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
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);
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.
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 Least 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:

If the payment is not received move down to the next step:

- Payment has not arrived on due date
  - Give a reminder telephone call
  - Send a reminder SMS
- Reminder ignored
  - Send a Letter of Demand
- SMS Ignored
  - Give a final Call
- Letter of Demand Ignored
  - Record customer on the negative bureau
- Legal Action
  - Send debt to First Collection Agent
- Legal Action Ignored
  - Send debt to Second Collection Agent
- First Collection Ignored
  - Send debt to Third Collection Agent
- Second Collection Ignored
  - Selling Debt
- Third Collection Ignored
  - $ Received
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

Payment Table
Customer ID
Loan ID
Payment Date
Payment Value

Daily History Table
Customer ID
Loan ID
History Date
Loan Amount
Interest Rate
Loan Term
First Instalment Due

New Variables:
Balance
Previous Balance
Days Delay
Overdue Balance
Previous Overdue Balance
Days on Book

Borrower's Personal Details
Customer ID
Loan ID
Application Date
Date of Birth
Gender
Income
Employment Start Date
Marital Status
Telephone Numbers

Collection Table
Customer ID
Loan ID
Collection Action Date
Collection Action Code

New Variables:
Balance
Previous Balance
Days Delay
Overdue Balance
Previous Overdue Balance

Call1 SMS SPC Call2
Col1 Col2 Col3

New Variables:
Call1
Call2
Col1
Col2
Col3
Age
### DATA DESIGN

#### 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 Close 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

---

**Change Daily into Weekly**

---

**Collection Actions are Lagged**
# VARIABLES

<table>
<thead>
<tr>
<th>Cross Sectional</th>
<th>Time Series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>First Call at 12 DPD</td>
</tr>
<tr>
<td>Gender</td>
<td>SMS at 15 DPD</td>
</tr>
<tr>
<td>Income</td>
<td>SPC at 25 DPD</td>
</tr>
<tr>
<td>Marital Status</td>
<td>Second Call at 30 DPD</td>
</tr>
<tr>
<td>Employment Length</td>
<td>First Collect Agency at 60 DPD</td>
</tr>
<tr>
<td>Product Category</td>
<td>Second Collect Agency at 90 DPD</td>
</tr>
<tr>
<td>Ratio of Instalment/Income</td>
<td>Third Collect Agency at 180 DPD</td>
</tr>
<tr>
<td>Ratio of Loan Amount/Income</td>
<td></td>
</tr>
</tbody>
</table>
2011 ARREARS POPULATION

Payment by Action

In House Payment by Action

Outside House Payment by Action
CHARACTERISTICS OF THE SAMPLE

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

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

- Unbalanced
  - missing observations
STATISTICAL ANALYSIS

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

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

\[
Y_{2it} = \alpha_i + \beta_1 D_{Call1}^{it} + \beta_2 D_{SMS}^{it} + \beta_3 D_{SPC}^{it} + \beta_4 D_{Call2}^{it} + \beta_5 D_{Col1}^{it} + \beta_6 D_{Col2}^{it} + \beta_7 D_{Col3}^{it} + \sum_{N=1}^{N} \delta_N X_{Ni} + \epsilon_{i,t}
\]

Collection Actions over time
Customer Details
Bal/L Amt
Error

Dependent Variables:
- Overdue Balance → \( Y_1 \) (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 → \( Y_2 \) (Overdue balance divided by loan amount)
  - Model 3: Neither outstanding balances nor loan amounts were used
- Bal Ratio → \( Y_2 \) (Overdue balance divided by outstanding balance)
  - Model 4: Neither outstanding balances nor loan amounts were used
## RESULTS

<table>
<thead>
<tr>
<th>Exploratory Variables</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overdue Balance</td>
<td>Overdue Balance</td>
<td>Rec Ratio</td>
<td>Bal Ratio</td>
</tr>
<tr>
<td>Coef.</td>
<td>P-value</td>
<td>Coef.</td>
<td>P-value</td>
<td>Coef.</td>
</tr>
<tr>
<td>Lag1 call1</td>
<td>-256.159</td>
<td>0.000</td>
<td>-359.5600</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag1 sms</td>
<td>-260.97</td>
<td>0.000</td>
<td>-367.5741</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag1 spc</td>
<td>-57.5381</td>
<td>0.000</td>
<td>-108.8432</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag1 call2</td>
<td>-197.38</td>
<td>0.000</td>
<td>-316.1666</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag1 col1</td>
<td>-159.551</td>
<td>0.000</td>
<td>-209.9550</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag1 col2</td>
<td>-105.43</td>
<td>0.000</td>
<td>-121.4622</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag1 col3</td>
<td>56.05598</td>
<td>0.000</td>
<td>79.3606</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag2 sms</td>
<td>-185.821</td>
<td>0.000</td>
<td>-285.1386</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag2 call2</td>
<td>-179.288</td>
<td>0.000</td>
<td>-291.5136</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag2 col1</td>
<td>-164.627</td>
<td>0.000</td>
<td>-212.8397</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag2 col2</td>
<td>-110.406</td>
<td>0.000</td>
<td>-124.8195</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag2 col3</td>
<td>52.48934</td>
<td>0.000</td>
<td>77.7802</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag3 call2</td>
<td>-136.942</td>
<td>0.000</td>
<td>-229.5463</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag3 col1</td>
<td>-165.675</td>
<td>0.000</td>
<td>-211.5525</td>
<td>0.000</td>
</tr>
<tr>
<td>Lag3 col2</td>
<td>-113.749</td>
<td>0.000</td>
<td>-127.2828</td>
<td>0.000</td>
</tr>
</tbody>
</table>
## RESULTS

<table>
<thead>
<tr>
<th>Model</th>
<th>Overdue Balance</th>
<th>Rec_ratio</th>
<th>Bal_ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.7621</td>
<td>0.5060</td>
<td>0.2709</td>
</tr>
<tr>
<td>Adj R-squared</td>
<td>0.7621</td>
<td>0.5060</td>
<td>0.2709</td>
</tr>
<tr>
<td>Root MSE</td>
<td>415.81</td>
<td>599.18</td>
<td>0.0303</td>
</tr>
<tr>
<td>MAE</td>
<td>297.5692</td>
<td>474.9476</td>
<td>0.2667</td>
</tr>
<tr>
<td>MSE</td>
<td>466.8474</td>
<td>650.0159</td>
<td>0.3142</td>
</tr>
</tbody>
</table>
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