Insper Instituto de Ensino e Pesquisa
Faculdade de Economia e Administração

Juracy Britto Magalhães Junior

REPLICANDO RATINGS DE CREDITO COM BASE EM INDICES FINANCEIROS

São Paulo
2014

Juracy Britto Magalhães Junior

Replicando ratings de crédito com base em índices financeiros

Monografia apresentada ao curso de Ciências Econômicas, como requisito parcial para a obtenção do Grau de Bacharel do Insper Instituto de Ensino e Pesquisa.

Orientador: Prof. Dr. João Manoel P. de Mello

São Paulo
2014

Magalhães Jr., Juracy Britto

Replicando ratings de crédito com base em índices financeiros / Juracy Britto Magalhães Jr. – São Paulo: Insper, 2014.

32 f.

Monografia: Faculdade de Economia e Administracão Insper Instituto de Ensino e Pesquisa.

Orientador Prof. Dr. João Manoel P. de Mello

1.Rating 2. Risco de crédito 3. Probabilidade de default

Juracy Britto Magalhães Junior

Replicando ratings de crédito com base em índices financeiros

Monografia apresentada ao curso de Ciências Econômicas, como requisito parcial para a obtenção do grau de Bacharel do Insper Instituto de Ensino e Pesquisa.

Aprovado em Dezembro 2014

Examinadores

Prof. Dr. João Manoel P. de Mello

Prof. Dr. Rinaldo Artes

Prof. Dr. Rodrigo Moita

Resumo

Magalhães, Juracy. Replicando ratings de crédito com base em índices financeiros. São Paulo, 2014. 13p. Monografia – Faculdade de Economia e Administração. Insper Instituto de Ensino e Pesquisa.

Esse estudo procura preencher um vazio acadêmico: replicar o rating de crédito, dado por empresas de rating, para empresas do mercado brasileiro. O principal objetivo é conseguir através de um método multivariado, com base em diversos índices financeiros, replicar os rating de créditos de agências como Moody's, Fitch e S&P, de forma mais acurada possível. Com isso conseguiremos entender melhor como cada uma das variáveis financeiras afetam a probabilidade de default, possibilitando, ainda, que tomadores de decisão possam deixar suas empresas mais atrativas, da ótica dos analistas de risco.

Palavras-chave: Rating, Risco de crédito, Probabilidade de default.

Abstract

Magalhães, Juracy. Replicating credit ratings based on financial ratios. São Paulo, 2014. 13p. Monograph – Faculdade de Economia e Administração. Insper Instituto de Ensino e Pesquisa.

This study tries to fill an academic gap: replicate the credit rating given by rating agencies to companies in the Brazilian market. The main goal is, through a multivariate method based on various financial ratios, replicate the credit rating given by agencies such as Moody's, Fitch and S & P, the most accurate as possible. Through this analysis it's possible to understand how each financial variable affects the probability of default, enabling further decision-makers to make their companies more attractive, in the perspective of risk analysts.

Keywords: Rating, Credit risk, Probability of default

Sumário

1. **Introdução** 8

2. **Fundamentação Teórica** 10

3. **Metodologia** 12
   3.1 Variáveis Propostas 14
   3.2 Dados 15
   3.3 Modelagem Estatística 15
   3.4 Resultados Esperados 16

4. **Plano de Trabalho** 18

5. **Resultado** 19
   5.1 Modelo 19
   5.2 Efeito Marginal Médio 20
   5.3 Comparação do Modelo Proposto Com o Modelo CS ...21
   5.4 Goodness-to-Fit Tests 22
   5.5 Backtests 24

6. **Conclusão** 25

   Referências.....26

1. Introdução

Risco pode ser definido como desvios nos objetivos empresariais decorrentes de incertezas nas variáveis que impactam os mesmos. Risco seria portanto os desvios com relação ao valor esperado. Gestores e empresas do setor financeiro como um todo procuram implementar a gestão de risco para diminuir o risco de default, o que, muitas vezes, leva a falência. Westgaard e Wijst (2001) definem default como o descumprimento de um contrato que vincula devedor a credor.

Modigliani and Miller (1958) sugerem que as políticas de endividamento da empresa serão irrelevantes, ou até desaconselháveis, caso certas premissas se cumpram, tais como: os investidores possam administrar o risco com custo e eficiência iguais aos da empresa, não haja o benefício fiscal de deter capital de terceiros (dívida), não haja custo de transação e nem assimetrias no mercado. No entanto, como levantando pelos próprios autores na conclusão de sua obra, o que se observa na realidade é que nem todas as hipóteses se verificam.

Consecutivas crises econômicas mundiais ressaltaram a importância de se ter uma ativa e competente gestão de risco. Tratados, como o de Basiléia, obrigam instituições financeiras, se quiserem sobreviver de forma competitiva no mercado atual, a se atualizarem constantemente nesse assunto. Cada vez mais empresas adotam portanto práticas de análise de risco, especialmente a análise do risco de crédito, que tem como objetivo principal tentar mensurar a probabilidade de default, e com isso montar suas carteiras de forma a minimizar este risco.

Empresas e agências de rating adotam modelos diversos para tentar atingir esse objetivo. A Moody's, por exemplo, adota o modelo BET[1], enquanto que o Credit Suisse utiliza do modelo FACT, uma variação da metodologia UTADIS. Segundo profissionais da área o desenvolvimento desses modelos, na maioria das vezes, é um processo custoso e trabalhoso, dispondo, na maioria dos casos, de uma base de dados escassa para se modelar probabilidade de default.

O objetivo desse estudo é tentar replicar ratings de crédito, uma representação de probabilidade de default, de agências de crédito, com base em indicadores financeiros, dados que esses na maioria das vezes, são mais fáceis de obter no mercado. Pretende-se ainda preencher a lacuna levantada por Murcia (2013) que está presente nos estudos desse tópico, uma vez que poucos estudos abordam o assunto tentando replicar os ratings de crédito, e menos ainda são aqueles que o fazem no mercado brasileiro. O objetivo final é tentar gerar

[1] Woo and Siu (2003)

um modelo que seja menos custoso e menos trabalhoso, e que ainda assim seja preciso na quantificação da probabilidade de default, podendo ser usado por gestores e empresas na alocação de ativos em suas respectivas carteiras.

Na seção 2 será discutido o referencial teórico no qual o estudo se embasou: teorias e estudos já realizados no tópico que ajudaram no desenvolvimento do trabalho. Na seção 3 será mostrado o modelo que usaremos para medir essa probabilidade de default, a metodologia que sustenta o modelo e aspectos como a origem da base de dados e o período amostral. Na seção 4 será apresentado um plano de trabalho, seguido por um modelo proposto e devidamente testado na seção 5. Por fim, na seção 6 serão feitas as considerações finais.

2.Fundamentação Teórica

Nos últimos anos a gestão do risco de crédito obteve um papel cada vez mais importante nas empresas, especialmente nas instituições financeiras. Crises financeiras, como as de 2001 e 2008, trouxeram a tona a importância dessa prática, e geraram acordos de regulamentação e monitoramento, como o Basileia 1,2 e 3.

Junto com a propagação dessa prática vieram os avanços na metodologia de se medir o risco de crédito e, consequentemente, a probabilidade de default. Woo and Siu (2003) citam as duas principais escolas de pensamento metodológicas: os modelos estruturais com base no valor dos ativos da carteira, originalmente proposto por Merton (1974), e os *Intensive-Based Reduced Form Models*. A primeira foca em avaliar variações no valor dos ativos da carteira, que se assume seguir um movimento geométrico Browniano, que quando atinge um valor *threshold* gera o default, ou o gerará em algum momento próximo. As dificuldades impostas pelo modelo são que tanto o valor real da firma quanto o valor do *threshold* dos ativos não são observáveis. Já a segunda escola metodológica foca em medir a probabilidade de default em um certo momento instantâneo, sem se preocupar em modelar as variáveis que levaram ao evento de default.

Os autores ainda citam o modelo desenvolvido pela agência de rating Moody's: Binomial Expansion Technique (BET). O modelo consite em montar uma carteira de ativos arriscados, possíveis de dar default. Duas limitações do modelo são a estaticidade do modelo, que não permite capturar mudanças intertemporais da probabilidade de default, e a hipótese de independência entre as probabilidades de default dos ativos da carteira, o que na prática não se observa. Para mais detalhes no modelo BET consultar Moody's (1996), Bluhm *et al*. (2003), Duffie and Singleton (2003) e Schonbucher (2003). Em seu trabalho os autores inserem um elemento intertemporal no modelo BET para capturar as variações do preço na carteira ao longo do tempo, e com isso calcular a perda esperada e o value-at-risk de crédito (CVAR), e usam um modelo de filtro Bayesiano para conseguir calcular a probabilidade das firmas sobreviverem no mercado, o complementar da probabilidade de default. Para mais detalhes no modelo proposto consultar Woo and Siu (2003).

Doumpos and Pasiouras (2004) usam uma outra abordagem para tentar replicar, da forma mais precisa possível, os ratings das agências, teste ainda pouco trabalhado em trabalhos anteriores sobre o mesmo tópico. Esses mostram em seu trabalho que é possível se atingir resultados satisfatórios com dados públicos de empresas, chegando a um rating próximo ao dado pelas agências especializadas.

Os autores se distinguem dos demais em dois aspectos: dividindo as empresas em cinco grupos, dado sua probabilidade de default e o uso de um vetor de indicadores financeiros para chegar a um rating. O nosso estudo consiste em tentar replicar esses resultados usando dados de empresas brasileiras, tendo em vista que o estudo usou como base de dados empresas britânicas. A metodologia usada no estudo foi a *multicriteria decision aid (MCDA*), especificamente a UTADIS, muito usada no setor bancário especialmente para relatórios internos (Treacy and Carey, 2000). Para maiores detalhes sobre o método UTADIS consultar Doumpos and Zopounidis (2002).

É importante ainda, dado que a metodologia usa de diversos indicadores financeiros, identificar os indicadores que se mostram significantes no período analisado, para não se ter problemas de sobreespecificação do modelo (Doumpos and Pasiouras, 2004).

Florez and Ramon (2013) apontam que uma grande limitação desse tipo de estudo é a característica que se vem observado nas carteiras de ativo, especialmente no pós crise. Segundo os autores, as carteiras cada vez mais, tem se caracterizado como *Low Default Portfolio* (LDP), ou seja, carteiras que têm tido menos eventos de defaults. Isso dificulta a modelagem da *Probability of Default* (PD) devido ao tamanho reduzido da amostra, que gera tanto uma assimetria, entre a amostra de empresas que tiveram o evento de default e empresas que não tiveram, e o problema de má especificação no modelo.

É necessário que a amostra seja a mais diversificada possível, abrangendo empresas de um maior número de mercados, para que se exclua o risco sistêmico de algum mercado específico (Breiman, 2001). Outra condição necessária é que a amostra escolhida tenha o mínimo de correlação possível entre as empresas, para poder captar os efeitos das variações dos indicadores financeiros e não algum problema pontual em um setor ou em um grupo de empresas (Dietterich,1997). Essas condições combinadas tendem a otimizar a capacidade de medir a PD (Florez and Ramon, 2013).

Nesse estudo procuramos replicar os ratings dados pelas agências de crédito, entendendo as variáveis que realmente afetam a sua formação, sua variação durante um dado período e tentando superar a limitação levantada por Florez and Ramon (2013).

3. Metodologia

O principal intuito desse estudo é conseguir gerar um "score" de crédito, com base em uma probabilidade de default, para um grupo de empresas e com base nesse "score" traduzi-lo em um rating, similar aos atribuídos por agencias de rating. Procura-se replicar os ratings dado pelas agências usando como variáveis explicativas indicadores financeiros, atribuindo-os pesos (Doumpos and Pasiouras, 2004). Diferentemente dos ratings de crédito gerado por agências de rating, serão usadas apenas variáveis quantitativas e não subjetivas (i.e.: grau de governança coorporativa da empreas), podendo assim, no final do estudo, avaliar o quão necessário essas variáveis de caráter subjetivo são para a mensuração da probabilidade de default.

A amostra será composta apenas de empresas brasileiras, ou que atuem no Brasil, com indicadores financeiros no período no qual foi atribuído o ultimo rating. Para empresas que no período *t* for atribuído um rating "D", situação na qual a empresa se encontra em default, serão utilizados então os indicadores financeiros do período *t-1*, uma vez que o objetivo do estudo é gerar um probabilidade de default *ex-ante.* Pretende-se, portanto, encontrar variáveis antecedentes para o evento de insolvência (default).A base de dados será coletado dos setores de Relação com Investidor das próprias empresas, setor responsável pela apresentação dos demonstrativos financeiros das empresas ao público, e também dados internos do sistema de risco do Credit Suisse no Brasil, o Risk Portal Insight. Serão escolhidas apenas empresas que tiverem um rating no Credit Suisse, para poder ter assim uma base de comparação com o rating proposto pelo modelo. Na prática, dado que o modelo utilizado pelo CS tem um forte peso para variáveis qualitativa (i.e. qualidade de governança, posicionamento de mercado, etc), caso o modelo proposto mostre aderência e forte capacidade de previsibilidade é possível concluir que as variáveis qualitativas não contribuem para a mensuração da probabilidade de default

O modelo proposto é similar ao de Doumpos and Pasiouras (2004), levando em consideração as precauções levantadas por Florez and Ramon (2013). Primeiramente serão escolhidos diversos indicadores financeiros, que reflitam liquidez, alavancagem, desempenho operacional, despesas com juros e quaisquer outros aspectos que se julguem importantes. Os indicadores financeiros serão coletados no período em que cada rating correspondente foi dado, cuja grande maioria foi no ano fiscal terminado em Dez/13, data na qual foi publicado o último Demonstrativo Financeiro Padronizado das empresas. Para as empresas em que se observou o evento do default, foi recolhido também o último rating

atribuído à ela antes de ter se observado o evento ("rating pré-default"), e seus indicadores financeiros no período do rating, como explicado anteriormente. O intuito é de conseguir prever o evento do default *ex-ante*, e com isso testar a aderência do modelo e sua capacidade de previsibilidade.

Vale também ressaltar uma forte hipótese simplificadora que foi assumida no estudo: empresas com atraso de pagamento de dívida acima de 180 dias são consideradas insolventes, consequentemente tendo apresentado o evento de default, independente de posteriormente terem renegociado sua dívida ou terem entrado no processo de recuperação judicial. Essa é uma hipótese costumeiramente assumida nesse tipo de estudo e ajuda com o maior empecilho nesse tipo de base de dados que é a escassez de empresas em que se observa o evento do default, características de portfolios pouco arriscados que os bancos e empresas vem montando, especialmente depois de diversas crises mundiais como a de 2008. Essa hipótese também é razoável pois dado um evento desse tipo a instituição credora incorre em perdas com a operação de crédito concedida, situação que a instituição procura prever e evitar.

O modelo proposto será rodado com base em um modelo de regressão logística. Na literatura muitas vezes o modelo Probit é usado ao invés do Logit (regressão logística), no entanto pelo modelo Logit não assumir certas hipóteses[2] (i.e. normalidade das variáveis independentes, correlação zero entre essas variáveis, etc) e por uma maior simplicidade dos cálculos, o usado no estudo será o Logit.

A regressão logística é um modelo cuja variável resposta é uma variável Bernoulli, ou dicotômica, que assumi o valor 1 caso o evento seja observado e 0 caso o contrário (no caso: 1 no caso do evento de default/insolvência ser observado e 0 caso o contrário)[3]. A função de distribuição logística (acumulada) apresenta dois aspectos que são interessantes para o modelo: (1) quando X_i aumenta, $P_i = E(Y=1 \mid X)$ aumenta, mas nunca fica fora do intervalo $\{0,1\}$; e (2) a relação entre P_i e X_i é não linear, ou seja, capta efeitos diferentes de X_i sobre P_i conforme o nível no qual X_i ou qualquer uma das variáveis se encontrem[4].

O modelo apresenta a seguinte forma[5]:

$$P(y=1 \mid x) = P(y=1 \mid x_1, x_2, \ldots, x_k) = G(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k) = G(\beta_0 + x\beta)$$

[2] Hosmer and Lemeshow (1989)
[3] Gujaratti and Porter (2011)
[4] Gujaratti and Porter (2011)
[5] Wooldridge (2009)

$$G(\beta_0 + x\beta) = \exp(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k) / [1 + \exp(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k]$$

Tendo o método estatístico definido a próxima etapa será rodar o modelo usando a base de dados obtida e o programa Stata[6] . Para definir as variáveis que são relevantes ao modelo será usado o *Teste de Wald*, similar ao teste T usado em modelos lineares, só que seguindo uma distribuição *Qui-Quadrado*. O nível de significância adotado é de 10%, um nível de significância comumente observado na literatura.

3.1 Variáveis Propostas

As variáveis a serem propostas deverão ser dívidas 6 categorias de interesse, de forma a conseguir cobrir o máximo da saúde financeira da empresa e assim gerar um rating de crédito mais apurado possível. As categorias das variáveis são: (i) Fluxo, (ii) Porte, (iii) Cobertura, (iv) Performance, (v) Estrutura de Capital, e (vi) Liquidez.
As variáveis e suas respectivas categorias são:

1. **Fluxo**
   I. Retorno/Capitalização
   II. Fluxo de Caixa Operacional/Dívida Bruta
   III. Dívida Bruta/Capitalização
2. **Porte**
   I. Faturamento
   II. EBITDA
   III. PL
   IV. Ativo Total
3. **Cobertura**
   I. EBITDA/Despesa Financeira Líquida
   II. EBIT/Dívida Líquida
   III. EBIT/Passivo Total
4. **Performance**
   I. Resultado Após Despesas Financeira/Faturamento
5. **Estrutura De Capitais**
   I. PL /Passivo Total
   II. Dívida Bruta/PL
   III. Dívida de CP/PL
6. **Liquidez**
   I. Ativo Circulante/Passivo Circulante
   II. (Ativo Circulante-Passivo Circulante)/Ativo Total

[6] © Copyright 1996–2014 StataCorp LP

III. (Caixa + Contas a Receber)/Passivo Circulante

3.2 Dados

Como já discutido anteriormente os dados serão recolhidos no último período no qual foi atribuído um rating de crédito à empresa, exceto para empresas que apresentaram o evento de default, que terão dados recolhidos no período em que foi atribuído o ultimo rating de crédito anterior a observação do evento. Os dados serão retirados do sistema interno de risco do Credit Suisse no Brasil, o Risk Portal Insight. A amostra terá que ser condicionada também a empresas que tenham rating de crédito disponível no Credit Suisse, para possibilitar testar a aderência do modelo posteriormente. Dado esses critérios, totalizaram-se 101 observações na amostra.

3.3 Modelagem Estatística

O modelo usado para tentar medir a probabilidade de default será um modelo Logístico. O modelo proposto tem a seguinte apresentação[7]:

$P(Y_i=1 \mid x_1, x_2, \ldots, x_{17}) = \exp(ß_0 + xß) / [\, 1 + \exp(ß_0 + xß)\,]$

$(ß_0 + xß) = ß_0 + ß_1RC_i + ß_2FFO_i + ß_3TD_i + ß_4FAT_i + ß_5EBITDA_i + ß_6PL_i + ß_7AT_i + ß_8EBITDA/DESP.FIN.LIQ_i + ß_9EBIT/DÍV.LÍQ._i + ß_{10}EBIT/PASSIVOTOTAL_i + ß_{11}RES.APÓSDESP.FIN./PL_i + ß_{12}PL/PASSIVOTOTAL_i + ß_{13}DÍV.BRUTA/PL_i + ß_{14}DÍV.CP/PL_i + ß_{15}ATIVO\ CIRCULANTE/PASSIVO\ CIRCULANTE_i + ß_{16}(ATIVO\ CIRCULANTE – PASSIVO\ CIRCULANTE)/ATIVO\ TOTAL_i + ß_{17}(CAIXA + CONTAS\ A\ RECEBER)/PASSIVO\ CIRCULANTE_i + ERRO_i$

Onde:

Y_i : variável dicotômica que assumi 1 no caso da empresa i apresentar o evento de default e 0 caso contrário
RC_i: Razão retorno capitalização da empresa i
FFO_i: Razão fluxo de caixa operacional dívida bruta da empresa i
TD_i: Razão dívida bruta capitalização da empresa i
FAT_i: Faturamento da empresa i
$EBITDA_i$: EBITDA da empresa i
PL_i: Patrimônio Líquido da empresa i
AT_i: Ativo Total da empresa i
$EBITDA/DESP.FIN.LIQ._i$: Razão EBITDA Despesa Financeira Líquida da empresa i
$EBIT/DÍV.LÍQ._i$: Razão EBIT Dívida Líquida da empresa i
$EBIT/PASSIVOTOTAL_i$: Razão EBIT Passivo Total da empresa i
$RES.APÓSDESP.FIN/PL_i$: Razão resultado, descontado das despesas financeiras, PL da empresa i

[7] Wooldridge (2009)

PL/PASSIVOTOTAL$_i$: Razão PL Passivo Total da empresa i
DÍV.BRUTA/PL$_i$: Razão Dívida Bruta PL da empresa i
DÍV.CP/PL$_i$: Razão Dívida de Curto Prazo PL da empresa i
ATIVOCIRCULANTE/PASSIVOCIRCULANTE$_i$: Índice de liquidez corrente da empresa i
(ATIVOCIRCULANTE – PASSIVOCIRCULANTE)/ATIVOTOTALo$_i$: Razão Ativo de circulante, descontado o passivo circulante, ativo total da empresa i
(CAIXA + CONTAS A RECEBER)/PASSIVOCIRCULANTE$_i$: Índice de liquidez seca da empresa i
ERRO$_i$: Erro do modelo que por hipótese segue uma normal padrão

3.4 Resultados Esperados

Para as variáveis propostas no modelo esperam-se os seguintes resultados:

Variáveis	Sinal Esperado
Fluxo	
Retorno / Capitalização	MENOR QUE ZERO
Fluxo de Caixa Operacional / Dívida Bruta	MENOR QUE ZERO
Dívida Bruta / Capitalização	MAIOR QUE ZERO
Porte	
Faturamento	MENOR QUE ZERO
EBITDA	MENOR QUE ZERO
PL	MENOR QUE ZERO
Ativo Total	MENOR QUE ZERO
Cobertura	
EBITDA / Desp.Fin.Líq.	MENOR QUE ZERO
EBIT / Div.Líq.	MENOR QUE ZERO
EBIT / Passivo Total	MENOR QUE ZERO
Performance	
Res. Após Desp. Fin. / Faturamento	MENOR QUE ZERO
Estrutura de Capital	
PL / Passivo Total	MENOR QUE ZERO
Dívida Bruta / PL	MAIOR QUE ZERO
Dívada de CP / PL	MAIOR QUE ZERO
Liquidez	
AC / PC	MENOR QUE ZERO
(AC - PC) / Ativo Total	MENOR QUE ZERO
(Caixa + Contas a Receber) / PC	MENOR QUE ZERO

As variáveis que tem o sinal esperado maior que zero representam variáveis que conforme vão aumentando pior fica a situação financeira da empresa, consequentemente, pioram o rating de crédito, e, em última estância, aumentando a probabilidade de default. Já para as variáveis que apresentam o sinal esperado menor que zero, essas, conforme a variável assume valores maiores, melhor fica a situação financeira da empresa, o rating de crédito e, consequentemente, diminui a probabilidade de default.

4. Plano de Trabalho

Na próxima seção será gerado o modelo obedecendo todos os critérios e requisitos definidos anteriormente. Como também já explicitado, para tal será usado o programa Stata[8] e com base dos resultados será usado o *Teste de Wald*, adotando um nível de significância de 10%, para se excluir as variáveis que não se mostrarem significantes, chegando por fim ao modelo final. Ainda nessa seção serão feitas uma análise dos Efeitos Marginais Médios de cada uma das variáveis, serão comparados os resultados obtidos com os resultados observados no modelo do Credit Suisse, além de feitos *fit-tests* (Teste de Hosmer e Lemeshow (1989)[9] e Teste de Pearson[10]) para testar se o modelo está de fato bem especificado e se os resultados apresentaram aderência em comparação aos dados obtidos do Credit Suisse. Por fim, na seção 6 serão feitas as considerações finais.

5. Resultados

5.1 Modelo

[8] © Copyright 1996–2014 StataCorp LP
[9] Hosmer and Lemeshow (1989)
[10] Hosmer, Hosmer, Le Cessie, and Lemeshow (1997)

O primeiro modelo gerado, ainda com todas as variáveis, apresentou o seguinte resultado:

Regressão Logística

	Número de obs	=	101
	LR $chi^2(13)$	=	30.44
	Prob > chi^2	=	0.0041
Log Likelihood = -23.559512	Pseudo R^2	=	0.3924

Variáveis	Coef.	Std. Error	Z	P > | Z |	[90% Confidence Interval]	
Faturamento	-1.29E-09	-1.50E-09	0.86	0.391	-1.65E-09	-4.22E-10
EBITDA	-2.27E-08	1.42E-08	-1.60	0.110	-5.06E-08	5.16E-09
PL	8.61E-10	3.55E-09	0.24	0.809	-6.10E-08	7.82E-09
Ativo Total	-1.20E-09	2.26E-09	-0.53	0.597	5.64E-09	3.24E-09
EBITDA / Desp.Fin.Líq.	0.42481	0.02421	1.75	0.079	-0.00496	0.89932
EBIT / Div.Líq.	0.02290	0.66527	0.03	0.973	-1.28100	1.32679
EBIT / Passivo Total	2.08916	4.93859	0.42	0.672	-7.59030	11.76892
Res. Após Desp. Fin. / Faturamento	-2.50151	1.62167	-1.54	0.123	-5.67992	0.67690
PL / Passivo Total	0.32516	1.28160	0.25	0.800	-2.18674	2.83705
Dívida Bruta / PL	-0.69839	0.55671	-1.25	0.210	-1.78953	0.39275
Dívada de CP / PL	0.87433	0.95428	0.92	0.360	-0.99603	2.74468
AC / PC	1.49953	1.73085	0.87	0.386	-1.89287	4.81922
(AC - PC) / Ativo Total	-1.96229	2.44732	-0.80	0.423	-6.75895	2.83437
(Caixa + Contas a Receber) / PC	-3.85528	2.25859	-1.71	0.088	-8.28203	0.57147
Retorno / Capitalização	-1.09453	4.48247	-0.24	0.807	-9.88000	7.69095
Fluxo de Caixa Operacional / Dívida Bruta	0.09564	0.81927	0.12	0.907	-1.51011	1.70138
Dívida Bruta / Capitalização	0.49731	0.98522	0.50	0.614	-1.43368	2.42830
Constante	-2.12267	2.13536	-0.99	0.320	-6.30790	2.06255

Como descrito anteriormente foi então efetuado um *Teste Wald* para cada uma das variáveis para testar a sua significância. Análogo ao *Teste T* em modelos lineares, o *Teste Wald* tem as seguintes hipóteses nula e alternativa:

$H_{0:}\ \beta_j = 0$

$H_1: \beta_j \neq 0$

Adotou-se um α=10%. Como pode ser observado na tabela acima a variável com o maior p-valor foi EBIT/Dívida Líquida, portanto foi a primeira variável a ser retirada do modelo. Então foi novamente rodado o modelo sem a variável que se mostrou não significante. Essa metodologia foi repetida diversas vezes até que todas as variáveis se mostrassem significantes, dado o α. O resultado final obtido, onde todas as variáveis foram significantes, está apresentado na tabela abaixo [11] :

Regressão Logística

	Número de obs	=	101
	LR $chi^2(13)$	=	35.97
	Prob > chi^2	=	0
Log Likelihood = -20.79207	Pseudo R^2	=	0.4638

Variables	Coef.	Erro Pad.	Z	P > | Z |	[Intervalo de Confiança de 90%]	
EBITDA	-2.42E-08	-9.91E-08	-2.44	0.015	-4.36E-08	-4.76E-09
EBIT / Div.Líq.	-2.64901	1.45213	-1.82	0.068	-5.49513	0.19710
EBIT / Passivo Total	26.49331	10.37757	2.50	0.011	6.15365	46.83296
Res. Após Desp. Fin. / Faturamento	-3.77469	1.62323	-2.33	0.020	-6.95616	-0.59323
(Caixa + Contas a Receber) / PC	-2.01504	1.12494	-1.79	0.073	-4.21988	0.18980
Constante	-0.94285	0.87065	-1.08	0.279	-2.64929	0.76359

[11] Os estimadores foram obtidos a partir do método de Máxima Verossimilhança (MLE): para maior detalhamento consultar o capítulo 17 de Wooldridge (2009)

Vale a pena chamar a atenção de que muitas variáveis podem ter se mostrado não significantes, pois a amostra é relativamente pequena (101 observações), problema que foi anteriormente discutido, assim como falado na literatura.

O programa calcula automaticamente o Pseudo $R^2(\rho^2)$, que está no intervalo (0,1), sendo igual a 1 quando o modelo se ajusta perfeitamente à amostra e igual a 0 caso contrário. No entanto, diferentemente do R^2 no caso de modelos lineares, o ρ^2 não tem significado intuitivo entre os valores. Ou seja, o ρ^2 mede a porcentagem da incerteza dos dados que é explicada pelos resultados empíricos[12].

5.2 Efeito Marginal Médio

Efeito Marginal Médio

Número de obs = 101

Expressão: Pr(default), previsao ()

∂y/∂x com relação a: EBITDA ; EBIT/Div. Liq. ; EBIT / Passivo Total ; Res.Após Desp. Fin. / Faturamento ; (Caixas + Contas a Receber) / PC

Variables	Coef.	Erro Pad.	Z	P > | Z |	[Intervalo de Confiança de 90%]	
EBITDA	-1.51E-08	5.48E-10	-2.75	0.006	-2.41E-09	-6.05E-09
EBIT / Div.Líq.	-0.16508	0.08340	-1.98	0.048	-0.30227	-0.02790
EBIT / Passivo Total	0.65101	0.55476	2.98	0.003	0.73852	2.56351
Res. Após Desp. Fin. / Faturamento	-0.23523	0.08886	-2.65	0.008	-0.38140	-0.08906
(Caixa + Contas a Receber) / PC	-0.12557	0.65952	-1.90	0.057	0.23406	-0.01709

Diferentemente de modelos lineares, no modelo Logit não se consegue analisar o impacto exato do efeito marginal de cada variável na variável resposta, devido a sua distribuição. Como o modelo Logit segue uma distribuição logística o efeito marginal de cada variável depende do nível de todas as outras variáveis, para saber em que ponto da distribuição o modelo se encontra. Para efeito de análise olha-se o efeito marginal de cada uma das variáveis adotando para as outras variáveis o seu valor médio[13]. Esse resultado está exposto na tabela acima.

O resultado sugere que para um aumento de BRL 1.000 no EBITDA das empresas (os dados de porte foram registrados em milhares) a probabilidade de default diminui em cerca de 0.00000151%. Analogamente, se EBIT/Dív.Líq. , Res.ApósDesp.Fin./Faturamento ou (Caixa + Contas a Receber)/PC aumentarem em uma unidade, a probabilidade de default diminui em cerca de 16.5%, 23.5% e 12.6%, respectivamente.

Por outro lado, diferentemente do esperado, um aumento de uma unidade no EBIT/Passivo Total aumentaria em cerca de 65.1% a probabilidade de default. Possivelmente,

---

12 Hosmer and Lemeshow (1989)

13 Gujaratti and Porter (2011)

devido a choques exógenos, as empresas que apresentaram essa variável em um nível alto podem ter incorrido no evento de default, algo que não seria captado pelo modelo e poderia, possivelmente, gerar um viés no modelo final. Uma possível solução para esse problema seria retirar essas empresas da base de dados para conseguir medir de forma mais apurada o efeito da variação desse indicador sobre a probabilidade de default.

5.3 Comparação do Modelo Proposto Com o Modelo CS

Na tabela abaixo se encontram as empresas nas quais de fato se observou o evento de default, com seus respectivos ratings (no período anterior ao default), a probabilidade de default associado ao respectivo rating e a probabilidade de default gerado pelo modelo:

Empresa	Rating Pre-default	Prob. De Default Associada ao Rating	Probabilidade do Modelo
1	BB	1.10%	6.49%
2	BB-	1.89%	2.54%
3	BB-	1.89%	48.83%
4	BB-	1.89%	80.18%
5	B+	3.25%	50.00%
6	B+	3.25%	100.00%
7	B	5.58%	89.89%
8	CCC	28.23%	95.69%
9	CCC-	28.23%	27.87%
10	CC	28.23%	27.44%
11	CC	28.23%	44.74%
12	C	28.23%	19.41%
13	C	28.23%	37.03%
14	C	28.23%	99.54%

O modelo parece ter um caráter mais conservador, uma vez que quase todas as empresas que deram default tiveram uma probabilidade de default associadas a elas maior pelo modelo do que pelo rating a elas atribuídas. Isso provavelmente ocorre pelo caráter subjetivo do rating, algo que o modelo não cobre. Muitas vezes o rating pode ter sido atribuído sobre a perspectiva de fluxos de caixa/pagamentos futuro que não se concretizaram de fato, só que na visão do analista era o suficiente para sustentar o rating atribuído e, consequentemente, uma probabilidade de default que os números sozinhos não suportariam. Outro caráter que não tem como ser medido pelo modelo são as "manipulações" de ratings que eventualmente ocorrem em bancos para poder não só cobrar uma taxa mais atrativa do cliente, taxa sem a qual muitas vezes o *deal* em questão não sairia, quanto para manipular a quantidade de provisão que é requerido pelo Banco Central nos balanços dos bancos.

No entanto, quando analisado a probabilidade de default média gerada pelo modelo, agrupando as empresas por rating, vê-se que a probabilidade é similar com as probabilidades associadas aos ratings:

Rating	Número de Empresas	Prob. De Default Associada ao Rating	Prob. Média Estimada Pelo Modelo
A	1	0.07%	0.00%
BBB	22	0.30%	2.24%
BB	36	1.21%	9.94%
B	23	6.14%	15.03%
CCC	13	24.30%	23.77%
CC	4	28.23%	21.67%
C	2	28.23%	45.49%

É interessante observar que a probabilidade de default média, estimada pelo modelo proposto, para empresas que tiveram um rating CC apresenta uma queda quando comparada a probabilidade média estimada pelo modelo para empresas com rating CCC. Isso na prática não deveria ocorrer, uma vez que um rating CCC deveria ser atribuído a empresas que apresentem uma saúde financeira melhor, consequentemente apresentando uma probabilidade de default menor.

No caso específico em questão uma das quatro empresas da amostra de empresas CC é uma empresa que acabou incorrendo no evento de default mas por causa de uma variável exógena, especificamente nesse caso uma fraude. Esse tipo de variável não tem como ser captada pelo modelo, portanto essa observação possivelmente deveria ser excluída da amostra, uma vez que não contribui para o modelo e o seu resultado.

5.4 Goodness-to-Fit Tests

Em seguida foram rodados dois testes de especificidade, o teste de Hosmer e Lemeshow (1989) e o teste de Pearson; ambas as estatísticas usadas nos testes seguem uma distribuição *Qui-Quadrado*.

No teste de Hosmer e Lemeshow avalia-se o modelo comparando as frequências observadas com as esperadas, dividindo a amostra em subgrupos. Se as frequências observadas se assimilarem as frequências esperadas diz-se então que o modelo está bem calibrado[14]. As hipóteses são:

H_0: As frequências observadas se assemelham as frequências esperadas nos subgrupos
H_1: As frequências observadas não se assemelham as frequências esperadas nos subgrupos

[14] Hosmer and Lemeshow (1989)

Modelo Logístico para default, goodness-of-fit test

Número de obs	=	101
Número de agrupamentos	=	10
Hosmer-Lemeshow chi^2 (8)	=	3.38
Prob > chi^2	=	0.9084

Como o teste apresentou um p-valor alto (0.9084), conclui-se que a hipótese nula se verifica, logo, o modelo está bem calibrado.

Já o segundo teste, teste de Pearson, é um teste que analisa os resíduos do modelo, sendo útil para avaliar o quão bem o modelo se ajustou aos dados. Ele testa se existe independência entre os subgrupos nos quais a amostra é divida com base na frequência, condição necessária para garantir a precisão do modelo[15]. As hipóteses são similares as do teste de Hosmer e Lemeshow:

H_0: As frequências observadas não são diferentes das esperadas, portanto não existe diferença entre as frequências e há independência entre os subgrupos

H_1: As frequências observadas são diferentes das esperadas, portanto existe diferença entre as frequências e não há independência entre os subgrupos

Modelo Logístico para default, goodness-of-fit test

Número de obs	=	101
Número de padrões covariáveis	=	100
Pearson chi^2 (8)	=	75.39
Prob > chi^2	=	0.9206

Assim como verificado no primeiro teste, como o resultado foi um p-valor elevado (0.9206), se valida a hipótese nula, por tanto, os subgrupos se mostraram independentes, o ajuste dos dados de fato, estatisticamente, se mostra bom e o modelo bem especificado.

5.5 Backtest

Foram feitos também diversos *backtests* para testar a aderência do modelo à base de dados.

Primeiro foram rodados testes de aderência e acerto do modelo:

[15] Hosmer, Hosmer, Le Cessie and Lemeshow (1997)

	- Verdadeiramente -		
Classificação	**D**	**~ D**	**Total**
+	5	2	**7**
-	8	86	**94**
Total	**13**	**88**	**Total**

Pr (+ | ~ D)	2.27%
Pr (- | D)	61.54%
Pr (~ D | +)	28.57%
Pr (D | -)	8.51%
Corretamente Classificados	**90.10%**

Classificados + se Pr(default) > 0.5
Verdadeiramente definido como D ! = 0

Empresas que foram classificadas como " + " são empresas para as quais o modelo estimou uma probabilidade de default maior que 50%, enquanto que empresas classificadas como " – " tiveram uma probabilidade default menor que 50%.

Das 7 empresas que apresentaram uma probabilidade de default maior que 50%, 5 de fato deram default, enquanto que 2 não deram. Já das 94 empresas que tiveram uma probabilidade de default menor que 50%, 8 deram default, enquanto que 86 não. É razoável supor que é maior a probabilidade de uma empresa apresentar uma probabilidade de default menor que 50% sendo que ela de facto entrou em default (Pr(- | D) = 61.54%) do que uma empresa apresentar uma probabilidade de default maior que 50% e não apresentar o evento (Pr(+ |~ D) = 2.27%). Isso pode ocorrer devido a fatos exógenos que comprometem a capacidade de uma empresa de honrar seus pagamentos ou a fatos subjetivos (i.e: qualidade do corpo diretivo da empresa), que nem sempre são capturados através de números. O modelo, por fim, classificou corretamente 90.10% da amostra, mostrando que de fato aderiu bem aos dados.

6. Conclusão

Apesar de o modelo rodado ter se mostrado consistente com os dados o caráter subjetivo apresentado em modelos comumente usados no mercado se mostrou significante e importante quando feita uma análise de crédito. Quando feita uma análise desse cunho é importante se considerar aspectos que nem sempre são transparentes através dos indicadores financeiros da empresa, podendo ser desde aspectos micro (i.e.: governança corporativa,

ambiente de trabalho) até aspectos macro (i.e.: condições do setor no qual a empresa atua, perspectivas para o mercado no qual a empresa está situada). Uma evidencia empírica dessa afirmação é a diferença entre as probabilidades de default atribuída às empresas pelo CS em comparação com as probabilidades de default geradas pelo modelo.

Indo além, é importante também ressaltar que muitas das variáveis que foram excluídas do modelo, por se mostraram insignificantes estatisticamente, na prática se mostram importantes e significantes para a saúde financeira da empresa e, consequentemente, na sua probabilidade default. Muito possivelmente a sua insignificância estatística se deu pelo tamanho limitado da amostra, algo que possa vir a ser trabalhado em outros estudos em que se tenham a oportunidade de desfrutar de uma amostra mais volumosa.

Por fim, pode-se concluir que apesar das limitações do estudo, que estão também presentes no tema como um todo, o modelo gerado é satisfatório, aderi bem aos dados e apresenta resultados coerentes com o que se foi observado na prática, ainda que no dia-a-dia das análises de crédito é importante de fato considerar aspectos intangíveis e não numéricos.

Referências

Amemiya, T. "Advanced Econometrics." *Havard University*. Havard University Press, 1985. Web.

Bluhm, C., Overbeck, L. and Wagner, C. (2003), *An Introduction to Credit Risk Modelling*, Chapman & Hall, London.

Breiman, L. (2001), Random Forests, *Machine Learning*, **45**(1), 5-32

Dietterich, T. (1997), Machine Learning Research: Four Current Directions, *Al Magazine*, **18**, 97-136

Doumpos, M. and Pasiouras, F. (2004), *Developing and Testing Models for Replicating Credit Ratings: A Multicriteria Approach*, Financial Engineering Laboraty, Chania, GR.

Doumpos, M. and Zopounidis, C. (2002), *Multicriteria Decision Aid Classification Methods*. Kluwer Academic Publishers, Dordrecht.

Duffie, D. and Singleton, K.J. (2003), *Credit Risk: Pricing, Measurement and Management,* Princeton University Press, Princeton, NJ.

Florez, R. And Ramon, J. (2013), Modelling Credit Risk with Scarce Default Data: On the Suitability of Cooperative Bootstrapped Strategies for Small Low-Default Portfolios, *Journal of the Operational Research Society*, **65**, 416-434

He, Z. and Xiong, W. (2012), Rollover Risk and Credit Risk, *The Journal of Finance*, **67**, 391-429.

Hosmer, D.; Hosmer, T.; Le Cessie, S. And Lemeshow, S. (1997). A Comparison of Goodness-Of-Fit Tests For The Logistics Regression Model. *John Wiley and SONS, Inc.*

Hosmer, D. and Lemeshow, S. (1989). Applied Logistic Regression. *John Wiley and SONS, Inc.*

Krahnen, J. and M. Weber (2001), Generally Accepted Rating Principles: A Primer, *Journal of Banking and Finance*, **25**, 3-23.

"Logistic Regression: Statnotes." *North Carolina State University, Public Administration Program*. Statistics Associate Publishers. Web. <http://faculty.chass.ncsu.edu/garson/PA765/logistic.htm>.

Merton, Robert C. (1974) "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *The Journal of Finance* 29.2: 449-70. Print.

Modigliani, F. and Miller, M. (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *American Economic Review* **48** (3): 261–297.

Moody's Investors Service (1996), *The Biomial Expansion Method applied to CBO|CLO*, December.

Murcia, F.C. de S. (2013). "Rating De Crédito No Brasil: Fatores Determinantes E Impacto De Anúncios Nos Preços Das Ações." *Pós-Graduação Em Em Administração*. *Repositório UFSC*. Web. <https://repositorio.ufsc.br/bitstream/handle/123456789/107160/318126.pdf?sequence=1>.

N. Gujaratti, Damodar, and Dawn C. Porter. *Econometria Básica*. Quinta Edição ed. Vol. Único. New York: Mc Graw Hill, 2011. 924. Print.

Rosch, D. and Scheule, H. (2012), Forecasting Probabilities of Default and Loss Rates Given Default in the Presence of Selection, *Journal of the Operational Research Society*, **65**, 393-407.

Schonbucher, P.J. (2003), *Credit Derivatives Pricing Models: Models, Pricing and Implementation*, John Wiley & Sons Ltd, Chichester.

Treacy, W.F. and Carey, M. (2000), Credit risk rating systems at large US banks, *Journal of Banking and Finance* **24**, 167-201.

Westgaard, S. and Wijst, N. V. D. (2001), Default probabilities in a corporate bank portfolio: a logistic approach, *European Journal of Operational Research*, **135**, 338-349.

Woo, W.H. and Siu, T.K. (2003*), A dynamic binomial expansion technique for credit risk measurement: a Bayesian filtering approach*, Liu Bie Ju Centre for Mathematical Sciences, Kowloon, HK.

Wooldridge, Jeffrey M. "Limited Dependent Variable Models and Sample Selection Corrections." *Introductory Econometrics*. 4th ed. Mason: South-Western CENGAGE Learning, 2009. 865. Print.

Zopounidis, C. and Doumpos, M. (1999), A multicreteria decision aid methodology for sorting decision problems: The case of financial distress, *Computational Economics*, **14**, 197-218.