All Good?: How Credit Scoring Works in Myanmar

Dean Caire, CFA

December 2018

This paper examines the role of credit scoring in very low-default environments based on experience in both commercial banks and MFIs in Myanmar , between 2016-2018. It argues that, even in low-default environments, financial institutions stand to gain from digitization and automation of the loan-application process and from the consistent and efficient application of credit policy that a credit-scoring model facilitates. However, quantitative risk measurement is not possible until financial institutions have lent to a sufficient number of borrowers who actually fail to pay back. In some market segments in Myanmar, such as traditional microfinance, such borrowers have been very hard to find. The paper shares a few examples of the unique features of Myanmar and suggests how the 'ecosystem' for credit scoring may be changing in the near-term future.

Introduction

Credit scorecards rank loan applicants based on their expected likelihood to repay a loan on time. Creating such a ranking is technically possible only if some past borrowers did not repay on time. When historic repayment rates are at or near 100 percent[1], the historic data suggests that all successful loan applicants are 100% likely to repay a new loan.

One way to break out of this virtuous cycle is to 'lend to learn'[2]. In other words, a financial institution lends to some borrowers who would otherwise be denied credit (by current policies) with a primary goal of learning—will they also repay as promised?

The willingness to lend and possibly lose money is sometimes called 'risk appetite'. In developed and competitive financial markets, lenders have been hungry to expand and segment the credit market—finding there is more money to be made in measuring and pricing the risk of non-repayment into various products and business models targeting various types of borrowers—such as prime (lower risk) and subprime (higher risk).

Myanmar's regulators have helped to curb financial institutions' risk appetite in a variety of ways including:

---

[1] See chapter 11 'Microfinance in Burma' in Turnell, Sean. *Fiery dragons: Banks, moneylenders and microfinance in Burma*. No. 114. NIAS Press, 2009.

[2] The term 'lend to learn' was borrowed from Mark Flaming, Chief Digital Officer of Yoma Bank, December 2018 (personal meeting).

- a floor on interest rates on deposits (8% for banks and 10-15% for MFIs) and caps on loan interest rates (13% for banks and 30% for MFIs)
- various limits (by products and sectors) on loan size and tenor
- Lack of a credit bureau [3]

Nevertheless, banks and MFIs in Myanmar continue to develop new credit products and delivery channels, and many are looking at how credit scoring technology can help to assess repayment risk. The body of this paper shares some simple credit-scoring lessons learned from work with financial institutions supported by the USAID private sector development project (PSDA).

Note: Since this paper was first drafted, the Central Bank of Myanmar has eased bank lending rates somewhat, and allows rates of up to 16% for unsecured loans or loans not secured with traditional collateral (real estate, cash, and cash equivalents). This suggests that banks may want to be looking more loans at that would be riskier so credit scoring could be a more significant tool in credit evaluation.

How Credit Scoring Works in the 'Textbook' Case

The Myanmar Times reported in December 2018 that Myanmar Credit Bureau signed an agreement with Equifax to establish a credit bureau and to provide a 'loan data service'.[4] Equifax describes credit scoring as follows:

> A *credit score* is a tool used by lenders to help determine whether you qualify for a particular credit card, loan, mortgage, or service. Using the information on your credit report and any additional information you supplied as part of your application, lenders use a mathematical model to calculate a score that represents your credit history. This helps to indicate what kind of borrower you are, and how likely it is that you will manage your repayments.[5]

The 'mathematical model', commonly called a 'credit scorecard', is developed by looking at a sample of data on past loans that have been classified as either 'good' or 'bad'.[6] When such data on past loans is available, there are a myriad of 'machine-learning' methods available to build credit scorecards, all with surprisingly similar results (and thus favoring the use of

---

[3] These, and many more examples of regulation are presented in "Myths and Maths:
The Impact of Financial Regulations on Agriculture in Myanmar" by Roger Thomas Moyes and Kenneth Shwedel; https://www.mekongbiz.org/wp-content/uploads/2017/06/The-Impact-of-Financial-Regulations-on-Agriculture-in-Myanmar.pdf.

[4] https://www.mmtimes.com/news/credit-bureau-be-and-running-within-next-12-months.html#.XBEXCQOKpQI.linkedin; Accessed December 15, 2018.

[5] https://www.equifax.co.uk/resources/what_we_do/what-is-a-credit-score.html; Accessed December 15, 2018.

[6] "Credit Scoring" section of the International Financial Corporation's new handbook: DATA ANALYTICS AND DIGITAL FINANCIAL SERVICES (pp. 79-100). http://www.ifc.org/wps/wcm/connect/22ca3a7a-4ee6-444a-858e-374d88354d97/IFC+Data+HandBook+FINAL.pdf?MOD=AJPERES

relatively simple models).[7] Yet, when most or all loans are 'good', as they seem to be in Myanmar[8], humans, and not machines, must build the first scorecards.[9]

Credit scorecards built without reference to data on past 'good' and 'bad' loans are called 'Expert Scorecards'. Experienced lenders/consultants (the "experts") build these scorecards by identifying a comprehensive set of borrower characteristics they believe, from experience, are related to loan repayment. A point scheme is used to indicate the strength and direction of each characteristic's expected relationship to credit risk. Table 1 shows an example of a 'one-factor' risk-ranking model for "Years in Business"—where the more years in business, the more points assigned, and the lower the expected credit risk.

Table 1: An Expert Score for 'Years in Business'

CRITERIA	SCORE
> 5 years	10
> 1 to <=5 years	5
<=1 year	0

Suppose an expert scorecard is built with "Years in Business" (from Table 1) and 4 other factors, also each with scores of 0 to 10 points. In this case, the scorecard's points range from 0, for highest risk, to 50, for lowest risk.

How can the lender know if its expert scorecard is 'ranking risk' accurately?

For illustration, assume:

- The lender scores and disburses 1,000 12-month term loans.
- A year later, 50 of the loans were 'bad'[10] (seriously delinquent) and the other 950 'good'. The average 'bad' rate, or risk, in the portfolio is 5% (or 50/1,000 = 5%).

Table 2 presents counts of the 'good' and 'bad' loans grouped into three score ranges

Table 2: Scorecard Results with Three Risk Groups

Credit Score	>=	<=	Goods	Bads	Total	Bad Rate
Low Risk	33	50	248	2	250	1%
Average Risk	18	32	500	25	525	5%

[7] David Hand, 'Classifier technology and the illusion of progress', Statistical Science, Vol. 21.1 (2006): 1-14
[8] In the author's personal experience
[9] because using past data a machine will not anticipate possible changes in risk and thus will continue to estimate present near-perfect repayment rates.
[10] Say, for example, borrowers were delinquent for 90 consecutive days or more at some point during the life of the loan. However, the 'bad' definition used for scorecard modelling should be determined for each business case as a loan that is loss-making and the financial institution would chose to avoid in the future.

High Risk	0	17	202	23	225	10%
TOTAL			950	50	1,000	5%

The bad rates (number of bad loans divided by total loans in the group) confirm that scorecard accurately 'ranks risk' because:

- high scoring loans (33-50 points) are lower-than-average risk (a bad rate of 1%).
- The majority of loans score between 18-32 points and have 'average' risk (a 5% bad rate, equal to the total portfolio bad rate).
- low scoring loans (0-17 points) are higher-than-average risk (a bad rate of 10%).

How to use the risk-ranking information?

In this 'textbook' case, it is clear how the scorecard could help make lending decisions. Specifically, in a Myanmar context, consider the following assumptions:

1. The average loan amount is expected to be 1,000 currency units (column F)
2. The average interest rate margin is 5% (13% - 8% cost of funds, Column G)
3. Bad loans lose the full principal value and no interest is earned (Column I)

In such a scenario, Table 3 illustrates how the lender should not extend credit to borrowers in the high risk group, where:

Column H indicates the interest income expected per risk band (H=D*G*F)
Column J indicates the total expected loss per risk band (J = I * C)
Column K indicates the expected gross margin (K = H – J)

Table 3: Gross Margin Framework for Determining Decision Policy

A	B	C	D	E	F	G	H	I	J	K
Risk Group	Goods	Bads	Total	Bad Rate	Ave Loan Amo	Ave Margin	nterest Income	e. Total Charge	Total Loss	Gross Margin
Low	248	2	250	1%	1,000	5.0%	12,400	1,000	2,000	10,400
Average	500	25	525	5%	1,000	5.0%	25,000	1,000	25,000	0
High	202	23	225	10%	1,000	5.0%	10,100	1,000	23,000	-12,900
	950	50	1,000	5%			47,500		50,000	-2,500

The example shows that with Myanmar's capped interest margin of 5% for banks, lending to any risk band with an expected bad rate exceeding 5% would be loss-making on average. In 'standard' retail-banking environments, higher interest rates might be charged to lend to the risk groups with higher bad rates. In the absence of such risk-based pricing, rejecting the 225 applicants scoring in the high-risk group would improve the lender's gross margin (and lower its average bad rate to 3.5%, not shown in the table).

Challenges of Credit Scoring in Myanmar

What happens in the previous example if the lender uses an expert scorecard to score and disburse 1,000 12-month term loans, but after one year only 3 of the loans are 'bad'? Furthermore, what if the causes of non-repayment for those three 'bad' loans are known and are not related to the characteristics being used to measure credit risk, namely:

1. One borrower had family problems and was forced to stop running his business.
2. A second borrower decided to 'take the money and run'—deliberately changing his name and fleeing to another community (that is, an act of deliberate fraud, which, when possible, would be measured separately by a fraud risk model).
3. A third borrower had considerable undisclosed borrowing from other financial institutions—something lenders in most markets would understand from a credit-bureau report, but which will still not be available to banks and other lenders in Myanmar for at least a year.

First of all, the scorecard results table might look like Table 4.

Table 4: Scorecard Results with Three Risk Groups and Three Bad Loans

Credit Score	>=	<=	Goods	Bads	Total	Bad Rate
Low Risk	33	50	248	2	250	0.8%
Average Risk	18	32	524	1	525	0.2%
High Risk	0	17	225	0	225	0.0%
TOTAL			997	3	1,000	0.3%

Do those results mean the scorecard does not work? Empirically, it ranks risk backwards, exactly wrong. But, not only are there far, far too few bad loans to reasonably perform such analysis, also two of the three 'bad' loans are extraordinary cases (and in Myanmar, often family members repay loans in hardship cases, including death of the borrower).

Learning from Experience: How to Improve Credit Scoring Results

Assuming for a moment that a financial institution's funding and financial capacity are unlimited, the gross margin framework presented in Tables 5 and 6 suggests what may not be well understood by risk averse financial institutions—namely that taking more risk can increase profits, even as portfolio 'bad rates' rise.

Table 5: **Gross Margin Framework for Low-Risk Portfolio**

A	B	C	D	E	F	G	H	I	J	K
Risk Group	Goods	Bads	Total	Bad Rate	Ave Loan Amo	Ave Margin	nterest Income	. Total Charge	Total Loss	Gross Margin
Low	248	2	250	1%	1,000	5.0%	12,400	1,000	2,000	10,400
Average	524	1	525	0%	1,000	5.0%	26,200	1,000	1,000	25,200
High	225	0	225	0%	1,000	5.0%	11,250	1,000	0	11,250
	997	3	1,000	0.3%			49,850		3,000	46,850

With 997 good and 3 bad loans (and the same interest rate and cost assumptions presented earlier), gross margin is 46,850. Table 6 presents a possible result of making an additional 1,000 loans (or a total of 2,000 loans) where risk has increased to 2.5% (or 50 bad loans). Although the delinquency rate of the portfolio in Table 6 is much higher (2.5% vs. 0.3%), gross margin has (slightly) increased to 47,500.

Table 6: **Gross Margin Framework for Higher-Risk Portfolio**

A	B	C	D	E	F	G	H	I	J	K
Risk Group	Goods	Bads	Total	Bad Rate	Ave Loan Amo	Ave Margin	nterest Income		Total Loss	Gross Margin
Low	448	4	452	0.9%	1,000	5.0%	22,400	1,000	4,000	18,400
Average	922	16	938	1.7%	1,000	5.0%	46,100	1,000	16,000	30,100
High	580	30	610	4.9%	1,000	5.0%	29,000	1,000	30,000	-1,000
	1,950	**50**	**2,000**	**2.5%**			97,500		50,000	47,500

But the greater return to the bank in this example is not the additional 650 in gross margin (47,500 vs. 46,850). It is in the risk-taking and learning that will allow the lender to use its scoring model—which in Table 6 also ranks risk 'correctly' (Low, Average, and High risk groups have bad rates of 0.9%, 1.7% and 4.9% respectively).

Once there are enough delinquent loans[11], the lender will be able to find a decision policy that maximizes its profit. Using the results of the 2,000 loans in Table 6 to adjust a model, the bank may be able anticipate better results for its next wave of loans by making the changes shown in Tables 6a and 6b.

Table 6a: **Gross Margin Framework after Adjusting the Scorecard**

A	B	C	D	E	F	G	H	I	J	K
Risk Group	Goods	Bads	Total	Bad Rate	Ave Loan Amo	Ave Margin	nterest Income		Total Loss	Gross Margin
Low	448	4	452	0.9%	1,000	5.0%	22,400	1,000	4,000	18,400
Average	1,072	20	1,092	1.8%	1,000	5.0%	53,600	1,000	20,000	33,600
High	430	26	456	5.7%	1,000	5.0%	21,500	1,000	26,000	-4,500
	1,950	**50**	**2,000**	**2.5%**			97,500		50,000	47,500

Table 6a is an example of adjusting the scorecard risk groupings so that the 'High' risk group has a bad rate of 5.7% (vs. 4. 9% in Table 6).[12] Table 6b shows that with this adjusted model, it can potentially materially improve its gross margin by rejecting all 'High' risk borrowers – which when applied to the past data would have increased its gross margin by nearly 10%

[11] Which in reality is more likely to be a number like 500, rather than 50, delinquent loans—to facilitate the use of sampling in scorecard development

[12] Scorecard power comes from large differences in bad rates in different groups. Technically these differences (and scorecard power) can be increased by better modelling the individual features in the model and by adjusting the risk groupings based on total score.

((52,000 – 47,500)/47,500 = 9.5%) and lowers the future expected portfolio bad rate by nearly a full percentage point to 1.6% (from 2.5% in Table 7a).

Table 6b: **Gross Margin Framework after Adjusting the Lending Policy**

A	B	C	D	E	F	G	H	I	J	K
Risk Group	Goods	Bads	Total	Bad Rate	Ave Loan Amo	Ave Margin	nterest Income		Total Loss	Gross Margin
Low	448	4	452	0.9%	1,000	5.0%	22,400	1,000	4,000	18,400
Average	1,072	20	1,092	1.8%	1,000	5.0%	53,600	1,000	20,000	33,600
High			0	0.0%	1,000	5.0%	0	1,000	0	0
	1,520	**24**	**1,544**	**1.6%**			76,000		24,000	52,000

But how can the lender know if the scorecard will in fact produce similar results for the next wave of borrowers? It can only develop a degree of confidence in a model through out-of-sample testing. In other words, once there is enough data to build a statistical model, it should be built using on part of the data (commonly 60-70% of the historic data) and tested (or 'validated') on the rest of the data (the 30-40% not used to build the model). If results are similar on for those two sets of historic data, there can be a degree of confidence that the model will produce similar results in the future. Nevertheless, the future may not always resemble the past, for many reasons, and all scorecards should be monitored, periodically validated, and adjusted or re-developed as appropriate.[13]

All Good: How Can Credit Scoring Work in Myanmar?

As long as most all loans are 'good', scoring may still work to:

1. **Improve the collection and (ideally) digitalization of borrower data relevant to credit risk assessment.** Even if the lack of delinquent loans so far has not confirmed the value of different types of data for today's credit risk models, worldwide experience indicates that someday such data will add value. In any case, better borrower data can also enable more sophisticated segmentation and product and service offer-differentiation.
2. **Enforce consistency and efficiency in the lending process.** Many financial institutions (both banks and MFIs) until now have used relatively labor-intensive processes to assess borrowers of relatively small amounts. The introduction of a scorecard (even an expert one) can focus the analysis of loan officers and promote consistent loan decisions across loan officers, branches, and regions.
3. **Introduce the foundations of a quantitative approach to credit-risk measurement and management.** While it may still take time develop, eventually competitive pressures should expand the credit market in Myanmar to a point where more borrowers are unable repay their loans on time (and hopefully by this time the Myanmar credit bureau will be functioning and improving its coverage).

[13] See Credit Scorecards for SME Finance: The Process of Improving Risk Measurement and Management, April 2009 http://www.microfinance.com/English/Papers/Caire_Scorecards_Improving_Risk_Management_and_Measurement.pdf

Indeed, the main 'wins' to date for partners developing credit-scoring systems with support of the in the PSDA program have been in:

1. Designing streamlined processes for assessing borrowers in target segments;
2. Digitizing and systemizing the collection and storage of borrower data;
3. Building staff skills in credit risk management and basic data analytics.

Credit scoring models are only as good as the data that informs them. Building better credit scorecards involves:

- collecting more and better data;
- making loans to borrowers who become unwilling or unable to repay.

A Changing Landscape

Several concurrent developments in Myanmar are likely to help speed the potential utility of credit scoring:

1. The increased digitization of data collection and credit processes in financial institutions;
2. Wider use of mobile-banking services and digital-payment channels, including mobile wallets, point-of-sales (POS) data, and other non-cash payments;
3. The introduction of the Myanmar Credit Bureau;
4. Increased competition for clients among banks, MFIs and 'fintechs';
5. The recent liberalization of bank lending rates and collateral requirements for banks (as noted above) may encourage making more loans that are currently perceived as untested and risky.

However, any immediate short-cuts to better credit scoring in Myanmar are unlikely. Financial institutions are warned to treat with caution the promise of 'alternative data' credit-scoring vendors, until and unless they are able to convincingly prove a track record in Myanmar. Data, no matter how big, cannot tell a credit-scoring story until enough people fail in the timely repayment of their loans.

In conclusion, the PSDA experience recommends that financial institutions in Myanmar approach the market thoughtfully, understanding the need to accept some risk so as to be able to better measure risk and sustain retail lending at scale in the current regulatory environment. While local borrowers have historically been very diligent in loan repayment, new loan products for a new generation of borrowers may lead to the typically wider range of loan repayment behaviors found worldwide – and a greater role for credit scoring.

Acknowledgments: *Thanks to Mark Schreiner of Scorocs, L.L.C. and Mary Miller of Nathan Associates Inc. for editing this paper.*