Microcredit: stochastic and statistical approaches for understanding and scoring

MARC DIENER

Laboratoire J.A. Dieudonné

ISI-Delhi, october 2017

Abstract

Microcredit, as described by M. Yunus, allows efficient lending without collateral. One of its characteristics is that it is based on a large number of (frequent) settlements leading to little default risk but possible delay in settlements, creating randomness of actual interest rate. We take the example of the Yunus equation to examine the probability characteristics (law) of this interest rate, as a measure of the risks in microcredit. . . .

`http://math.unice.fr/~diener/Mifi/`

Markov chain models

- The simplest model
- Avoidind Strategic Default (ASD)
- Microcredit as a path towards inclusion into regular banking: Dhib Lessy model

Statistical approach of scoring

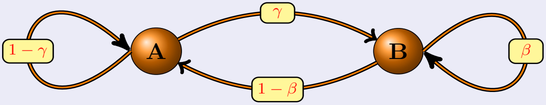
Variable Selection in Prediction of Repayment Outcome

Is microcredit efficient

References

The simplest model

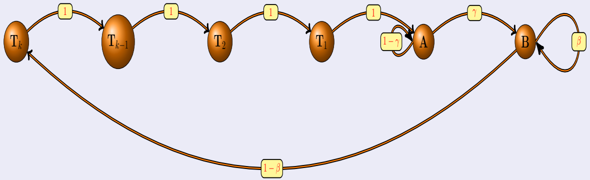
Figure 1: Only two states: Appliquant and Beneficiary



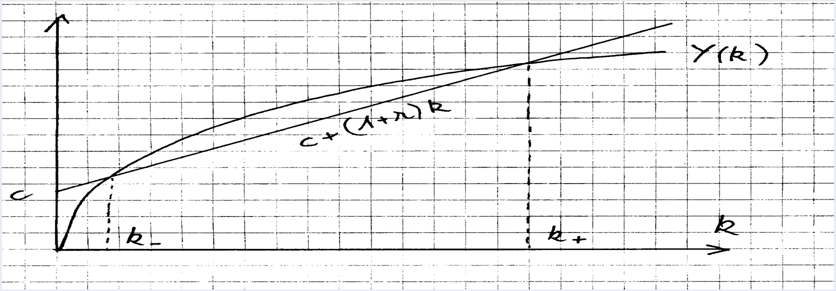
$$P = \begin{pmatrix} 1-\gamma & \gamma \\ 1-\beta & \beta \end{pmatrix}$$

k exclusion periods for discouraging strategic default

Figure 2: Tedeschi's idea: to avoid strategic default one can decide to exclude, for at least *k* time steps, a borrower who (pretends that she) is not able to pay



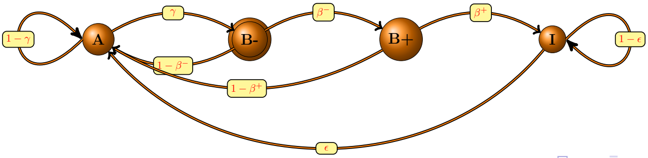
Yield and costs



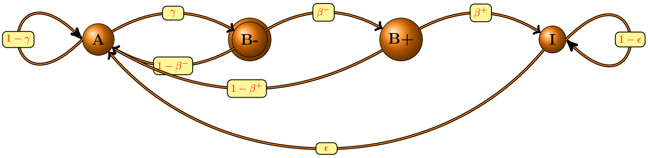
Microcredit as a path towards inclusion into regular banking

In this model we have a new state: I for “included (in regular banking)". We also have two beneficiary states: B^- and B^+ for “small loan" k_- and ”large loan" k_+.

- A (appliquant)
- B^- beneficiary of a small loan k_- with interest rate r)
- B^+ beneficiary of a large loan $k_=$ with same interest rate r)
- I (included in regular banking with same loan k^+, but with much smaller interest rate r')



Microcredit as a path towards inclusion into regular banking



$$P = \begin{pmatrix} 1-\gamma & \gamma & 0 & 0 \\ 1-\beta^- & 0 & \beta^- & 0 \\ 1-\beta^+ & 0 & 0 & \beta^+ \\ \varepsilon & 0 & 0 & 1-\varepsilon \end{pmatrix}$$

The 8 parameters of the model:

- stochastic: γ, β^-, β^+, ε,
- economic: α, r, r', δ.

Return and expected total discounted return

- $f(X_{t-1}, X_t) =$ return (benefit) for borrowing 1 when going from X_{t-1} to X_t.
- $\delta =$ discount factor
- any trajectory $(X_0, X_1, \ldots, X_t, \ldots)$, with $X_0 = x$, leads to the total discounted return
$$W(x) = \sum_{t \geq 1} \delta^t f(X_{t-1}, X_t)$$
-
- Theorem (Dhib/Diener)

$$w(x) := \mathbb{E}(W(x)) = \delta(I - \delta P)^{-1} Z(x) \text{ , where } Z(x) = \mathbb{E}(f(X_0, X_1) | X_0 = x)$$

Absence of Strategic Default (ASD) in Dhib-Lessy model

- $f(X_{t-1}, X_t) =$ return (benefit) for borrowing 1 when going from X_{t-1} to X_t: $f(x', x") = f_{II} = k_+(r - r')$ if $(x', x") = (I, I)$ and zero otherwise.
- $\delta =$ discount factor
- $X_0 = x$, leads to the total discounted return

$$W(x) = \sum_{t \geq 1} \delta^t f(X_{t-1}, X_t)$$

$$w(x) := \mathbb{E}(W(x)) = \delta(I - \delta P)^{-1} Z(x) \text{ , where } Z(x) = \mathbb{E}(f(X_0, X_1) | X_0 = x)$$

- ASD requires to compare $w(I)$ with $(1 + r)k_+ + w(A)$

Data and Variable Description

- BAAC group borrowers data collected and studied by Ahlin and Townsend [2007]
- The data set contains 219 samples with 25 variables.
- The 25 variables: REP, NOLNDPCT, COVARBTY, HOMOCCUP, SHARING, SHARNON, BCPCT, PRODCOOP, LIVEHERE, RELPRCNT, SCREEN, KNOWN, BIPCT, SNCTIONS, MEANLAND, AVGED, INTRAT, LOANSIZE, SQLOANSIZE, LNYRSOLD, MEMS, VARBTY, WEALTH, PCGMEM, CBANKMEM.

  REP= 0 if BAAC has ever raised interest rate as penalty for late repayment, and 1 otherwise.

Data and Variable Description

- BAAC group borrowers data collected and studied by Ahlin and Townsend [2007]
- The data set contains 219 samples with 25 variables.
- The 25 variables: REP, NOLNDPCT, COVARBTY, HOMOCCUP, SHARING, SHARNON, BCPCT, PRODCOOP, LIVEHERE, RELPRCNT, SCREEN, KNOWN, BIPCT, SNCTIONS, MEANLAND, AVGED, INTRAT, LOANSIZE, SQLOANSIZE, LNYRSOLD, MEMS, VARBTY, WEALTH, PCGMEM, CBANKMEM.

  REP= 0 if BAAC has ever raised interest rate as penalty for late repayment, and 1 otherwise.

Data and Variable Description

- BAAC group borrowers data collected and studied by Ahlin and Townsend [2007]
- The data set contains 219 samples with 25 variables.
- The 25 variables: REP, NOLNDPCT, COVARBTY, HOMOCCUP, SHARING, SHARNON, BCPCT, PRODCOOP, LIVEHERE, RELPRCNT, SCREEN, KNOWN, BIPCT, SNCTIONS, MEANLAND, AVGED, INTRAT, LOANSIZE, SQLOANSIZE, LNYRSOLD, MEMS, VARBTY, WEALTH, PCGMEM, CBANKMEM.

  REP= 0 if BAAC has ever raised interest rate as penalty for late repayment, and 1 otherwise.

Logistic Regression

- $Y = (Y_1, Y_2, \cdots, Y_n)'$ is an output vector , $Y_i \in \{0, 1\}$, $i = 1, 2, \cdots, n$.
- X is an $n \times (k+1)$ input matrix, $X_i = (x_{i0}, x_{i1}, \cdots, x_{ij}, \cdots, x_{ik})$.
- The logistic regression model:

$$\mathrm{logit}(\pi_i) = \ln\left(\frac{\pi_i}{1-\pi_i}\right) = X_i\beta + \varepsilon_i, \tag{1}$$

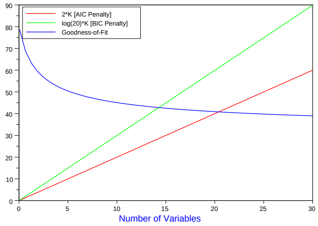
$$\pi_i = \pi(X_i) = \mathbb{P}(Y_i = 1 \mid X_i)$$

- Computation of $\hat{\beta}$:
  - Maximum likelihood method.
  - Iteratively.

Model Selection Criterion

- A parametric model: $f(y, \theta)$, $\theta \in \Theta$ and $\dim(\Theta) = k$.
- The Akaike Information Criterion: $\text{AIC} = -2 \ln f(y, \hat{\theta}) + 2k$.
- The Bayesian Information Criterion: $\text{BIC} = -2 \ln f(y, \hat{\theta}) + k \ln(n)$, n is the sample size.

Figure 4: Goodness-of-fit versus Penalty



- Optimal model is the one with minimum criterion.

AIC Optimal Model with AIC=223.56

AIC Optimal Model with AIC=223.56

▷ Apply AIC backward stepwise by the function `stepAIC()` in R-packages.

AIC Optimal Model with AIC=223.56

▷ Apply AIC backward stepwise by the function `stepAIC()` in R-packages.

Table 1: AIC Optimal Model

Variable	Coefficient	Std.Error	z-value.	Pr(>|z|)
Intercept	2.6706	0.6453	4.139	3.50e-05 ***
NOLNDPCT	-2.6429	1.0862	-2.433	0.01496 *
SHARING	0.3598	0.2135	1.686	0.09187 .
SHARNON	-0.4812	0.2283	-2.108	0.03506 *
PRODCOOP	0.5193	0.2498	2.079	0.03759 *
BIPCT	1.5597	1.0077	1.548	0.12168
SNCTIONS	3.4370	1.7222	1.996	0.04597 *
LNYRSOLD	-0.9022	0.2247	-4.015	5.94e-05 ***
PCGMEM	-3.2314	1.0596	-3.050	0.00229 **

Codes: ***, **, *, and . denote significance at 0%, 0.1%, 5%, and 10% respectively.

BIC Optimal Model with BIC=246.52 (AIC=226.19)

BIC Optimal Model with BIC=246.52 (AIC=226.19)

Table 2: BIC Optimal Model

Variable	Coefficient	Std.Error	z-value.	Pr(>|z|)
Intercept	2.8705	0.5789	4.958	7.11e-07 ***
NOLNDPCT	-2.9414	1.0751	-2.736	0.006222 **
PRODCOOP	0.5278	0.2427	2.175	0.029631 *
SNCTIONS	3.7675	1.6715	2.254	0.024201 *
LNYRSOLD	-0.8392	0.2175	-3.859	0.000114 ***
PCGMEM	-2.7119	0.9597	-2.826	0.004715 **

Codes: ***, **, *, and . denote significance at 0%, 0.1%, 5%, and 10% respectively.

BIC Optimal Model with BIC=246.52 (AIC=226.19)

Table 2: BIC Optimal Model

Variable	Coefficient	Std.Error	z-value.	Pr(>|z|)
Intercept	2.8705	0.5789	4.958	7.11e-07 ***
NOLNDPCT	-2.9414	1.0751	-2.736	0.006222 **
PRODCOOP	0.5278	0.2427	2.175	0.029631 *
SNCTIONS	3.7675	1.6715	2.254	0.024201 *
LNYRSOLD	-0.8392	0.2175	-3.859	0.000114 ***
PCGMEM	-2.7119	0.9597	-2.826	0.004715 **

Codes: ***, **, *, and . denote significance at 0%, 0.1%, 5%, and 10% respectively.

These 5 variables are also appeared in the AIC optimal model.

Subsequent Steps of Backward Stepwise

Subsequent Steps of Backward Stepwise

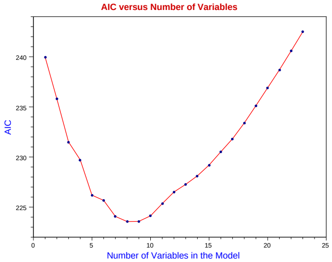
AIC	BIC	Var. Dropped
242.50	323.84	
240.58	318.53	- WEALTH
238.69	313.25	- MEANLAND
236.91	308.08	- HOMOCCUP
235.13	302.91	- KNOWN
233.39	297.78	- CBANKMEM
231.82	292.82	- BCPCT
230.52	288.14	- MEMS
229.2	283.43	- SCREEN
228.12	278.95	- LIVEHERE
227.27	274.71	- RELPRCNT
226.51	270.57	- VARBTY

AIC	BIC	Var. Dropped
225.35	266.02	- COVARBTY
224.13	261.41	- LOANSIZE
223.58	257.47	- AVGED
223.56	254.06	- INTRAT
224.07	251.18	- BIPCT
225.67	249.39	- SHARING
226.19	246.52	- SHARNON
229.71	246.66	- SNCTIONS
231.5	245.06	- PRODCOOP
235.81	245.98	- NOLNDPCT
239.97	246.74	- PCGMEM
255.19	258.58	- LNYRSOLD

AIC and BIC versus Number of variables in the Model

AIC and BIC versus Number of variables in the Model

Figure 5: AIC versus No. of Variables



AIC and BIC versus Number of variables in the Model

Figure 5: AIC versus No. of Variables

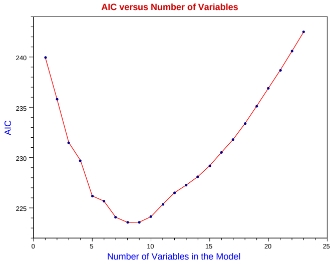
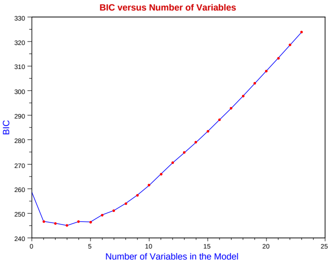


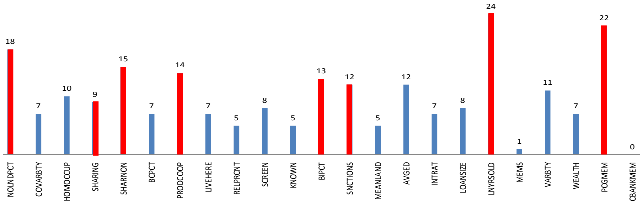
Figure 6: BIC versus No. of Variables



Validation of Optimal Models

Validation of Optimal Models

Figure 7: Frequency of Variables appeared in 25 AIC Optimal Models of Samplings



Validation of Optimal Models

Figure 7: Frequency of Variables appeared in 25 AIC Optimal Models of Samplings

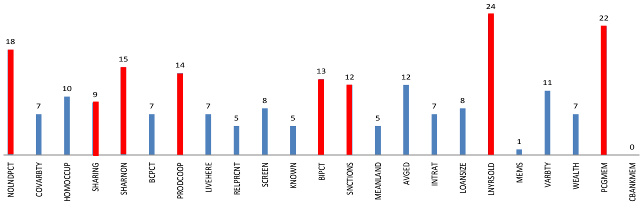
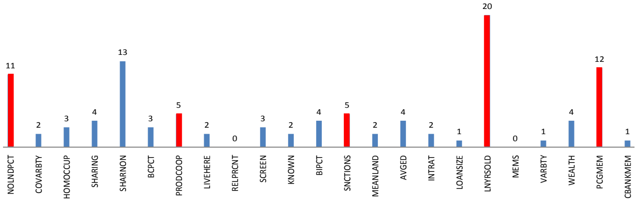


Figure 8: Frequency of Variables appeared in 25 BIC Optimal Models of Samplings



Adding INTRAT to the AIC Optimal Model, AIC=223.58

Adding INTRAT to the AIC Optimal Model, AIC=223.58

Table 3: The Final Model

Variable	Coefficient	Std.Error	z-value.	Pr(>|z|)
Intercept	4.168	1.282	3.252	0.00115 **
NOLNDPCT	-2.716	1.102	-2.464	0.01375 *
SHARING	0.384	0.221	1.736	0.08261.
SHARNON	-0.517	0.238	-2.176	0.02957 *
PRODCOOP	0.526	0.252	2.088	0.03682 *
BIPCT	1.473	1.013	1.453	0.14616
SNCTIONS	3.372	1.732	1.947	0.05155 .
INTRAT	-0.131	0.095	-1.373	0.16991
LNYRSOLD	-0.905	0.228	-3.970	7.17e-05 ***
PCGMEM	-3.250	1.062	-3.059	0.00222 **

Microcredit impact assessment

G. A. Tedeschi. Overcoming selection bias in microcredit impact assessments: a case study in Peru.
Journal of Development Studies, 44:4:504–518, 2008.

References

`http://math.unice.fr/~diener/Mifi/`

C. Ahlin and R. M. Townsend.
Using repayment data to test across models of joint liability lending.
The Economic Journal, 2007.

F. Diener, M. Diener, and N. Dhib.
Valeur espérée d'un microcrédit.
Technical report, Laboratoire Jean-Alexandre Dieudonné, 2015.

M. Diener, F. Diener, O. Khodr, and P. Protter.
Mathematical models for microlending.
In *16th Mathematical Conference of Bangladesh Mathematical Society, Dhaka, Bangladesh*, Dec 2009.

M. Diener and P. Mauk.
On the implicit interest rate in the Yunus equation.
In *Actes du colloque à la mémoire d'Emmanuel Isambert*, Paris, France, 2012.
4 pages (101-104).

O. Khodr.
Modèles dynamiques des innovations du microcrédit.
In *Thèse de Doctorat, EDSFA, Laboratoire J-A Dieudonné, UNS, Parc Valrose, 06108 Nice, France*, pages 1–61, 2011.

P. Mauk.
Modélisation Mathématique du Microcrédit.
PhD thesis, EDSFA, UNS, 2013.

G. Tedeschi.
Here today, gone tomorrow: Can dynamic incentives make microfinance more flexible ?
Journal of Development Economics, 80(1):84 – 105, 2006.

G. A. Tedeschi.
Overcoming selection bias in microcredit impact assessments: a case study in peru.
Journal of Development Economics, 44:4:504–518, 2008.

M. Yunus with Alan Jolis.
Banker to the Poor : micro-lending and the battle against world poverty.
Public Affairs, 1999.

Thanks

Thank you for your attention :-)

http://math.unice.fr/~diener/Mifi/