	TP/(TP+FN)	0.00	0.55	0.81	0.93	0.97	0.98	0.99	1.00
True negative rate	TN/(TN+FP)	1.00	0.71	0.37	0.19	0.11	0.06	0.03	0.00
Total true rate	(TP+TN)/N	0.09	0.57	0.77	0.87	0.89	0.90	0.91	0.91
Positive predictive value	TP/(TP+FP)	0.00	0.95	0.93	0.92	0.92	0.92	0.92	0.91
Negative predictive value	TN/(TN+FN)	0.09	0.13	0.15	0.21	0.24	0.26	0.25	0.00
Total predictive value	(TP+TN)/N	0.09	0.57	0.77	0.87	0.89	0.90	0.91	0.91

Note: There are 10,555 cases, 9,642 known goods and 913 known bads.
The known-good rate is 0.914, and the known-bad rate is 0.086.

Figure 1: Cumulative distributions of estimated risk of being bad for known bads and for known goods

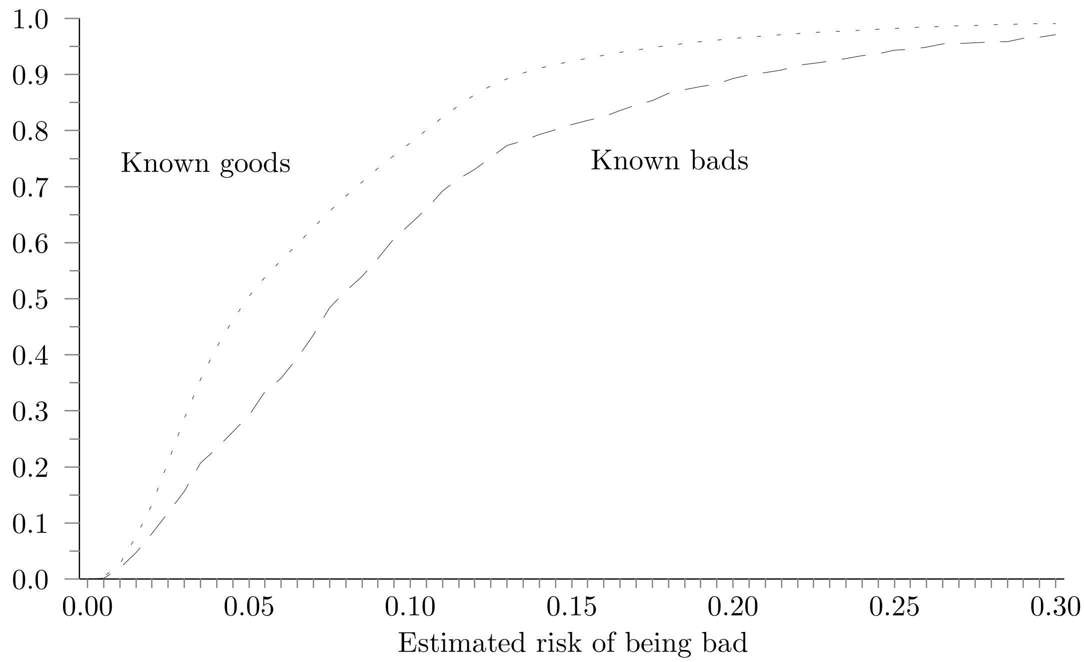

Figure 2: The trade-off between the true positive rate and the true negative rate