Can Scoring Help Attract Profit-Minded Investors to Microcredit?

Mark Schreiner

August 3, 2005

Microfinance Risk Management
6970 Chippewa St. #1W, Saint Louis, MO 63109-3060, U.S.A.
Telephone: (314) 481-9788, http://www.microfinance.com

2005 Financial Sector Development Conference
June 23 and 24, 2005, Frankfurt, Germany

Abstract

Finance in general—and microcredit in particular—is all about managing risk. Scoring quantifies repayment risk, and lenders in wealthy countries routinely use scoring to rationalize decision-making and increase profits. Can scoring help attract profit-minded investors to microcredit? Yes; explicit measures of risk facilitate informed, intentional management, and this not only increases profits but also weakens some institutional and governance barriers to private investment. At the same time, scoring uncollateralized loans for the self-employed in poor countries is less powerful than scoring credit cards, home mortgages, or car loans in wealthy countries. While scoring for microcredit is in its infancy, most adopters will probably be large microlenders who—due to competition—want to grow and improve profitability, as well as for-profit banks who want to ease their entry into microcredit markets.

Acknowledgements

I am grateful for funding from KfW Entwicklungsbank and for help from J.D. Von Pischke, Valerie Karplus, and Hanns Martin Hagen.

Can Scoring Help Attract Profit-Minded Investors to Microcredit?

1. Introduction

Microcredit is uncollateralized cash lending to the self-employed poor.[1] The central challenge of microcredit is to manage the risk that a client will behave "badly", whether by defaulting, paying late, or not returning for repeat loans. Indeed, microcredit was founded on two innovations that reduce the cost of managing these risks: joint-liability groups and careful evaluations by skilled loan officers of an individual applicant's business, chattel, and character.

Outside microcredit, wealthy countries developed a third risk-management innovation. *Scoring* relates the risk of behaving "badly" with indicators associated with the borrower (for example, type of business and debt/equity ratio) and the loan (for example, amount disbursed and number of installments). With sufficient data and care, scoring predicts risk more accurately and less expensively than non-automated methods. Moreover, scoring explicitly quantifies risk as a probability (for example, a 17-percent risk of ever reaching 30 days of arrears). Research shows that scoring increases not only profits but also the number of clients and the number of poor clients.[2] In general, scoring improves risk management, leading to a cascade of benefits.[3]

Unlike scoring in wealthy countries, scoring in microcredit has a short track record. There are pilots and proof-of-concept tests with historical data, but no documented long-term use in practice. This is a pity because the development of microcredit on a global scale has been stunted by dependence on the limited funding

---

[1] This excludes collateralized housing loans, uncollateralized consumer loans to salaried employees, and loans tied to the purchase of specific goods or services.

[2] Fishelson-Holstine (2004); Berger, Frame, and Miller (2002); Gates, Perry, and Zorn (2002); Longhofer (2002); Frame, Padhi, and Woosley (2001); Frame, Srinivasan, and Woosley (2001); Martel *et al.* (1999)

[3] Schreiner (2002a); McCorkell (1999); Lewis (1990)

available from social-minded donors. For two reasons, profit-minded investors have been biding their time. First, microcredit returns are too low to compensate for the risk of investing in a new industry. Second, there is uncertainty about exactly what are the risks of investing in microcredit.

Can scoring help attract profit-minded investors to microcredit? It can, by addressing the two barriers just mentioned. First, scoring can increase profits, especially for large microlenders who, because of competition, seek to grow and improve efficiency. Second, scoring can decrease uncertainty about the risk of microcredit as an investment because "traditional" bankers and investors understand lending based on scoring better than they understand lending based on joint-liability groups or detailed evaluations of individual applicants. Thus, if a microlender uses scoring, investors without special expertise in microcredit can more confidently evaluate investment risk. This is partly due to investors' familiarity with the scoring process and partly because scoring quantifies the risk of the microlender's main asset, its loan portfolio. Scoring also reduces uncertainty about investment risk by helping centralize decision-making, giving non-specialist investors more confidence that they can maintain effective control.

The next section explains how scoring works in microcredit and how it differs from scoring in wealthy countries. After a discussion of benefits and costs, the final two sections focus on policy issues: who has adopted scoring and why, how scoring can help attract profit-minded investors, and possible roles for donors.

2. How does scoring work in microcredit?

This section describes scoring in general as well as how scoring for microcredit differs from scoring for credit cards, home loans, and car loans in wealthy countries.

2.1 How scoring works

Scoring assigns points to indicators associated with a loan, adds up the points, and then links the points with a probability of going "bad". Figure 1 is a simple scorecard for the risk that a loan, if approved, will at some point have 30 days of arrears. Here, more points mean more risk. For the indicator "sector of business", retail or wholesale trade receives 0 points, while services or manufacturing receive 3 points. Likewise, "repeat" applicants get 0 points, and "new" applicants get 2 points. "Repeat" applicants also get one point (up to 7) for each day in the longest spell of arrears in their previous loan. (New applicants and "repeat" clients with no previous arrears get 0 points.) Finally, applicants with a savings account get 0 points, while others get 4 points. For example, a "new" applicant in retail trade without a savings account would score $0 + 2 + 0 + 4 = 6$ points. A "repeat" applicant in manufacturing who had 4 days of arrears in the previous loan and who has a savings account would score $3 + 0 + 4 + 0 = 7$ points.

The score itself is not the probability of going "bad", but it may be associated with such a probability. For the example scorecard, Figure 2 associates a score of 6 points with a risk (that is, a probability of going "bad") of 9 percent. Likewise, a score of 7 points is associated with a risk of 12 percent.

Of course, risk describes uncertain future events. Like weather forecasts, scoring is accurate if the average realized risk for a group of similar loans is "close" to predicted risk. For the example scorecard, about one in two loans with a score of 17 (risk 52 percent) should go "bad". Likewise, about one in 50 loans with a score of 0 (risk 2 percent) should go "bad".

2.2 How are scorecards constructed?

Given data and constraints on what users will accept, indicators and points are selected so that the score is highly correlated with "bad" behavior.

Scorecards vary in scope and source of information (Figure 3). In terms of scope, scorecards can be *generic* (made for multiple lenders) or *tailor-made* (custom-built for a specific lender). In terms of source of information, indicators and points can be derived based on *judgment* (expert opinion) or *data* (statistics).

Tailor-made, data-based scorecards are the most accurate (Figure 3), but they are also the most difficult to construct (and they require data).[4] Generic and/or judgmental scorecards are less accurate, but they require little (or no) data, are easier to build, may have adequate accuracy,[5] and are more readily accepted by users (especially if the users themselves help choose the indicators and points).

Only the largest microlenders who make loans to individuals[6] are ready for tailor-made, data-based scorecards. Smaller microlenders could start with generic, data-based scorecards or (more likely) tailor-made, judgmental scorecards. At the same time, smaller microlenders can prepare for the future by setting up policies and processes to improve data quantity and quality. Appendix 1 presents guidelines in this regard.

While predictive accuracy matters, it is not the only thing that matters. For example, about many scoring projects in wealthy countries fail, not because the scorecard is inaccurate but because front-line workers do not embrace scoring, whether because they did not know how to use it, because they feared it would change the status quo, or because they believed it would increase their workload or threaten their

---

[4] There is no statistical law that dictates how many historical cases are needed, but a common (probably too liberal) rule-of-thumb is at least 500 "bads" and 500 "goods".
[5] Lovie and Lovie (1986); Kolesar and Showers (1985); Stillwell, Hutton, and Edwards (1983); Dawes (1979); Wainer (1976).
[6] Scoring probably will not work for joint-liability groups or village banks. Groups are less "score-able" because they are like a diversified portfolio, their indicators are difficult to collect, and their arrears are more likely to arise from unpredictable social dynamics.

jobs.[7] Thus, scoring must be embedded seamlessly in processes, policies, and information systems that are easy-to-use and immediately and transparently beneficial to front-line workers. Success in scoring requires persistent training and follow-up with the people who use scoring.

2.3 How do microlenders use scoring?

The most common type of scoring in microcredit is *pre-disbursement*, using information available before disbursement to predict repayment problems after disbursement. A loan officer gathers information on an individual borrower as usual, applying all “traditional” screens (Figure 4). If the loan officer believes the application merits review by the credit committee, then this data—along with any credit-bureau data and savings history—is keyed-in, and the information system computes the risk of going “bad”. If the credit committee rejects the case using “traditional” criteria, it is rejected, regardless of the score. If the credit committee conditionally approves the case using “traditional” criteria, it then applies policies according to the value of the score. For example, it might waive fees to encourage loyalty for “very safe” cases with risk of less than 5 percent, proceed to disbursement as usual for “regular” cases with risk of 5–20 percent, send back for extra review “borderline” cases with risk of 21–39 percent, and reject “very risky” cases with risk of 40 percent or more.

Thus, unlike scoring in wealthy countries, scoring for microcredit does not save time by eliminating the loan officer’s “traditional”, individualized evaluation (Schreiner, 2002a). Instead, scoring adds a final hurdle, detecting some high-risk cases that slipped through standard screens. Scoring cannot approve applications, but it can reject applications that would otherwise have been approved or flag applications for additional analysis and possible modifications to the loan contract.

For three reasons, scoring does not replace loan officers and the “traditional” screening process. First, scoring for microcredit is less powerful than scoring for credit cards, home loans, and car loans in wealthy countries. There is less information about

[7] This fits the social theory of diffusion of technological innovation in Rogers (1983).

borrowers; for example, some countries lack credit bureaux, and if there are bureaux, many microcredit borrowers are not on the rolls. Furthermore, borrowers are subject to more risk from more varied sources, as they are self-employed, poor, and living in a poor country. Also, microcredit cannot fall back on collateral when it loans go sour. Microenterprises cannot present audited financial statements, nor can they be expected to fill out written applications without assistance. All this means that while scoring in wealthy countries can approve applications based on 10–15 indicators from credit-bureau reports and a few responses supplied by an applicant, scoring for microcredit must use 30–50 indicators gathered by a loan officer.

Second, loan officers are believed to be needed to sniff out applicants with qualitative characteristics (such as dishonesty) that scoring might overlook. In wealthy countries, scoring detects such qualitative characteristics indirectly via previous arrears in comprehensive credit-bureau reports, but this is not yet feasible for microcredit. Of course, as information improves, scoring will receive more weight, and loan officers may become less judges-of-character and more enumerators-of-surveys. (This will reduce costs by allowing microlenders to hire less-qualified loan officers, shorten their apprenticeships, and keep them on-board longer before burning out.) It may turn out that qualitative factors are highly correlated with quantitative factors in scorecards, further simplifying the role of loan officers. For now, however, no one knows how important loan officers' "sixth sense" is, so the conservative course is to keep them.

Third, few microlenders store data on rejected applicants. Thus, data-based scorecards can be derived only from approved applicants. But approved applicants differ from rejected applicants, partly in qualitative ways (such as dishonesty) that scoring cannot detect directly. Applying such a scorecard to applicants who have not passed the same qualitative screens will understate risk to some unknown extent, possibly leading to disaster. As noted above, if quantitative indicators turn out to be good proxies for qualitative indicators, the issue will be moot. For now, however, the safe choice is to score only applicants conditionally approved by "traditional" standards.

Later, scoring can be experimentally walked back earlier in the evaluation process. Of course, microlenders who want to do this should start storing data on rejected borrowers.

The approach advocated here is admittedly conservative and sometimes unpopular; many lenders (especially banks "downscaling" into microcredit) want to avoid costly individualized, qualitative evaluations by loan officers. Wishing that scoring could do all the work, however, does not make it so.

There is one documented case in which scoring—and only scoring—was used to "downscale" into microcredit (Rhyne, 2001); it was a disaster. Backed by private investors, a consumer-finance company entered the Bolivian microcredit market in 1995 with an evaluation process based on a scorecard derived from experience with salaried employees in Chile. Scoring allowed low-cost, rapid evaluation, and growth at first was explosive. By 1998, however, inaccurate evaluations led to high arrears and default and then bankruptcy, a textbook case of ignoring qualitative factors and blindly applying a scorecard in a context unlike the experience that generated the data it was based on.

3. What are the benefits and costs of scoring?

If scoring does not replace "traditional" evaluation, then what does it do? This section discusses benefits, costs, and how "back-testing" can estimate scoring's impacts even before roll-out.

3.1 "Back-testing"

A strength of scoring is its "trialability" (Rogers, 1983); it can be tried out in a low-cost, reversible way. The idea of "back-testing" is to build a scorecard with cases up to a certain point in time and then to apply it to later cases. Predicted risk is then compared with actual outcomes in the test period, showing how well scoring would have performed if it had been used.

Figure 5 shows results from a "back-test" for a Latin American microlender. Here, "bad" is defined as ever being in arrears on the last day of the month (the day when the lender reported to banking regulators). In the test period (January to August 2002), 36 percent of 8,549 cases were "bad" at some point.

Rejecting cases with risk in excess of the "very risky" threshold of 80 percent would lead to rejecting 1,259 "bads" and 295 "goods" (4.3 "bads" avoided per "good" lost). It also implies approving 5,137 "goods" and 1,858 "bads" (2.8 "goods" per "bad") with risk below the "very risky" threshold of 80 percent. All in all for the 8-month test period, scoring would reject 40 percent of "bads", approve 95 percent of "goods", and reject 18 percent of all cases.

The "back-test" also shows trade-offs for other "very risky" thresholds (such as 40 percent or 60 percent), showing managers the likely consequences of potential scoring policies before they are implemented. The "back-test" also provides the basic data required to estimate the impacts of scoring on profits and deal flow.

Given a "very risky" threshold of 80 percent, Figure 6 shows "back-test" estimates of changes in 30-day portfolio-at-risk. On average, scoring would have reduced portfolio-at-risk from 6.7 percent to 4.9 percent, a 27-percent reduction.

3.2 Benefits of adopting scoring

3.2.1 Profits and volume

As in the example "back-test", the most immediate benefit of pre-disbursement scoring is better risk management; fewer high-risk loans are approved.[8] While this does not save time up-front in evaluation, it does save time for loan officers in collections after disbursement, as well as reducing write-offs and loan-loss provisions. With less time in collections, loan officers have more time seek out and evaluate new clients.

This can increase profits and volume a lot. In the "back-test" in Figure 5, rejecting cases with risk above 80 percent would reduce "bads" by 40 percent. If loan officers spent two days per week on collections before scoring, they would save 6.4 hours per week (40 percent of two days). If loan officers also spent two days per week before scoring on marketing and evaluation, and if they use their new-found 6.4 hours as productively as before, they could increase evaluations by 40 percent. Netting off the 18 percent of cases rejected due to scoring (81 percent of them "bad") leaves a 40-percent decrease "bads" and about a 22-percent increase in disbursements, a win-win in terms of both profits and outreach.

The quantitative impact of scoring on profit considers decreased costs due to approving fewer "bads" as well as increased costs due to approving fewer "goods":

Change in profit = (Cost per "bad" x Number of "bads" avoided) –
(Benefit per "good" x Number of "goods" lost).

[8] The microlender can also give better service/lower prices to low-risk borrowers.

For the Latin American microlender in the "back-test", the assumed cost of a "bad" is \$100, and the assumed benefit of a "good" is \$100.[9] Figure 5 shows that an 80-percent "very risky" threshold would successfully reject 1,259 "bads" and mistakenly reject 295 "goods", for a net impact of scoring on profit in 8 months of:

Change in profit = (\$100 x 1,259) – (\$100 x 295) = \$96,400.

This gives scoring a pay-back period of less than half a year. Right now, "back-tests" such as this one and in Schreiner (2002a) are the best estimates of scoring's impact on profit and volume. As far as I know, there is no publicly available data on scoring impacts in practice, mostly because the first projects are still in their pilot stages.

3.2.2 Scoring as a teaching tool

Especially for "downscaling" banks, scoring can be a teaching tool (Caire and Kossman, 2003). Collecting data and seeing the resulting scores can help novice loan officers focus on key factors when analyzing applications.

Furthermore, the head office sets scoring policy, promoting consistency and reducing the effects of discrimination and prejudice by individual loan officers (Schreiner, 2002a). Scoring also gives new investors greater control. Microcredit loan officers are known for their autonomy, acting like mobile mini-branches with implicit—and often widely varying—credit policies. By making policy more explicit, scoring helps owners keep tabs on loan officers and detect portfolios about to deteriorate. It is also possible to detect branches or loan officers who ignore scoring policy or "cook" data, as they will approve a high share of high-risk loans and/or obtain much worse results than their borrowers' scores would indicate.

3.2.3 Quantified risk and a culture of rational decision-making

[9] This is conservative; the cost of "bads" probably exceeds twice the benefit of "goods".

Scoring quantifies risk explicitly as a probability; this more than just automates, it improves decision-making. For example, "back-testing" lets managers set scoring policy bands based on actual trade-offs among profits and deal flow. After tasting rational decision-making, they will resist seat-of-the-pants judgments for other decisions, seeking instead new ways to get more information. In short, scoring gets managers addicted to a culture of explicit, data-based decision-making.

Microlenders' data bases are useful not only for predicting repayment risk but also for promoting "mass customization" of products and services (segmentation) and strengthening client loyalty (drop-out analysis).[10] As a culture of rational decision-making takes hold, large microlenders will purposely gather data (rather than accumulating it only as a by-product of loan evaluations) and set up in-house data-analysis departments to inform decisions on a continuous basis. In-house analysis need not be fancy, and quick turn-around will let managers go beyond the bounds of the reports that the information system currently produces, allowing them to design reports that provide information to address specific business questions.

3.2.4 Beyond pre-disbursement scoring

As described above, individual microlenders will adopt pre-disbursement scoring first. Microlenders can also benefit from other types of scoring that can predict the risk of other uncertain future events.

Pre-visit scoring: Before the loan officer visits a "new" applicant in the field, data from the written application can be used to predict the risk of eventual rejection. Visits to clients who are almost certain to be rejected might be cancelled (saving the loan officer's time), or loan officers could be alerted to specific "risky" characteristics that could then be double-checked during the visit. As noted above, this is the type of scoring most microlenders think of first. Those who want to use pre-visit scoring must start to store data on rejected applicants, record precise reasons for rejections, and

---

[10] Schreiner and Sherraden (2005); Schreiner (2003, 2002c, and 2002d); Berry and Linoff (2000)

make sure that their written application collects as much information before the visit as is feasible and reliable.

Loyalty scoring: Before a borrower pays off a current loan, information up to that point can be used to predict the risk of not returning to apply for a "repeat" loan (drop-out). Figure 7 is an example policy matrix combining loyalty scoring with "traditional" rules and pre-disbursement scoring to target (costly) loyalty incentives where they matter most (desirable clients who are at-risk of drop out). Example incentives include a field visit to exhort "repeating", reduced application fees, access to a line of credit, or explicit recognition as a valued client. No incentives are offered to "kick-outs", that is, clients whose current repayment problems preclude a "repeat" loan. Among clients who qualify to "repeat", "loyalists" have low drop-out risk and so, regardless of pre-disbursement risk, do not receive loyalty incentives. Nor are incentives offered to "unsafe waverers" who are at-risk of both drop-out and repayment problems. Loyalty incentives are targeted only to desirable "safe waverers" who have a low risk of repayment problems but a high risk of drop-out.

Collections scoring: Given a loan x days in arrears, information up to that point can be used to predict the risk of reaching y days. Together with value-at-risk (amount outstanding), this can guide collections efforts. Figure 8 shows a simple example policy. Cases with high risk *and* high value-at-risk receive immediate, firm visits. Cases with low risk and low value-at-risk are left alone for a few days in the hope that they cure themselves. All other cases receive immediate visits with a gentle tone.

Collections scoring is a simple extension of pre-disbursement scoring. After all, the only new information available after disbursement concerns arrears so far in the current loan. If arrears have not yet been a problem, any collections score will be highly correlated with the pre-disbursement score. If arrears have already been a problem, then scoring is hardly needed to know that the current spell of arrears has a high risk of becoming problematic. A pre-disbursement scorecard can easily serve double-duty as a collections scorecard; there is no need for a distinct collections scorecard.

Loan-size or loan-term scoring: Some microlenders have bought scorecards that claim not to predict what loan size or term-to-maturity *will be* but rather what they *should be*. The promise of avoiding the work of setting size and term is tantalizing, but such scorecards may or may not work (and they cannot be "back-tested"). They can work only if the scorecard builder knows more than the microlender itself about setting size and term, because such scorecards cannot be based on the microlender's data.[11] This is because size and term in the past were determined by the microlender's qualitative assessment of risk (low-risk applications were approved for larger amounts and longer terms) and the size and term requested by the applicant. Setting size and term with scoring extrapolates from that context to a different one. If large, long loans in the past went to clients with exceptionally low-risk qualitative characteristics, then scoring could mistakenly approve larger and longer loans to borrowers who look equivalent in terms of quantified characteristics but who are riskier in terms of qualitative characteristics. (If scoring just sets a maximum cap on size and term—rather than sometimes increasing it—then all that is needed is a simple policy matrix applied to pre-disbursement scores.) Thus, using scoring to set size and term is like scoring applicants who have not passed through a qualitative screen using a scorecard derived from data on approved applicants who did pass a qualitative screen; it may or may not work, and "back-testing" is not available.

Net-present-value scoring: Available information is used to predict, for a given client from now on, dollar-years borrowed, number of loans, number and length of spells of arrears, risk of default, dollar-years of savings, and the usage of other services from the microlender. Then detailed data is collected from the microlender on revenue per dollar-year borrowed and the costs of evaluating and disbursing loans, collecting from clients in arrears, and provisioning for loan losses. Combining the client's predicted usage with the microlender's revenues and costs produces an estimate the client's net present value to the microlender. This enables microlenders to explicitly consider the

[11] That is, unless the lender ran an experiment to generate data for this type of scoring.

long-term bottom line in every decision related to a client. While this type of scoring is complex and has yet to be tried in microcredit, it is used routinely by banks in wealthy countries, and Schreiner (2002e) takes a first step in this direction.

3.3 Costs of adopting scoring

Beyond human resources, the most obvious cost/risk is that scoring might not predict accurately. After all, scoring for microcredit has plenty of skeptics, no documented successes, and one documented failure. Not all scorecards work, and scoring—even if used conservatively as recommended here—may still overstate risk and reject many "good" applicants. This risk can be eliminated, however, with "back-testing", and no scorecard should be used without thorough "back-testing".

Scoring how a microlender performs its central business task (evaluating risk), and like any large-scale, long-term change, it must be managed carefully and intentionally. Adopting scoring is not a project but a process, requiring commitment (not just funding) and a high-level champion. The key challenge is not predictive accuracy but rather commitment to training and follow-up.

Training helps break down employee resistance to change. To get "buy-in" from branch managers, loan officers, and key-punch operators requires training so that they know scoring works, how to use it, and its immediate benefits. Loan officers in particular cannot be expected, after years as the stars, to trust that a computer can improve on—and sometimes reverse—their judgments. Thus, they require repeated demonstrations of predictive power, both in "back-tests" and "live tests" with their own loans. A useful exercise is to show how the use of scoring would have affected their previous-month bonus.

At first, loan officers and branch managers will seize any excuse to avoid using scoring. There are several ways to address this. First, upper management must establish a simple scoring policy, document it in a "Scoring Policy Manual", and teach it to front-line workers. Second, the scoring software must be impeccable, without clunky interfaces, misspelled words, or mysterious error messages. Any visible mistake or inconvenience—however trivial—will allow users to tell themselves that the invisible inner workings cannot be trusted. Third, users will want to tinker with scoring by adding indicators, adjusting points, or making (usually cosmetic) modifications to the software. Making these changes helps users take ownership and regain a feeling of control. Fourth, upper management must closely monitor overrides and track, by branch and loan officer, deviations between predicted and realized risk, as these most likely reflect careless (or deliberately false) data collection. Fifth, management must minimize the impression that scoring creates "extra" work. This means avoid asking loan officers to collect indicators that they do not already collect until after scoring is accepted. It also means integrating scoring seamlessly in the "traditional" evaluation process and in the existing information system, which avoids the need to key-punch the same data twice or to request scoring reports manually.

On a technical level, scoring is intensely mathematical, and crunching the numbers requires software and adjustments to the information system. Furthermore, the scoring software draws on the microlender's computerized data base. All this tends to give the impression, especially at the beginning, that scoring is about information technology when in fact is about evaluating risk. Technology is a tool that makes it easier for users to accept scoring and use it properly. Adopters should guard against getting lost in the technology and forgetting the people and the process (Rogers, 1983)

Scoring has potential public-relations costs. On the surface, scoring looks inhuman, consciously trading off one person's cost versus another's benefit, statistically judging a particular individual based on his or her similarity with others, and explicitly revealing (via "back-testing") that some "good" applicants are mistakenly rejected. Of course, all lenders make these trade-offs, all evaluation methods compare individuals to others, and some "goods" are always mistakenly rejected. Scoring makes these facts of lending life explicit (and thus more susceptible to improvement), whereas "traditional" approaches allow them to be swept under the rug.

In terms of legal issues, scoring must take care to avoid indicators (such as age, gender, race/ethnicity, language, or marital status) whose use is prohibited by local law. Furthermore, microlenders who use scoring must secure data on its clients in accordance with local consumer-protection laws. (Of course, this is necessary with or without scoring.)

All in all, adopting scoring is a complex and costly process. Success hinges not on software nor on statistics but on careful change management, persistent training, and long-term follow-up. While the technical tools can be tested and tuned until they work perfectly, the human aspects are more delicate, and mistakes the first time around just make front-line workers even more skeptical that scoring can work.

Microlenders adopting scoring must not allow themselves believe that it is somehow different than other change processes. Scoring improves risk management, but it is not magic, and no lender should bet the bank on any scorecard without careful "back-testing" and without a plan for training front-line workers.

4. Who adopts scoring for microcredit and why?

So far, the adoption of scoring in microcredit been driven by networks, in particular ACCION (in Bolivia, Ecuador, and Peru) and Women's World Banking (in Colombia and the Dominican Republic). Some competitors of members of these networks in Bolivia, Colombia, and Peru have also started to adopt scoring. Finally, scoring is being adopted by some Central and Eastern European banks who are "downscaling" into small-business lending (Caire, 2004).

Scoring in microcredit is still in its infancy, perhaps because it seems—at least on the surface—so different from the two qualitative risk-management innovations that spawned microcredit in the first place. As explained by Everett Rogers in his 1983 *Diffusion of Innovations* (p. 4), "An important factor affecting the adoption rate of any innovation is its compatibility with the values, beliefs, and past experiences of a social system." While scoring scores low on compatibility with microcredit experience, it scores high on compatibility with traditional banking, consistent with "downscaling" banks making scoring a central part of their approach while only a handful of microlenders are testing scoring on a small scale. Lacking a sunk investment in traditional microcredit innovations, "downscaling" banks are more open to scoring.

Furthermore, while a strength of scoring is its "trialability" (via "back-testing") and its ability to quantify its impact on profit and outreach, an innovation is rarely evaluated "on the basis of scientific studies of its consequences, although such objective evaluations are not entirely irrelevant, especially to the very first individuals who adopt. . . . The heart of the diffusion process is modeling and imitation by potential adopters of their network partners who have adopted previously" (Rogers, p. 18).

As it turns out, the networks driving scoring in microcredit fit the theory; they are imitating banks who adopted earlier in wealthy countries. In turn, the two networks are planning, once they perfect scoring with their initial adopters, to replicate the process among their other affiliates.

The networks have several features that encouraged them to help a few affiliates adopt scoring. First, networks can exploit economies of scale; once they perfect scoring, they can divide the costs of research and development (and multiply the benefits) over many affiliates. Second, the networks have a reputation for successful innovation, increasing the likelihood that their affiliates would agree to be first adopters. Third, the networks have the time and resources to think strategically, and they are plugged into the wider microcredit community that has been talking about scoring for some time. Furthermore, the networks have had financial resources to pay the cash costs of setting up scoring systems. Fourth, the networks need ways to make themselves valuable to their affiliates, a perhaps surprisingly difficult goal that scoring serves well. Fifth, the networks compete against each other and others for status in the larger microfinance community, pushing them to embrace innovations to stay one step ahead.

Competition also motivated early-adopting affiliates. They are all large, profitable microlenders in Latin America with effective "traditional" systems in place but who want greater profits and volume to beat the competition, spread to empty markets to pre-empt competitors, and/or relax limits on loan-officer productivity.

The first few second-generation adopters were also driven by competition. These lenders in Bolivia, Colombia, and Peru are not affiliates of international networks but are adopting scoring largely because they fear falling behind their competitors, the first-adopting affiliates ACCION or Women's World Banking.

In sum, one group of early adopters has been "downscaling" banks whose background is more compatible with scoring than with "traditional" microcredit approaches. A second group of early adopters has been large, profitable affiliates of international networks who seek greater size and profitability to beat competitors. Overall, competition—among banks forced to "downscale", among networks seeking prestige, and among microlenders in a given country—has spurred adoption.

Who will adopt scoring next? Beyond the affiliates of ACCION and Women's World Banking, prime candidates include the affiliates of the German firm IPC and of the international credit-union movements (for example, WOCCU out of the United States and Desjardins out of Canada). These networks have many affiliates who lend to individuals, a reputation for quality technical assistance, and a focus on growth and profitability. Countries with large public banks that make microloans (such as India or China) may also test scoring. And finally, of course, most commercial banks trying to "downscale" into microcredit will likely include scoring as part of their strategy.

5. Conclusion: Scoring, investors, and donor roles

Like any type of finance, microcredit is tasked with managing risk. To a degree of accuracy, scoring quantifies the risk that clients will behave "badly". Lenders in wealthy countries routinely use scoring to reduce arrears and default, to target collections, to attract and keep "good" clients, and in general to increase profit and deal flow.

Can scoring help attract profit-minded investors to microcredit? Yes, and not only by increasing profits. By quantifying portfolio risk and rationalizing decision-making, scoring reduces uncertainty about the risk of microcredit as an investment. Furthermore, scoring is simpler, easier-to-learn, and more familiar for "traditional" bankers and investors than microcredit's joint-liability groups or detailed evaluations of individuals. Scoring also gives the home office greater control over branches and loan officers, increasing investors' confidence in their ability to govern a microlender. Likewise, covenants that benchmark current and future portfolio performance to scoring-based measures can assure investors that a microcredit portfolio will not suddenly sour without a chance for a pre-emptive response. By quantifying risk in different segments of the loan portfolio, scoring may also eventually open the door to securitization, facilitating liquidity and thus investment. In sum, scoring makes microcredit less an art and more a science (or at least moves the art from the loan officers in the field to the managers in the central office). This not only increases profits and volume but also reduces some of the institutional and governance barriers to investment.

5.1 Who will adopt scoring?

Scoring is most attractive to large, profitable microlenders who can spread the costs (and reap the benefits) of adoption over many clients and many years. Growth requires simplification and specialization until run-of-the-mill workers can do most tasks routinely, and scoring allows microlenders to hire less-qualified loan officers with shorter apprenticeships. Thus, scoring helps large microlenders and feeds their growth. All in all, large microlenders—especially if they face competition or want to pre-empt it—have the strongest incentives to adopt scoring and to use tailor-made, data-based scorecards. Of course, investors are likely to look at large microlenders first.

Banks "downscaling" into microcredit (or small-business lending) may also combine scoring with new (to them) qualitative, microcredit-specific techniques. This is especially the case for banks that use scoring already. Still, banks will need loan officers and qualitative methods for a while, and a lack of historical data will force them to start with generic and/or judgmental scorecards. For regulated microlenders, the need to grade loans by risk for Basel II also creates an incentive to use scoring.

Mid-size microlenders (5,000 to 15,000 borrowers) who lend to individuals and who seek to grow, become regulated, and attract private capital may start to prepare to use scoring as a central tool. This means solidifying current lending processes and information systems, as well as improving data quality and starting to record credit-bureau histories in their own information systems. Successful adoption of scoring and graduation to tailor-made, data-based scorecards might help propel some microlenders from goodness to greatness.

Finally, scoring is less attractive to small, not-for-profit microlenders who work with joint-liability groups, who provide non-financial services, and/or who lack an impetus to grow and become profitable.[12] These microlenders already have processes that work for them. Also, they resist explicit trade-offs between risk and outreach in the short term, even though scoring can mitigate such trade-offs in the long term.

5.2 What can donors do?

Ultimately, scoring is a profitable proposition for large, stable microlenders. Lenders in wealthy countries have adopted scoring without help from donors or government, and, sooner or later, microlenders will do the same. If this were the whole

[12] They may, however, use scorecards to identify poor clients (Schreiner *et al.*, 2004).

story, however, then there would be no need for this donor-funded paper, other than to tell donors to focus elsewhere.

But there are three possible short-term roles for donors. The first is to sponsor a few “proof-of-concept” projects as well as education about what scoring is and how it works. While education and demonstration would speed diffusion, they are not strictly necessary; scoring works, so microlenders will eventually adopt it on their own.

A second role is to sponsor the design of a standardized software “scoring module”. With adjustments to parameters and other minimal customization, this “scoring module” could be plugged into a specific microlender’s information system to implement generic, tailor-made, judgmental, and/or data-based scorecards (Data Mining Group, 2004; Schreiner, 2002b). This would speed up adoption by reducing start-up costs and improving ease-of-use.

The third possible role for donors is support for credit bureaux. Accurate scoring rests on good data; without credit bureaux, data is limited to what the microlender itself can collect from new borrowers or accumulate from repeat borrowers. Furthermore, credit bureaux produce a public good, so each lender wants other lenders to contribute their data (and pay for set-up costs) without contributing themselves. Thus, donors might pay initial costs and support governments in developing laws to mandate participation. Setting up a credit bureau is less innovation than replication; there are many examples around the world of how to address the challenges of uniquely identifying people, safeguarding privacy, and managing the information itself.[13]

5.3 Conclusion

Scoring helps attract profit-minded investors to microcredit by increasing profits and by decreasing uncertainty about the risk of microcredit as an investment. Scoring is a proven technique, and microlenders need not invent, only adapt. At the same time, scoring is less powerful in microcredit than in wealthy countries, and it will not single-

[13] Luoto, McIntosh, and Wydick (2004); Miller (2003); Staten (2001); Guillamón, Murphy, and Luna (2000); Jappelli and Pagano (1999).

handedly open the floodgates of private investment, nor will it replace loan officers or joint-liability groups. For large, stable microlenders, however, scoring will be routine.

References

Berger, Allen N.; Frame, W. Scott; and Nathan H. Miller. (2002) "Credit Scoring and the Availability, Price, and Risk of Small-Business Credit", Federal Reserve Bank of Atlanta Working Paper No. 2002–06, `http://www.frbatlanta.org/filelegacydocs/wp0206.pdf`.

Berry, Michael J.A.; and Gordon Linoff. (2000) *Mastering Data Mining: The Art and Science of Customer Relationship Management*, New York: John Wiley and Sons, ISBN 0–471–33123–6.

Caire, Dean. (2004) "Building Credit Scorecards for Small Business Lending in Developing Markets", Bannock Consulting, `http://www.microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf`.

Caire, Dean; and Robert S. Kossman. (2002) "Credit Scoring: Is It Right for Your Bank?", Bannock Consulting, `http://www.microfinance.com/English/Papers/Scoring_For_Your_Bank.pdf`.

Data Mining Group (2004). "Predictive Model Markup Language", `http://sourceforge.net/projects/pmml`.

Dawes, Robyn M. (1979) "The Robust Beauty of Improper Linear Models in Decision Making", *American Psychologist*, Vol. 34, No. 7, pp. 571–582.

Fishelson-Holstine, Hollis. (2004) "The Role of Credit Scoring in Increasing Homeownership for Underserved Populations", paper for "Building Assets, Building Credit: A Symposium on Improving Financial Services in Low-Income Communities", Joint Center for Housing Studies, Harvard University, `http://www.jchs.harvard.edu/publications/finance/babc/babc_04-12.pdf`.

Frame, W. Scott; Padhi, Michael; and Lynn Woosley. (2001) "The Effect of Credit Scoring on Small Business Lending in Low- and Moderate-Income Areas", Federal Reserve Bank of Atlanta Working Paper No. 2001–6, `http://www.frbatlanta.org/filelegacydocs/wp0106.pdf`.

Frame, W. Scott; Srinivasan, Aruna; and Lynn Woosley. (2001) "The Effect of Credit Scoring on Small-Business Lending", *Journal of Money, Credit, and Banking*, Vol. 33, No. 3, pp. 813–825.

Gates, Susan Wharton; Perry, Vanessa Gail; and Peter M. Zorn. (2002) "Automated Underwriting in Mortgage Lending: Good News for the Underserved?" *Housing Policy Debate*, Vol. 13, No. 2, pp. 369–391.

Guillamón, Bernardo; Murphy, Kevin X.; and Saúl Abréu Luna. (2000) "Risk mitigation as a cost-effective microfinance strategy: A Case Study of the IDB-Peru Global Microenterprise Credit Program", Inter-American Development Bank, `www.iadb.org/mif/v2/files/RiskMitigation.doc`.

Jappelli, Tullio; and Marco Pagano. (1999) "Information Sharing in Credit Markets: International Evidence", Inter-American Development Bank Working Paper No. R–371, `http://www.iadb.org/res/publications/pubfiles/pubR-371.pdf`.

Kolesar, Peter; and Janet L. Showers. (1985) "A Robust Credit Screening Model Using Categorical Data", *Management Science*, Vol. 31, No. 2, pp. 123–133.

Lewis, Edward M. (1990) *An Introduction to Credit Scoring*, San Rafael: Athena Press, LIC# 90–92258, ISBN 99–956–4223–9.

Longhofer, Stanley D. (2002) "Statement of Stanley D. Longhofer", Perspectives on Credit Scoring and Fair Mortgage Lending, Final Installment, Federal Reserve Bank of Saint Louis, `http://www.stlouisfed.org/community/assets/pdf/installment5.pdf`.

Lovie, A.D.; and P. Lovie. (1986) "The Flat Maximum Effect and Linear Scoring Models for Prediction", *Journal of Forecasting*, Vol. 5, pp. 159–168.

Luoto, Jill; McIntosh, Craig; and Wydick, Bruce. (2004) "Credit-Information Systems in Less-Developed Countries: Recent History and a Test", University of California at Berkeley, `http://are.berkeley.edu/courses/DEVELWORK/papers/Luoto.pdf`.

Martell, Javier; Panichelli, Paul; Strauch, Rich; and Sally Taylor-Shoff. (1999) "The Effectiveness of Scoring on Low-to-Moderate-Income and High-Minority Area Populations", San Rafeal, CA: Fair, Isaac Company.

McCorkell, Peter. (1999) "Credit Scoring 101", Presentation at a Federal Trade Commission Public Forum on "The Consumer and Credit Scoring", `www.ftc.gov/bcp/creditscoring`.

Miller, Margaret. (2003) *Credit Reporting Systems and the International Economy*, Cambridge, MA: MIT Press, ISBN 0–262–13422–5.

Rhyne, Elisabeth. (2001) *Mainstreaming Microfinance: How Lending to the Poor Began, Grew, and Came of Age in Bolivia*, Bloomfield: Kumarian, ISBN 1–56549–127–0.

Rogers, Everett M. (1983) *Diffusion of Innovations, Third Edition*, New York: Free Press, ISBN 0–02–926650–5.

Schreiner, Mark. (2002a) "Scoring: The Next Breakthrough in Microfinance?" Consultative Group to Assist the Poorest Occasional Paper No. 7, Washington, D.C., http://www.cgap.org/html/p-occasional-papers07.html.

Schreiner, Mark. (2002b) "Guía Técnica del Sistema de Scoring del Riesgo de Créditos", Saint Louis, MO: Microfinance Risk Management.

Schreiner, Mark. (2002c). "Who Drops Out of Microfinance Programs? A Segmentation Analysis for WWB/Cali", report to Women's World Banking.

Schreiner, Mark. (2002d). "Sales Growth Among Clients of WWB/Cali", report to Women's World Banking.

Schreiner, Mark. (2002e). "What Clients are Profitable at WWB/Cali?", report to Women's World Banking.

Schreiner, Mark. (2003) "Scoring Drop-Out at a Microlender in Bolivia", *Savings and Development*, Vol. 27, No. 2, pp. 101–118.

Schreiner, Mark; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) "The Power of Prizma's Poverty Scorecard: Lessons for Microfinance", http://www.microfinance.com/English/Abstracts/Scoring_Poverty_in_BiH.htm.

Schreiner, Mark; and Michael Sherraden. (2005) "Drop-out from Individual Development Accounts: Prediction and Prevention", Center for Social Development, Washington University in Saint Louis, http://www.microfinance.com/English/Papers/IDAs_Drop_Out.pdf.

Staten, Michael E. (2001) "The Value of Comprehensive Credit Reports: Lessons from the U.S. Experience", paper presented at the World Bank conference on "Profiting from Small Business Lending", April 2–3, Washington, D.C., www.worldbank.org/wbi/banking/creditscoring/.

Stillwell, William G.; Barron, F. Hutton; and Ward Edwards. (1983) “Evaluating Credit Applications: A Validation of Multi-attribute Utility Weight Elicitation Techniques”, *Organizational Behavior and Human Performance*, Vol. 32, pp. 87–108.

Wainer, Howard. (1976) “Estimating Coefficients in Linear Models: It Don’t Make No Nevermind”, *Psychological Bulletin*, Vol. 83, pp. 213–217.

Appendix 1: Collecting data for scoring in the future

Most microlenders do not yet have adequate data to support tailor-made, data-based scorecards. Still, with growth and planning, they can accumulate the necessary data, perhaps using judgmental, tailor-made scorecards in the interim. Improving data quality and quantity is unglamorous work, but successful scoring depends on it. This appendix offers some guidelines drawn from Schreiner (2002a).

Most microlenders who make individual loans already collect most of the indicators listed below. Assuming the indicators are recorded accurately for a sufficient number of cases, they should be enough for a tailor-made, data-based scorecard.

In general, more indicators improves accuracy, and all microlenders who plan to use scoring should start from now on to record (in electronic form) at least credit-bureau data, quantified measures of loan officers' subjective judgments, and indicators of household demographics and assets.

<u>Credit-bureau data</u>

- Arrears in current and paid-off loans
- Amounts owed to current creditors
- Number of inquiries in past year
- Identity of current and former creditors
- Dates disbursed (and dates paid off) of current and paid-off loans
- Amounts disbursed for current and paid-off loans
- Monthly installments for current and paid-off loans
- Maximum line of credit with current and former creditors

<u>Quantification of loan officers' subjective judgments</u>

Loan officers record their subjective impressions of applicants on a five-point scale (far below-average, below-average, average, above-average, or far above-average).

- Overall credit risk
- Honesty and transparency of responses
- Quality of references
- Entrepreneurial skill and oomph
- Prospects for line of business
- Variability and seasonality of cash flows
- Extent of recent investment in the home and business
- Grasp of the rules in the loan contract
- Quality of family relationships and informal support networks

Applicant demographics:

- Year of birth
- Gender
- Marital status (married, cohabiting, single/never-married/never-cohabited, divorced/separated/no-longer-cohabiting, widowed)
- Year of most recent change in marital status
- Last grade completed in school
- Number of household members (including applicant) age 18 or older
- Number of household members age 17 or younger
- Number of household members with salaried employment
- Number of enterprises run by household members

Contact information:

- Phone number to contact at home
  - Is it a neighbor's phone?
  - Is it a cell phone?
- Phone number to contact at the business
  - Is it a neighbor's phone?
  - Is it a cell phone?
  - Is it different from the phone number to contact at home?
- e-mail address

Household assets:

- Home tenure (owner, renter, other)
- Year took up current residence
- Number of rooms (excluding bathroom and kitchen)
- Hectares of non-homestead land owned
- Housing materials (context-specific)
  - Roof
  - Floor
- Household services
  - Electricity connection
  - Source of drinking water (pipe/well/other)
  - Fuel used for cooking (gas/electricity/wood)
  - Toilet (latrine/indoor with piped water/indoor with bucketed water/other)
- Vehicles (working)
  - Automobile, tractor, truck, or bus
  - Motorcycle
  - Bicycle

- Appliances (working, context-specific)
    - Refrigerator
    - Color television
    - Washing machine (clothes)
- Formal savings account
    - Date opened
    - Passbook or time
- Frequency of receipt of remittances

Business "demographics":
- Sector (manufacturing, services, trade, agriculture, other)
- Specific type of business (list of 30–50 context-specific)
- Year started
- Formalization
- Tenure of locale (owned, rented, other)
- Person-months of full-time-equivalent workers per year
    - Applicant
    - Family members (excluding applicant)
    - Non-family members

Business financial flows (monthly):
- Sales
- Expenses
- Installments on other debts (business and household)

Business financial stocks:
- Cash and savings-account balances
- Inventory
- Fixed assets
- Debts
    - Formal
    - Informal

Repayment record for each scheduled installment:
- Date due
- Date paid-off

Aspects of the loan contract:

- Date application submitted
- Date loan disbursed
- Date paid-off
- Amount disbursed
- Amount of average installment
- Number of installments
- Frequency of installments
- Refinanced status
- Type of guarantee
- Value of guarantee
- Identity of cosigner

Identity of the lender:

- Branch
- Loan officer

Figure 1: Simple scorecard for risk of having at some point 30 days of arrears

Indicator	Indicator values		Points
A. Sector of business	Retail or wholesale trade	Service or manufacturing	
	0	3	
B. Experience of applicant	"Repeat"	"New"	
	0	2	
C. Days in longest spell of arrears in previous loan	"New" applicant or no arrears	"Repeat" applicant with arrears	
	0	1 per day, up to 7	
D. Savings account	Yes	No	
	0	4	
		Total:	

Figure 2: Association between scores and the risk of at some point having 30 days of arrears

Score	Risk (%)
0	2
1	3
2	3
3	4
4	5
5	7
6	9
7	12
8	14
9	17
10	22
11	24
12	27
13	31
14	35
15	40
16	45
17	52

Figure 3: Trade-offs among accuracy, acceptance, and cost/difficulty according to a scorecard's scope and source of information

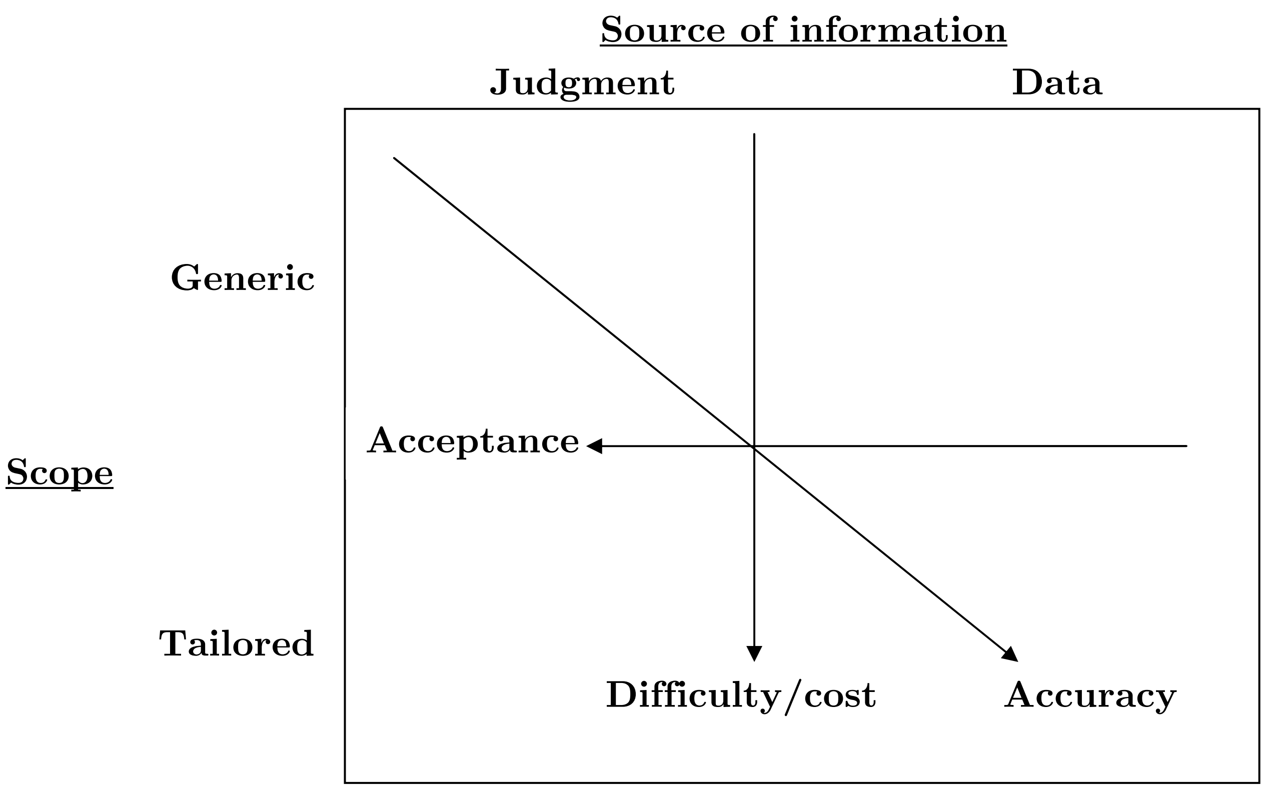

Figure 4: Process of pre-disbursement scoring in microcredit

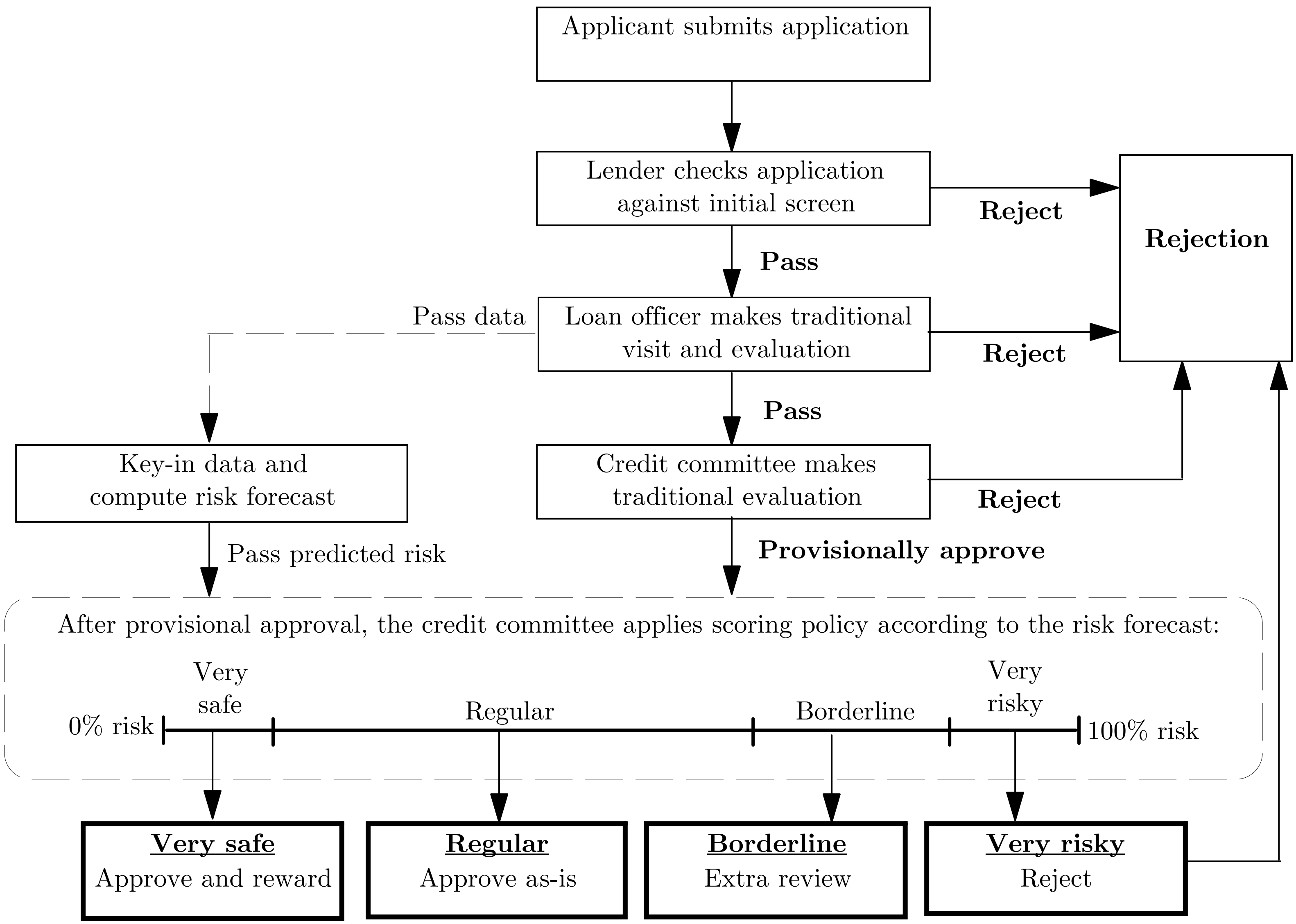

Figure 5: “Back-test” for Latin American microlender, scorecard made with cases up to December 2001, applied to cases from January to August, 2002

Criteria	Formula	'Very risky' threshold					
		0	20	40	60	80	100
'Bads' avoided	A	3,117	2,967	2,591	2,020	1,259	0
'Goods' lost	B	5,432	3,904	2,223	1,056	295	0
'Bads' avoided per 'good' lost	A/B	0.6	0.8	1.2	1.9	4.3	#N/A
'Goods' approved	C	0	1,528	3,209	4,376	5,137	5,432
'Bads' approved	D	0	150	526	1,097	1,858	3,117
% 'bads' avoided	100*A/(A+D)	100	95	83	65	40	0
% 'goods' approved	100*C/(C+B)	0	28	59	81	95	100
% cases rejected	100*(A+B)/(C+A+D+B)	100	80	56	36	18	0

Note: 8549 cases, with 5432 (64%) observed 'goods' and 3117 (36%) observed 'bads'.

Figure 6: "Back-test" of effects on portfolio-at-risk due to scoring with an 80-percent "very risky" threshold for Latin American microlender, scorecard constructed with cases up to Dec. 2001 and applied Jan.–Aug. 2002

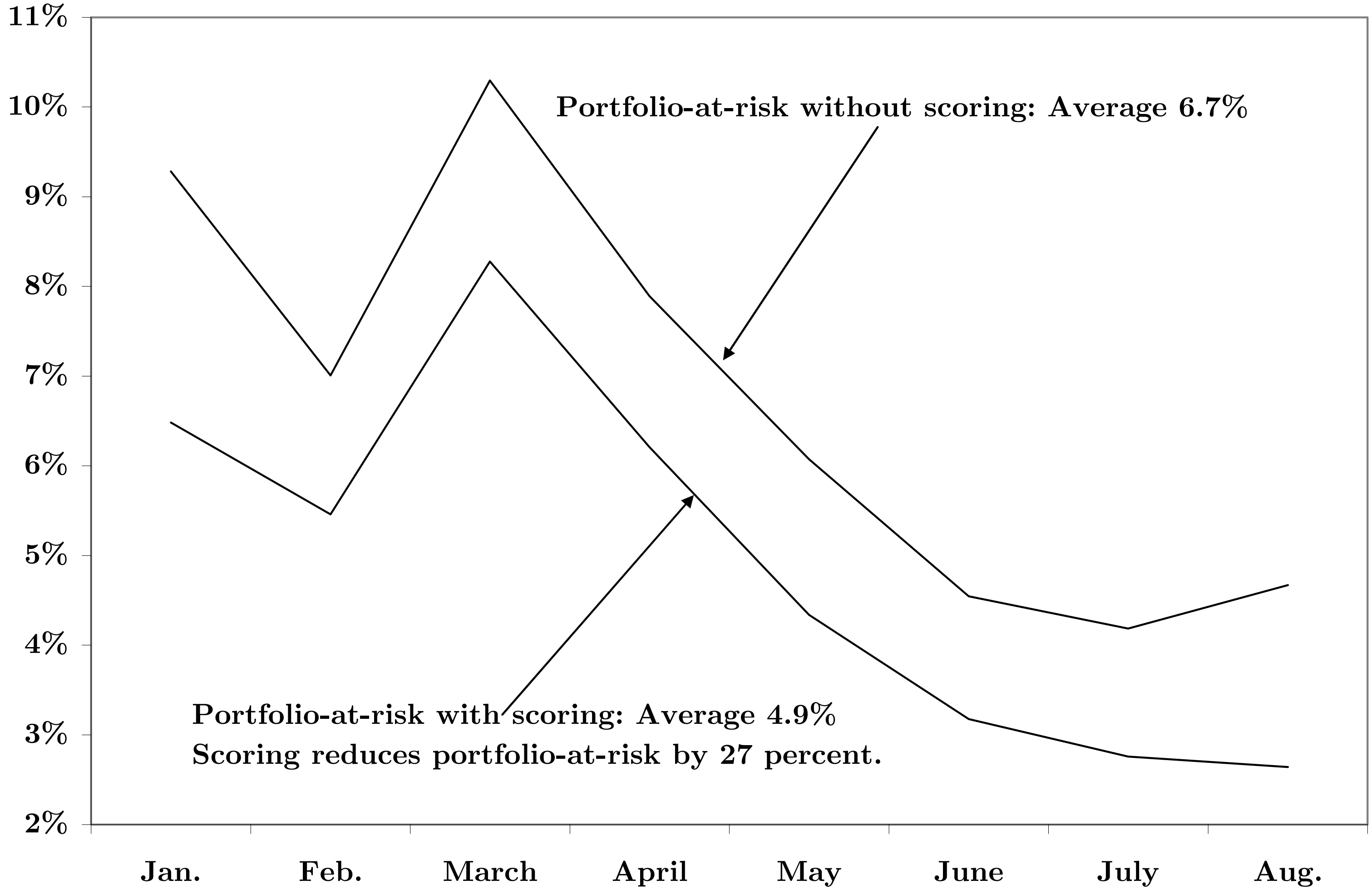

Figure 7: Example policy matrix for loyalty incentives

<table>
<tr><td colspan="2" rowspan="2"></td><td rowspan="2">Disqualified
under “traditional” rules</td><td colspan="2">Qualified
under “traditional” rules</td></tr>
<tr><td>High pre-disbursement risk</td><td>Low pre-disbursement risk</td></tr>
<tr><td rowspan="2">Drop-out risk</td><td>Low</td><td rowspan="2">“Kick-outs”
No incentives</td><td colspan="2">“Loyalists”
No incentives</td></tr>
<tr><td>High</td><td>“Unsafe waverers”
No incentives</td><td>“Safe waverers”
Incentives offered</td></tr>
</table>

Figure 8: Example policy matrix for collections efforts