MONTE DEI PASCHI DI SIENA

BANCA DAL 1472

How Reject Inference Can Improve the Credit Granting Process

Credit Scoring and Credit Control XIV

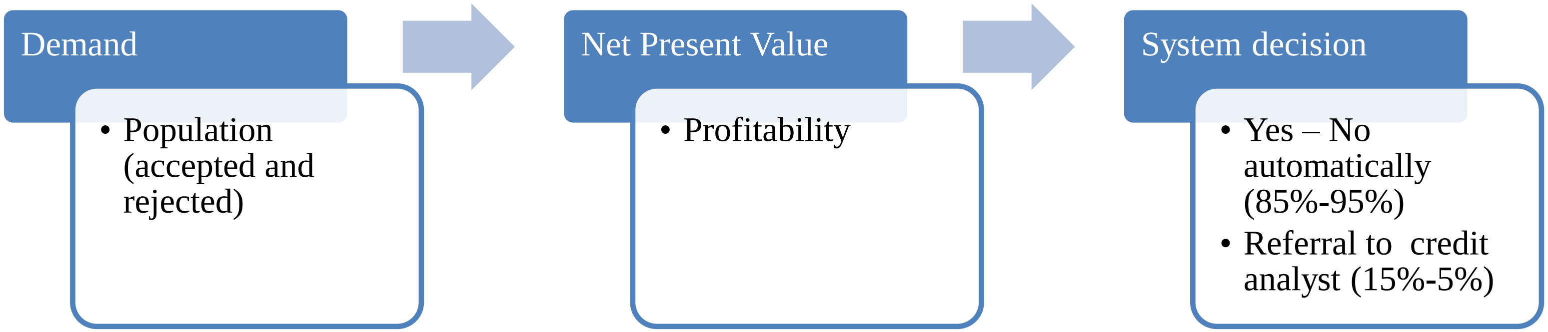
Edinburgh August 26-28, 2015

Lorenzo Quirini
Banca Monte dei Paschi di Siena
Area Servizi Specialistici del Credito-Servizio Credit Services

Claudio Carpenteri
Experian

Fabrizio Cipollini
University of Florence

CREDIT GRANTING DECISION PROCESS IS BASED ON THE PROFITABILTY OF THE ENTIRE PORTFOLIO



- A profitability index should be the core of an automated decision system
- Net Present Value (NPV) summarizes the customer’s behaviour and the bank‘s cost of capital
- NPV’s expectation can be evaluated by an index which condenses the way a customer repays her obligations (CWI – CreditWorthiness Index)
- CWI has to represent the entire population
- Reject Inference is requested and an internal model has been developed
- External credit bureau improves Reject Inference; Experian collaborated to the project

CONTENTS

- Creditworthiness Index (CWI) and Net Present Value

- Model formulation to estimate CWI including rejected loans

- Role of external information and integration in credit decision process

CWI COMPARES ACTUAL CASH FLOWS WITH CONTRACTUAL ONES

$$X_t = \frac{\sum_{h=1}^{t} R_h (1+i)^{-h}}{\sum_{h=1}^{t} r_h (1+i)^{-h}}$$

R_h random installment at time h

$h = 1, \ldots, n$ with n term of the operation

t evaluation time with $t = 1, \ldots, n$

r_h contractual installment at time h

i contractual rate of return

Quirini L., Vannucci L., (2010), "A new index of creditworthiness for retail products", in *Journal of the Operational Research Society,* 61, 455-461.

SOME INTERPRETATIONS CAN BE GIVEN TO DESCRIBE THE CUSTOMER'S BEHAVIOUR

Case 1

Deterministic percentage of each installment

$$R_h = a r_h \quad (a \le 1)$$

$$X_t = \frac{\sum_{h=1}^{t} R_h (1+i)^{-h}}{\sum_{h=1}^{t} r_h (1+i)^{-h}} = \ldots = a$$

Case 2

Systematic delayed repayment

$$X_{n+j} = \frac{\sum_{h=1}^{n} r_h (1+i)^{-(h+j)}}{\sum_{h=1}^{n} r_h (1+i)^{-h}} = (1+i)^{-j}$$

Case3

Lottery

$$E(X_t) = E\left(\frac{\sum_{h=1}^{t} R_h (1+i)^{-h}}{\sum_{h=1}^{t} r_h (1+i)^{-h}}\right) = \ldots = p$$

Case 4

Relationship with prob default and recovery

$$E(X_n) = 1 \cdot (1 - p_d) + p_d \cdot \frac{\sum_{h=1}^{t_d} r_h \cdot (1+i)^{-h} + z_d \cdot (1+i)^{-n}}{b}$$

$$\sigma^2(X_n) = 1 \cdot (1 - p_d) + p_d \cdot \left(\frac{\sum_{h=1}^{t_d} r_h \cdot (1+i)^{-h} + Z_d \cdot (1+i)^{-n}}{b}\right)^2 - E^2(X_n)$$

p_d Probabilty of default at time t_d

Z_d cash flow for the defaulted loan at time t_d

CWI MAY BE LINKED TO PROFITABILITY EXPRESSED IN TERM OF NPV

$$NPV(s,a) = -b + \sum_{h=1}^{n} a \cdot r_h (1+s)^{-h}$$

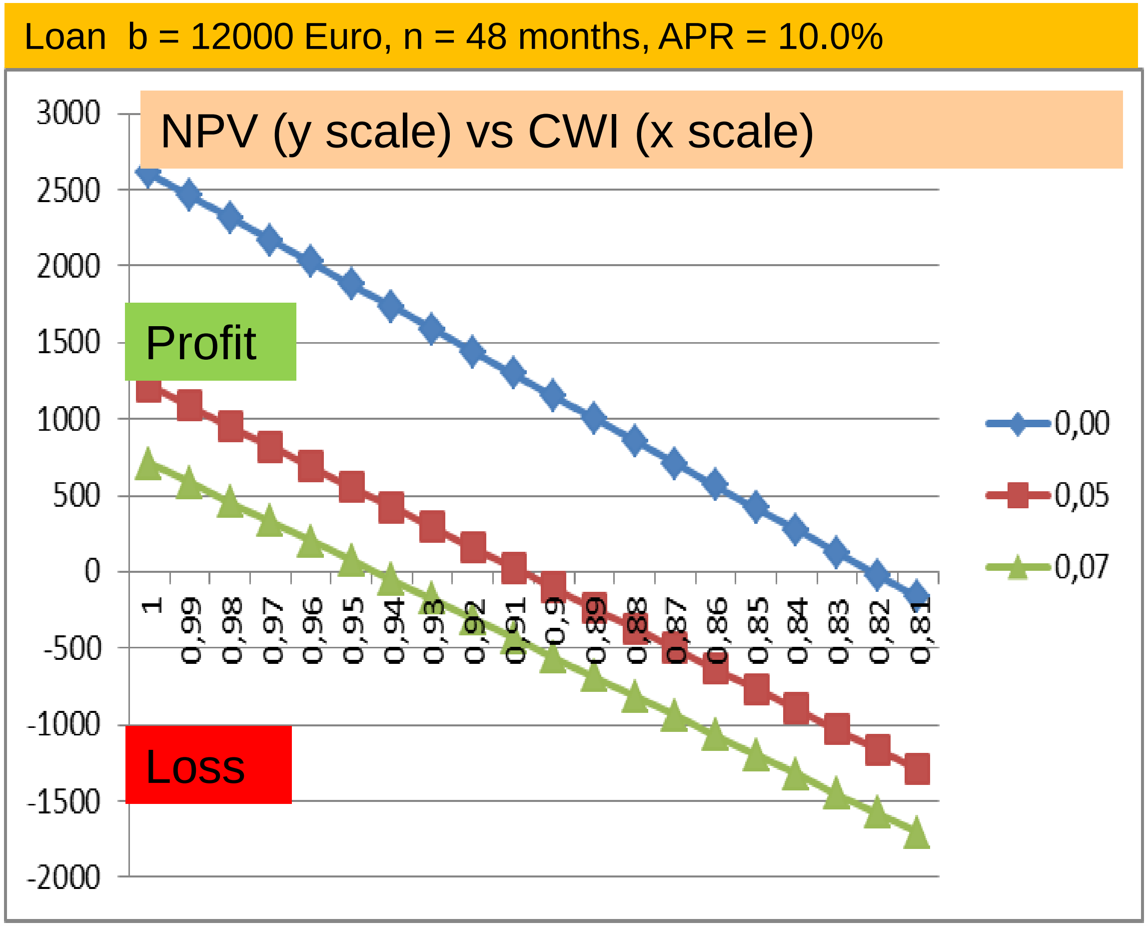
b granted amount

n term of the operation

r_h installment due at time *h*

s is the discount rate

a CWI (0≤ *a* ≤1)- first CWI interpretation



CONTENTS

- Creditworthiness Index (CWI) and Net Present Value

- Model formulation to estimate CWI including rejected loans

- Role of external information and integration in credit decision process

PROCESS WITH THREE STEPS: MODEL DEVELOPMENT, INFORMATION FROM A CREDIT BUREAU, INTEGRATION INTO BANK'S DECISION SYSTEM

1) CWI model taking into account the entire population (rejected and approved loans)

Multivariate Heckman Type Formulation

Relationship with credit decision (Yes/No)

Relationship with expected CWI

2) Comparison with the CWI's expected values given by Experian

3) The application in automated decision rules

Motivations: The Creditworthiness Index (CWI)

- Quirini & Vannucci (2010)
- Measure of the debtor's repayment quality increasing continuously in [0,1]
- Finer information than more usual default {0,1} - measures
- Data have many ties at extreme values, in particular at the unit value

Modeling strategy

- Reject Inference (RI) framework
- Heckman-type formulation for modeling the dependence of CWI from explanatory variables according to the CWI characteristics

Contributions

- Parameter interpretation
- Maximum Likelihood inference (no simulation based methods needed)
- Formulas for the conditional expectation of the outcome (fitted values, predictions)
- Goodness-of-fit diagnostics

MODEL FORMULATION

$$\left.\begin{matrix} S^* \\ Y^* \end{matrix} \right| x \sim N\left[\begin{pmatrix} \mu_S = x'_S\beta_S \\ \mu_Y = x'_Y\beta_Y \end{pmatrix}, \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix} \right]$$

$$S = \mathrm{I}(S^* > 0)$$

$$Y = Y^*\,\mathrm{I}(0 < Y^* < 1) + \mathrm{I}(1 \leq Y^*)$$

Main points

- Reject Inference (RI) framework
- *Selection* (*S*) and *outcome* (*Y*) quantities driven by latent variables (*S** and *Y**)
- *S** and *Y** structured as a bivariate linear regression model with Normally distributed, correlated errors
- Standard deviation of *S** fixed at 1 for identification
- Mapping between the latent and the observables from convenient transformations
- x_S, x_Y independent variables
- β_S, β_Y, ρ, σ parameters to be estimated

DETAILS ON MODEL INTERPRETATION

- Unifying interpretation of ρ for different Heckman-type formulations
- In general, if $g(.)$ is any function and

$$E\left[g\left(Y^*\right)|S^*>0\right]=\int g\left(\mu_Y+\sigma z\right)\phi(z)\underbrace{\frac{\Phi\left(\left(\mu_S+\rho z\right)/\left(1-\rho^2\right)^{1/2}\right)}{\Phi(\mu_S)}}dz$$

Accordingly, for two contracts having **the same** μ_Y, then μ_S influences this expectation only via the **ratio** under the brace.

- This ratio behaves as follows as function of z:

For $\rho < 0$ it decreases monotonically (more and more flat increasing μ_S)
For $\rho > 0$ it increases monotonically (more and more flat increasing μ_S)

- Ultimate **implication**: for a **fixed** μ_Y, the relation between above expectation and μ_S is **direct** for $\rho < 0$ and **inverse** for $\rho > 0$

DETAILS ON MODEL INFERENCE

- Inference based on *Maximum Likelihood* (ML)
- *Log-likelihood* and *score* functions computed analytically
- Sandwich variance-covariance matrix *robust* to some misspecification
- *Fitted* (in-sample) values

$$E(Y|S=1, x_S, x_Y)$$

and *predicted* (out-of sample) values

$$E(Y|x_Y)$$

have closed form expressions

- Goodness-of-fit diagnostics specific for this model are proposed: a pseudo - R^2 and a *Hosmer-Lemeshow* type statistic

EXPECTED NPV CAN BE EVALUATED AT LOAN LEVEL AND FOR A PORTFOLIO

Input:

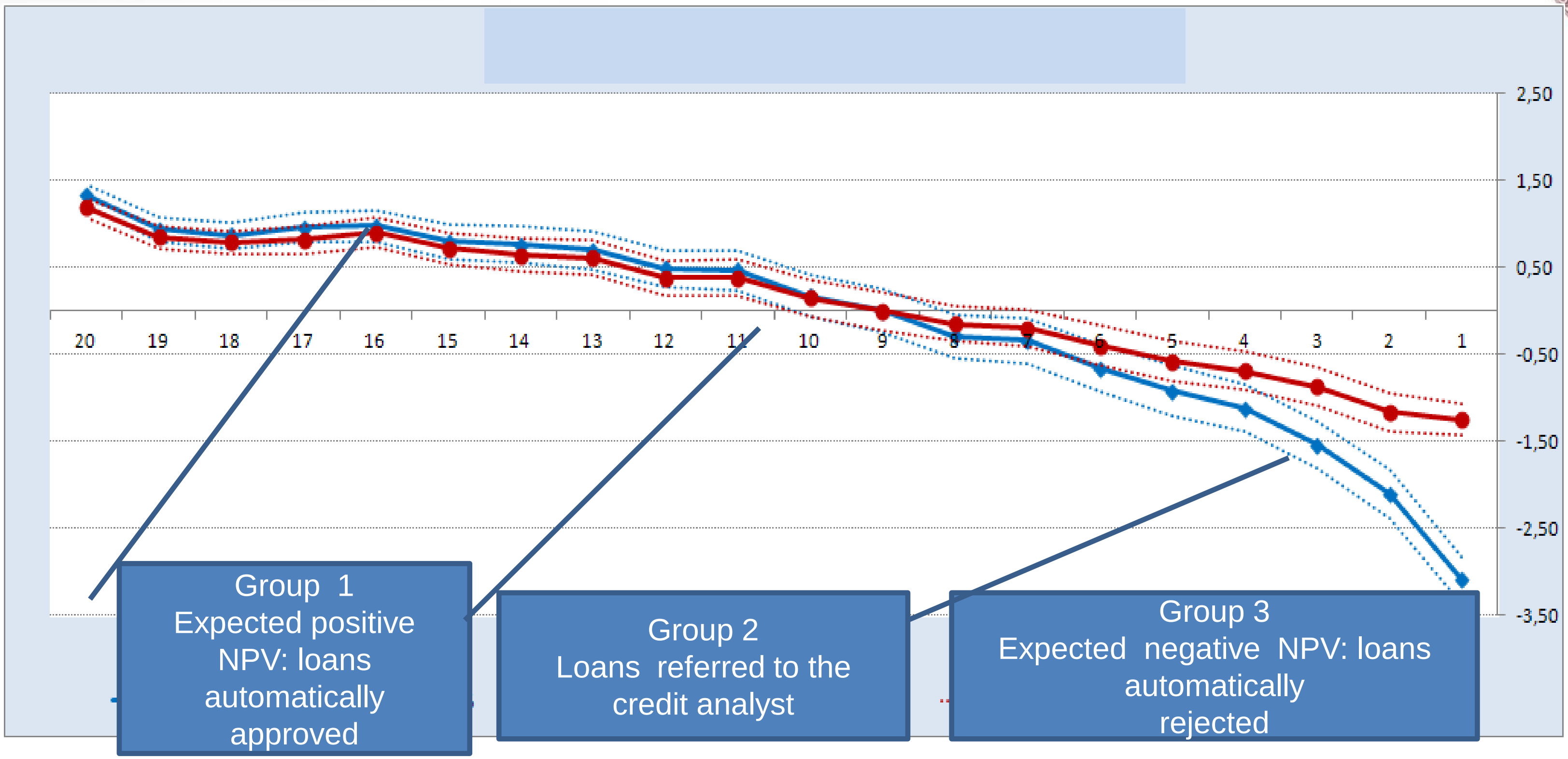
- Contractual elements: granted amount, term, installments
- Credit risk parameters: expected CWI over the term of the loan
- Cost of capital: constant

Output:

- Conctractual NPV
- Random NPV at loan level: expected value
- Random NPV for portfolio of similar risky loans : expected value

More details in Quirini, Vannucci, Cipollini (2013):
«Default and prepayment: an NPV analysis under a Markovian dynamics of the credit market» Credit Scoring and Credit Control XIII

POPULATION CAN BE DIVIDED INTO THREE SEGMENTS



$E(Y|x_Y)$ **All demand**

$E(Y|S=1, x_S, x_Y)$ **Granted loans**

CONTENTS

- Creditworthiness Index (CWI) and Net Present Value

- Model formulation to estimate CWI including rejected loans

- Role of external information and integration in credit decision process

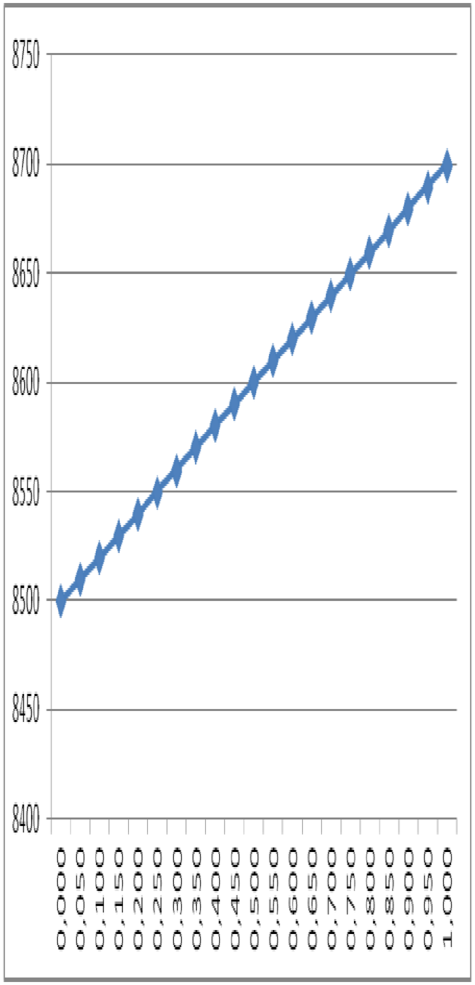
STRATEGY TO MENAGE THE UNCERTAINTY OVER THE RISK PROFILE

- Any loan can be seen as a mixture of two or more probabilities (see interpretation of CWI as a lottery)

- For example CWI equal to 8600 bps can be seen as a mixture (with equal weights) of a CWI equal to 8500 bps (scenario A) or a CWI equal to 8700 bps (scenario B)

- Internal information don’t help to reduce uncertainty on these two scenarios

- Sample made by Experian reduces such uncertainty



EXAMPLE FOR THE GROUP OF LOANS REFERRED TO CREDIT ANALYST

- Granted amount: 1.000 Euro, term: 12 months, installment: 100 Euro
- Parameters over the term of the loan : expected CWI 8500 *bps* (scenario A); expected CWI 8700 bps (scenario B)*;* both cases have the same probability
- CWI seen as a probability (lottery)
- Discount rate : 500 bps year base

- Sample given by Experian: 1000 loans (accepted and rejected) with an average CWI average equal to 8400 bps
- The sample has modified the uncertainty between the two scenarios: the probability of A increases from 50% to 97%, the probability for B drops from 50% to 3%

HOW THE BUREAU INFORMATION CAN BE APPLIED IN DECISION MAKING

	Internal Model **Pr(A) = 50%** **Pr(B) = 50%**	**Internal model + external information (Experian)** **Pr(A) = 97%,** **Pr(B) = 3%**
Contractual NPV	168 €	168 €
Expected NPV (mixture of scenarios)	5 €	-6 €
Decision	Referral to credit analyst	Reject

MONTE DEI PASCHI DI SIENA
BANCA DAL 1472

www.mps.it