LGD

Douglas Beserra Pinheiro[[1]](#footnote-2)\*

*Paketá Crédito*

**RESUMO**

*xxxxx.*

*xxxxxx*

*xxxxx*

*xxxxxx*

**Palavras-chave**: *IPO*, *IPO*, estabilização de preços, *aftermarket short covering*

**Classificação JEL**: G24

1. Introdução

Xxxxxxxxxxxxxxx xxxxxxxxxxxxx xxxxxxxxxxxxxxxx xxxxxxxxxxxxxxxxxxxxxx xxxxxxxx.

FALAR DO LGD, SUA IMPORTÂNCIA E DESAFIOS.

ESTUDOS NO BRASIL

ESTUDOS NO MUNDO E O USO DE ML

EXISTEM POUCOS ESTUDOS

|  |
| --- |
| . |

Este artigo está organizado da seguinte forma. A Seção 2 detalha o funcionamento do processo de estabilização. A Seção 3 descreve os dados e variáveis. A Seção 4 apresenta as hipóteses e a metodologia empregada. A Seção 5 apresenta e discute os resultados empíricos. Finalmente, a Seção 6 conclui este artigo.

2. Machine Learning e Crédito (Algo assim)

Segundo

2000) reporta que os *ASC* chega a ser realizado ao longo de meses.

3. Metodologia

**Explicar work-out**

4. Dados e Variáveis

Esta seção descreve a fonte de informação utilizada e apresenta uma discussão sumária dos dados e tratamentos.

4.1. Dados

A base com as informações dos empréstimos foi obtida do Kaggle[[2]](#footnote-3) e possui 141 variáveis e 2.925.493 empréstimos realizadas entre 2007 e 2020 pelo *Lending Club*. Utilizando o mesmo critério de Zhou et al. (2018) foram selecionados todos os empréstimos na condição de baixados por inadimplência, o que normalmente ocorre 120 dias após o vencimento da operação. No entanto, diferente de Zhou et al. (2018) e para melhorar a apuração do LGD foi incluído o custo de recuperação cobrado pelo *Lending Club* dos investidores, que representa em média 17,5% da recuperação observada. O LGD, portanto, foi definido considerando o seguinte critério:

|  |  |  |  |
| --- | --- | --- | --- |
| *LGD = 1 –* [ | *(Recuperação – Custo de Cobrança)* | ] | (1) |
| *(Valor do Empréstimo – Principal Pago)* |

A figura 1 apresenta a quantidade de operações baixadas, o que soma 363.309 empréstimos. O LGD médio gira em torno de 90%, porém a quantidade de registros com LGD igual a 100% (ou seja, aqueles casos em que não se observa nenhuma recuperação) aumenta significativamente após 2018 em função do menor período de apuração da recuperação, e é significativamente menor que a média anterior a 2014, portanto esses períodos foram excluídos da análise. Também foram excluídas 31 operações cujo valor principal do empréstimo foi integralmente pago, impedindo o cálculo do LGD como proposto aqui, o que resultou em uma amostra final de 266.515 observações.

**Figura 1**

**Quantidade de Empréstimos Anuais e LGD Médio**

Gráfico, Histograma

Descrição gerada automaticamente

A variável LGD refere-se aquela apresentada na equação 1, já a variável LGD=1 refere-se à proporção de empréstimos que não apresentaram recuperação.

Zhou et al. (2018) excluem operações cujos clientes não tiveram renda verificada e também operações com informações faltantes, porém optamos por mantê-las em nossas análises. No caso da renda verificada não há diferença relevante no LGD observado e para os casos de informações faltantes, além de serem pouco representativos também consideramos essas situações no desenvolvimento do estudo.

Quanto ao custo do dinheiro no tempo não foi considerado pois não há a informação de data de inadimplência e data de recuperação, além disso, o efeito não deve ser relevante pois o work-out do LGD se estabiliza em média 2 anos após a data da concessão, conforme observado no gráfico 1.

A figura 2 apresenta o LGD distribuído nos diferentes anos da amostra. Observa-se uma concentração da distribuição à direita, o que é esperado para um produto de varejo sem garantias associadas, que acarreta um elevado nível de perda após a ocorrência da *inadimplência. Além disso, não há variação relevante entre os diferentes anos.*

**Figura 2**

**Distribuição do LGD**

Interface gráfica do usuário

Descrição gerada automaticamente com confiança baixa

4.2. Variáveis

Uma deficiência do conjunto de variáveis utilizado é que a maioria se refere às informações no momento da concessão. A inclusão das informações de “ead”, “ead\_inc” e “ead\_loan”, cujo saldo é capturado no momento da inadimplência, reduz em parte essa deficiência.

No processo de modelagem, normalmente são selecionadas informações que possuem sentido econômico sobre o evento analisado. Optamos aqui por utilizar o mesmo conjunto de variáveis de Zhou et al. (2018) e a tabela 1 apresenta as mesmas, explicando o sentido o efeito esperado individualmente. Além delas, incluímos também algumas informações sobre o comportamento de crédito do cliente, saldo do contrato no momento da inadimplência e o uso declarado dos recursos, também detalhadas na tabela 1.

DETALHAR OS TRATAMENTOS DE ZEROS E CENSURA NAS VARIÁVEIS EXPLICADAS

|  |  |  |
| --- | --- | --- |
| **Tabela 1** | | |
| **Informações utilizadas na Modelagem Tradicional** | | |
| **Variável** | **Definição** | **Sinal esperado** |
| *lgd* | Perda percentual em relação ao saldo inadimplido |  |
| *loan\_amnt* | Valor do empréstimo concedido (para a análise de regressão, adotado o logaritmo natural dos valores em US$ mil). | No paper original em nível |
| *term\_c* | Assume 1 quando o prazo do empréstimo é de 60 meses e 0 quando o prazo é 36 meses. |  |
| *int\_rate\_n* | Taxa anual de juros do empréstimo (para a análise de regressão, adotado o logaritmo natural) | No paper original em nível |
| *grade* | Rating de crédito definido na concessão, a classificação é de A (menor risco de inadimplência) a G (maior risco) |  |
| *ead* | Valor do saldo do empréstimo no momento da inadimplência (para a análise de regressão, adotado o logaritmo natural dos valores em US$ mil). | No paper original em nível |
| *ead\_inc* | Valor do saldo do empréstimo no momento da inadimplência dividido pela renda anual do devedor |  |
| *ead\_loan* | Valor do saldo do empréstimo no momento da inadimplência dividido pelo valor concedido |  |
| *emp\_length\_n* | Anos de emprego do devedor, limitado a 0,5 se for menor que 1 ano e 10 se for maior. |  |
| *home\_c* | Classifica a moradia em (0) Alugada, (1) Financiada, (2) Outros e (3) Própria |  |
| *annual\_inc* | Renda anual do devedor (para a análise de regressão, adotado o logaritmo natural dos valores em US$ mil). |  |
| *dti* | Comprometimento mensal de renda (parcela/salário) excluindo outros empréstimos da Lending Club e Crédito Imobiliário. |  |
| *delinq\_2yrs* | Número de atrasos acima de 30 dias nos últimos 2 anos |  |
| *mths\_since\_last\_delinq* | Número de meses desde que ocorreu o último atraso > 30 dias pelo devedor |  |
| *mths\_since\_last\_delinq\_null* | Indicador de que não há histórico de atraso do devedor |  |
| *pub\_rec* | Número de registros negativos do devedor no mercado |  |
| inq\_last\_12m | Número de consultas de crédito do devedor nos últimos 12 meses |  |
| purpose | propósito do empréstimo, dividido em 12 categorias |  |
| fico\_range\_high (low) | The upper boundary range the borrower’s FICO at loan origination belongs to. |  |
| **Variável** indica o nome da variável utilizada na avaliação tradicional. **Definição** traz uma breve explicação sobre a construção da variável. **Sinal Esperado** traz uma breve explicação sobre o efeito esperado da variável sobre o LGD. Variáveis indicadas com \* foram utilizadas por Zhou *et al.* (2018). | | |

5. Metodologia

Descrever Hipóteses

.

6. Discussão e conclusão

A possibilidade

Referências

Zhou, G., Zhang, Y. e Luo S. (2018). P2P Network Lending, Loss Given Default and Credit Risks, *Sustainability* **10**(4): 1010.

URL: <https://doi.org/10.3390/su10041010>

1. \*Paketá Crédito, São Paulo, Brasil. Email: [douglas.pinheiro@paketa.com.br](mailto:douglas.pinheiro@paketa.com.br) [↑](#footnote-ref-2)
2. Utilizar o link <https://www.kaggle.com/code/ztrimus/loan-repayment-prediction/data?select=Loan_status_2007-2020Q3.gzip> para acesso à base e ao dicionário de dados. [↑](#footnote-ref-3)