Modelling the Impact of Collection Actions on Recovery Rates Using Retail Loan Level Panel Data

Angela Rita Freitas De Moraes

Dr. Galina Andreeva

Prof. Jonathan Crook

Credit Scoring & Credit Control 2017

INTRODUCTION

- Credit risk assessment plays an important role in the credit risk decisions of financial institutions.

- Basel II: Banks are permitted to develop and use their own internal risk ratings. The IRB approach is based on four key parameters used to estimate credit risk:

  1. PD - The probability of default of a borrower over one-year horizon;

  2. LGD – The Loss Given Default (or 1 minus recovery) as a percentage of exposure at default;

  3. EAD – Exposure at Default (an amount, not a percentage);

MOTIVATION

- Considerable increase in delinquency may cause lenders with low capital adequacy ratios to become insolvent. For this reason, decisions need to be taken on how to manage delinquency so that the likelihood of the account recovering is maximized and potential future losses due to write off are minimized (Crook and Banasik, 2012).
- It is known that recovery depends on the debt collection process, and the different options or actions that collection departments can take.
- There is practically no literature that explores the impact of the lender's collection actions on RR/LGD. For that reason, this work investigates the role of different collection actions at the loan level for a retail credit product, and estimates LGD models using Panel Data which is built by tracking customer payments following collection actions.

BACKGROUND

Credit Risk Management:

- ✓ Credit Risk Research – Focused on the estimation and validation of PD parameters (Loterman et al., 2012)

- ✓ LGD - Crucial input to Basel II regulatory capital calculation. Industry models for LGD , particularly for consumer lending portfolios, are often built using Ordinary Lest Squares regression or regression trees (Bastos, 2009; Bellotti & Crook, 2007; Caselli & Querci, 2009; Gupton & Stein, 2002)

- ✓ This illustrates one of the important issues in LGD modelling, namely that LGD depends not only on the uncertainty of whether a defaulter will repay, or how much they can afford to repay, but also on the lender's collection policy

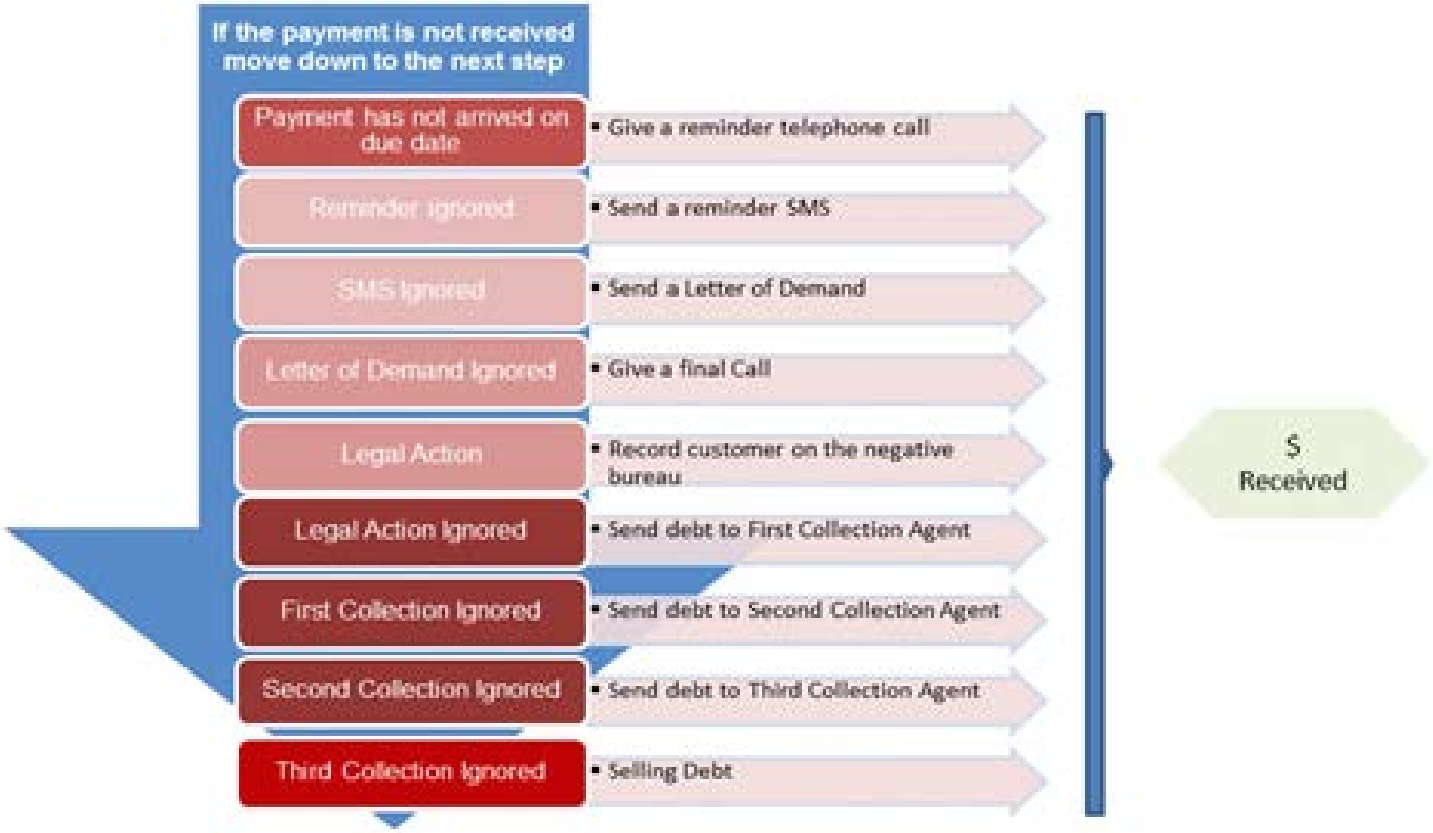
BACKGROUND

Collection Policy:

- ✓ Collection process steps
  - ➢ In house
  - ➢ Outside agents
  - ➢ Selling debt

BACKGROUND

Collection Policy:



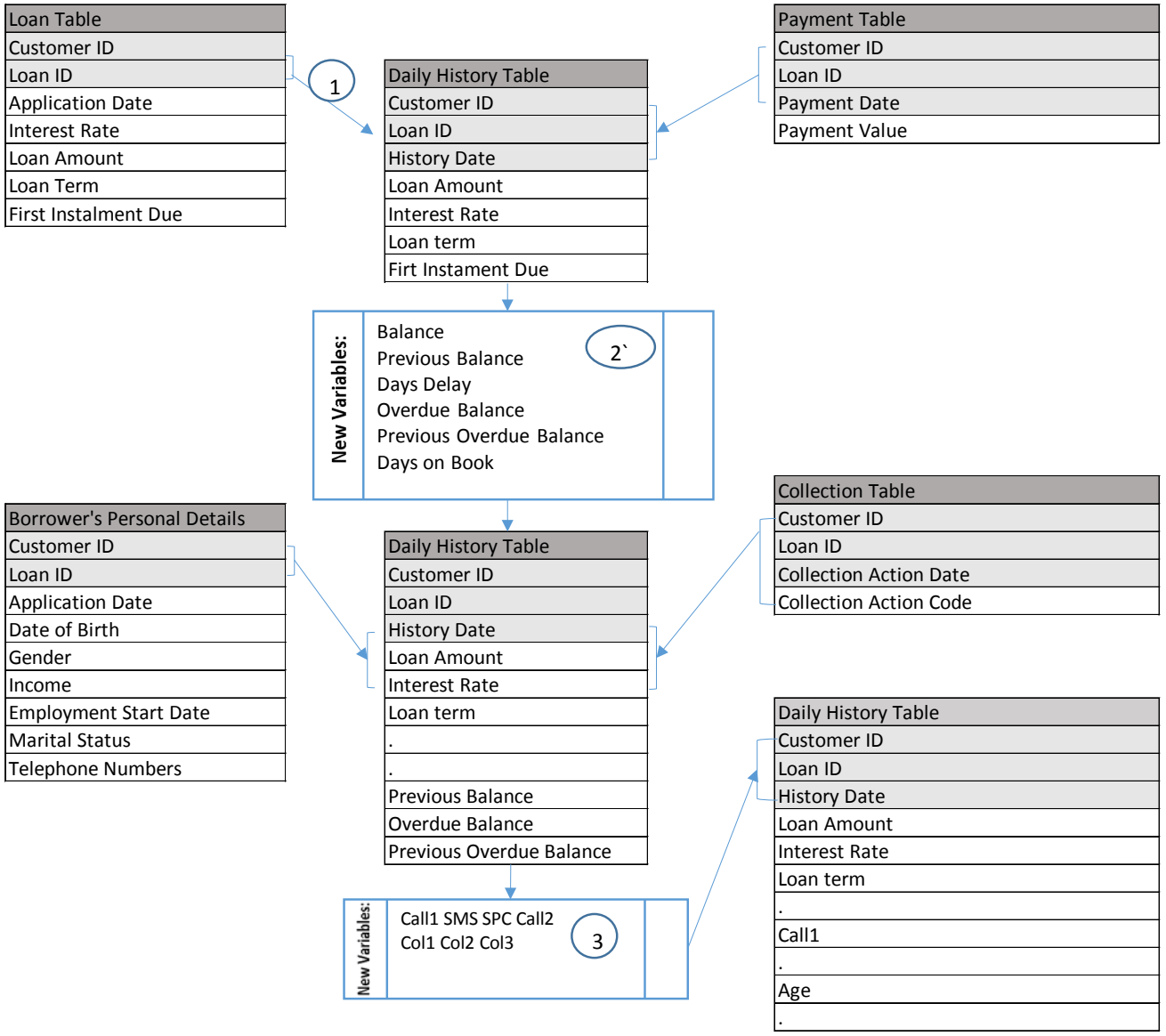
MODEL

DATA COLLECTION

Data Set

- Sample size →345,000 accounts from loan products of Brazilian lender
- Observation period → from January 2010 to August 2016
  - Personal customer details
  - Loan financial information
  - Collection information
- Training set → 2011 loan accounts
- Validation set → 2012 loan accounts

DATA DESIGN


Loan Table
Customer ID
Loan ID
Application Date
Interest Rate
Loan Amount
Loan Term
First Instalment Due
1
Daily History Table
Customer ID
Loan ID
History Date
Loan Amount
Interest Rate
Loan term
Firt Instament Due
Payment Table
Customer ID
Loan ID
Payment Date
Payment Value
New Variables:
Balance
Previous Balance
Days Delay
Overdue Balance
Previous Overdue Balance
Days on Book
2`
Borrower's Personal Details
Customer ID
Loan ID
Application Date
Date of Birth
Gender
Income
Employment Start Date
Marital Status
Telephone Numbers
Daily History Table
Customer ID
Loan ID
History Date
Loan Amount
Interest Rate
Loan term
.
.
Previous Balance
Overdue Balance
Previous Overdue Balance
Collection Table
Customer ID
Loan ID
Collection Action Date
Collection Action Code
New Variables:
Call1 SMS SPC Call2
Col1 Col2 Col3
3
Daily History Table
Customer ID
Loan ID
History Date
Loan Amount
Interest Rate
Loan term
.
Call1
.
Age
.

DATA DESIGN

Change Daily into Weekly

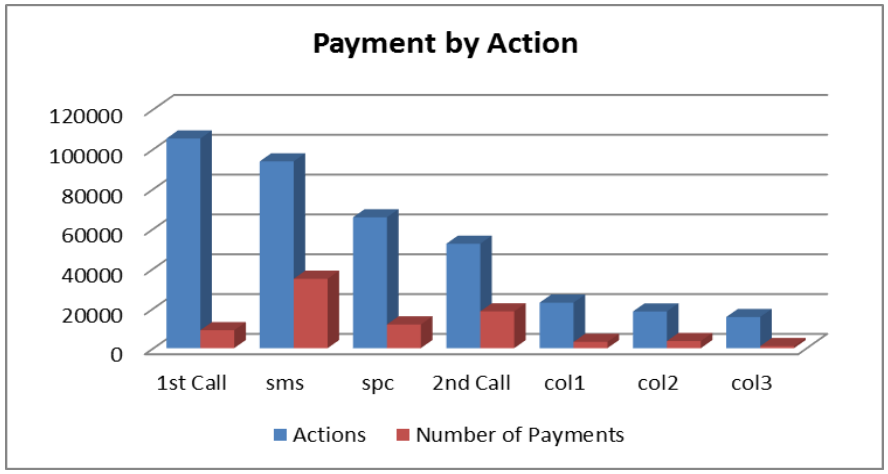
Daily History Table
Customer ID
Loan ID
History Date
Loan Amount
Interest Rate
Loan Term
First Instalment Due
Balance
Previous Balance
Overdue Balance
Previous Overdue Balance
Days Delay
Overdue Balance
Previous Overdue Balance
Days on Book
Account Clouse Date
Week
Call1
SMS
SPC
Call2
Col1
Col2
Col3
Application Date
Date of Birth
Gender
Income
Employment Start Date
Marital Status
Land Line Number
Mobile Number

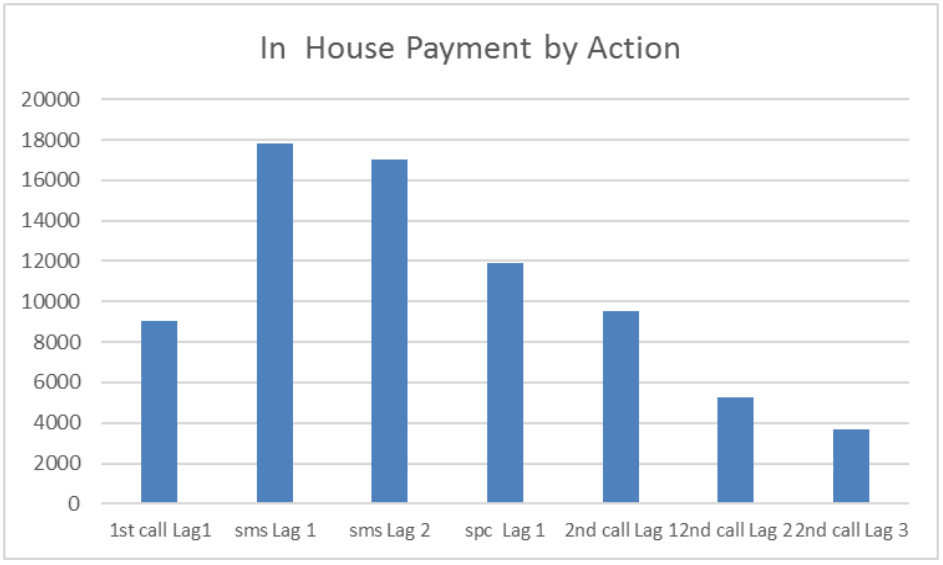
Collection Actions are Lagged

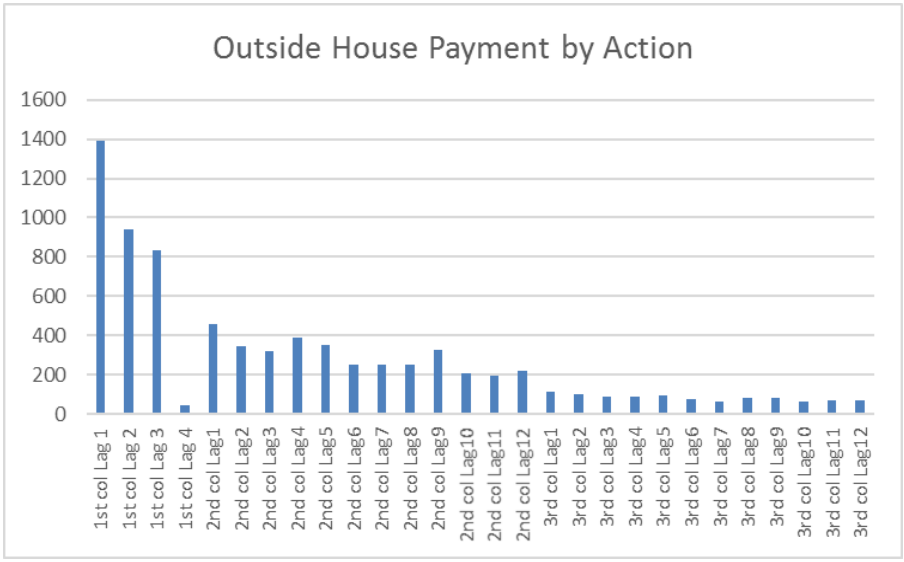
VARIABLES

Cross Sectional	Time Series
Age	First Call at 12 DPD
Gender	SMS at 15 DPD
Income	SPC at 25 DPD
Marital Status	Second Call at 30 DPD
Employment Length	First Collect Agency at 60 DPD
Product Category	Second Collect Agency at 90 DPD
Ratio of Instalment/Income	Third Collect Agency at 180 DPD
Ratio of Loan Amount/Income	

2011 ARREARS POPULATION







CHARACTERISTICS OF THE SAMPLE

- Cross section
  - Variation across individuals modelled
  - Unit of observation $i = 1, \dots, N$
  - Comparability / heterogeneity

- Panel
  - Variation across individuals *and* time modelled
  - Units of observation $i = 1, \dots, N; t = 1, \dots T$
  - Up to observations

- Unbalanced
  - missing observations

STATISTICAL ANALYSIS

The impact of the lender's collection actions on RR/LGD

$$Y1_{it} = \alpha_i + \beta_1 D_{it}^{Call1} + \beta_2 D_{it}^{SMS} + \beta_3 D_{it}^{SPC} + \beta_4 D_{it}^{Call2} + \beta_5 D_{it}^{Col1} + \beta_6 D_{it}^{Col2} + \beta_7 D_{it}^{Col3} + \sum_{N=1}^{N} \delta_N X_{Ni} + \lambda Z_i + \varepsilon_{i,t}$$

Collection Actions over time | Customer Details | Bal/L Amt | Error

$$Y2_{it} = \alpha_i + \beta_1 D_{it}^{Call1} + \beta_2 D_{it}^{SMS} + \beta_3 D_{it}^{SPC} + \beta_4 D_{it}^{Call2} + \beta_5 D_{it}^{Col1} + \beta_6 D_{it}^{Col2} + \beta_7 D_{it}^{Col3} + \sum_{N=1}^{N} \delta_N X_{Ni} + \varepsilon_{i,t}$$

Dependent Variables:

Overdue Balance $\rightarrow Y1$ (Sum of the values of instalments overdue)

- Model 1: outstanding balance as an exploratory variable
- Model 2: loan amount as an exploratory variable

Rec Ratio $\rightarrow Y2$ (Overdue balance divided by loan amount)

- Model 3: Neither outstanding balances nor loan amounts were used

Bal Ratio $\rightarrow Y2$ (Overdue balance divided by outstanding balance)

- Model 4: Neither outstanding balances nor loan amounts were used

RESULTS

Exploratory Variables	Model 1		Model 2		Model 3		Model 4	
	Overdue Balance		Overdue Balance		Rec Ratio		Bal Ratio	
	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value
Lag1 call1	-256.159	0.000	-359.5600	0.000	-0.2353	0.000	-0.1855	0.000
Lag1 sms	-260.97	0.000	-367.5741	0.000	-0.2379	0.000	-0.1857	0.000
Lag1 spc	-57.5381	0.000	-108.8432	0.000	-0.056049	0.000	0.0133532	0.000
Lag1 call2	-197.38	0.000	-316.1666	0.000	-0.1788	0.000	-0.0301	0.000
Lag1 col1	-159.551	0.000	-209.9550	0.000	-0.0934	0.000	0.0913	0.000
Lag1 col2	-105.43	0.000	-121.4622	0.000	-0.0282	0.000	0.1563	0.000
Lag1 col3	56.05598	0.000	79.3606	0.000	0.1129	0.000	0.3025	0.000
Lag2 sms	-185.821	0.000	-285.1386	0.000	-0.1722	0.000	-0.0925	0.000
Lag2 call2	-179.288	0.000	-291.5136	0.000	-0.1572	0.000	0.0092	0.000
Lag2 col1	-164.627	0.000	-212.8397	0.000	-0.0943	0.000	0.0913	0.000
Lag2 col2	-110.406	0.000	-124.8195	0.000	-0.0299	0.000	0.1540	0.000
Lag2 col3	52.48934	0.000	77.7802	0.000	0.1111	0.000	0.2983	0.000
Lag3 call2	-136.942	0.000	-229.5463	0.000	-0.1131	0.000	0.0441	0.000
Lag3 col1	-165.675	0.000	-211.5525	0.000	-0.0929	0.000	0.0866	0.000
Lag3 col2	-113.749	0.000	-127.2828	0.000	-0.0314	0.000	0.1489	0.000

RESULTS

	Overdue Balance		Rec_ratio	Bal_ratio
Model	**Model 1**	**Model 2**	**Model 3**	**Model 4**
P-Value	0.000	0.000	0.000	0.000
R-squared	0.7621	0.5060	0.2709	0.0238
Adj R-squared	0.7621	0.5060	0.2709	0.0237
Root MSE	415.81	599.18	0.0303	1.9663
MAE	297.5692	474.9476	0.2667	0.2921
MSE	466.8474	650.0159	0.3142	0.3100

CONCLUSION

- Panel modelling is a viable approach in modelling debt recovery
- The Model 1
  - Dependent variable: overdue balance
  - Exploratory variables: addition of balance
    - Higher R-squared 76%
    - Lower MAE and MSE
- Each collection actions has a significant impact on debt recovery

NEXT STEPS

- Incorporating macroeconomic variables into the model
  - GDP
  - Unemployment Rates
  - Consumer Price Index
  - Consumer Confidence Index