
Um Modelo de Previsão de Solvência Utilizando Regressão Logística

João Alberto Minussi
Cláudio Damacena
Walter Lee Ness Jr.

RESUMO

O processo de entrada de novos bancos estrangeiros no mercado brasileiro, aliado à estabilidade monetária do país, está requerendo do sistema financeiro uma mudança em seu perfil de atuação, principalmente na área de crédito. Neste sentido, este artigo representa uma importante contribuição ao apresentar os resultados do teste e comprovação de uma nova técnica (regressão logística) para avaliar o risco de crédito. Esta ferramenta estatística se mostrou mais robusta em relação a outras técnicas utilizadas em trabalhos desta natureza. Ademais, as características da amostra (tamanho, perfil e origem) consubstanciam-se como diferenciais deste estudo em relação aos anteriores, pois foram utilizados 323 clientes de uma instituição financeira, identificados como empresas do setor industrial. Em função da revisão da literatura, foram selecionados 49 indicadores financeiros para a análise de solvência. Através da aplicação da análise estatística, conhecida como regressão logística, foi obtido o modelo econométrico de previsão de solvência, composto por 5 variáveis. A precisão deste modelo foi bastante alta, pois 94,85% das empresas foram classificadas corretamente. A validação do modelo foi realizada por meio do método conhecido como *cross-validation*, ou seja, a subdivisão da amostra original: uma para a definição do modelo e outra para a sua validação.

Palavras-chaves: bancos; crédito; regressão logística.

ABSTRACT

The process of entry of foreign banks in the Brazilian financial system, together with the achievement of monetary stability, are requiring changes in the profile of action of Brazilian banks, especially in the area of credit. In this sense, this article represents an important contribution in testing a new technique (logistic regression) in evaluating credit risk. The statistical technique used shows itself more robust than other methods that have been used in previous studies of this nature. Additionally, characteristics of the sample used (size, profile, and source) differentiate this study from others. 323 clients from the industrial sector of a financial institution were used. Through the application of the statistical technique known as logistic regression, an econometric model of solvency prediction was developed that includes five variables. The precision of the model was relatively high, in that 94.85% of the sample was classified correctly. The validation of the model was performed by the method known as *cross-validation*, dividing the original sample in two: one part for constructing the model and the other for validating the model.

Key words: banking; credit; logistic regression.

INTRODUÇÃO

As questões norteadoras dos estudos em finanças, usualmente, contemplam o binômio risco e retorno. Segundo Bernstein (1997), a idéia revolucionária que define a fronteira entre os tempos modernos e o passado é o domínio do risco: a noção de que o futuro é mais do que um capricho dos deuses e de que os homens não são passivos ante a natureza. A capacidade de prever o que poderá acontecer no futuro e de optar entre as várias alternativas é fundamental às sociedades contemporâneas. A administração do risco é o guia que conduz a uma ampla gama de tomada de decisões. A forma de administrar o risco e a vontade de com ele fazer ou não opções ousadas são elementos-chaves da energia que impulsiona o sistema econômico. Não são poucas as dificuldades e controvérsias que envolvem a mensuração de risco e retorno. Esta questão assume maior relevância no caso de uma instituição financeira, ao se considerar que os reflexos de uma correta mensuração dos níveis de risco sobre uma carteira de ativos pode representar um diferencial competitivo. Por outro lado, o insucesso de um modelo de avaliação de risco acarreta efeitos danosos à instituição.

A análise sobre o nível de risco de crédito de determinada empresa ou investimento requer a definição de critérios cuidadosos que possam indicar a possibilidade de inadimplência do tomador de um crédito ou do lançador de um título, remetendo para mecanismos de avaliação da sua saúde financeira. Neste momento, estará em evidência a necessidade de prever a probabilidade de o pagamento ocorrer (Caouette, Altman e Narayanan, 1999).

Ao aplicar um modelo de previsão de solvência trabalha-se com a hipótese de que os problemas de uma empresa podem ser identificados antecipadamente, pois sua situação financeira tende a sofrer um processo de deterioração progressiva, até atingir o estágio de concordata ou falência. A eficácia de um modelo de previsão pode representar a própria reversão da insolvência, à medida que é admitida uma intervenção e são tomadas as decisões necessárias visando à correção do problema. Contribui, também, para um melhor equilíbrio no sistema de alocação de recursos do mercado.

Neste sentido, este artigo demonstra a importância da aplicação de um modelo econométrico no processo de deferimento de crédito para uma instituição financeira. Para a consecução deste objetivo, concorrem fatores inovadores que diferenciam este estudo dos anteriores. A diferenciação inicia no campo metodológico, ao aplicar a técnica de regressão logística, transita pelo tamanho,

origem e perfil da amostra e avança em direção à comprovação de que é recomendável incorporar novos indicadores econômico-financeiros, pois os índices tradicionais de balanços, pelo menos para concessão de crédito, não se mostraram tão representativos.

REFERENCIAL TEÓRICO

O Crédito e o Uso de Modelos Econométricos

Os sistemas de crédito são normalmente resultantes de atitudes, respostas e padrões comportamentais que derivam do nível estratégico organizacional e, muitas vezes, das concepções ou das idéias do principal executivo responsável pela área. A filosofia da organização, a sua tradição e os padrões existentes são fatores adicionais incorporados. Assim, o comportamento do crédito tem seu próprio ciclo e suas peculiaridades, flutuando desde um conservadorismo defensivo até uma agressividade responsável.

Para Caouette, Altman e Narayanan (1998) crédito é a “expectativa de uma importância em dinheiro, num prazo definido”; o risco de crédito, portanto, será a possibilidade de que esta expectativa não se cumpra.

Referenciadas em probabilidades surgiram, no mercado financeiro brasileiro, as técnicas de análise matemática/estatística. Embora ainda estejam distantes de um processo consolidado, são utilizadas como auxiliares e, em muitos casos, como determinantes na decisão de crédito (Minussi, 2001). No seu aparecimento era bastante compreensível que o valor dessas técnicas como instrumento de decisão fosse protelado, devido à enorme quantidade de cálculos exigida para se obterem resultados consistentes. Seu uso prático só foi possível com o desenvolvimento da informática. Conforme Caouette, Altman e Narayanan (1998), as técnicas mais utilizadas no campo econométrico são a análise discriminante linear e múltipla, análise *logit* e análise *probit*. O modelo *logit* assume que a probabilidade cumulativa de perda de um crédito esteja situada entre 0 e 1, e que a probabilidade de perda seja logisticamente distribuída. A análise *probit* é semelhante ao modelo *logit*; porém assume que a probabilidade de perda tenha uma distribuição normal. Essas três técnicas modelam a probabilidade de inadimplência ou o prêmio de inadimplência, como variável dependente, cuja variância é explicada por um conjunto de variáveis independentes. Entre as variáveis independentes estão razões financeiras e outros indicadores, bem como as variáveis externas que mensuram as condições econômicas.

Outras técnicas muito empregadas na avaliação de crédito são as Redes Neurais e os Sistemas Especialistas (Caouette, Altman e Narayanan, 1998). As Redes Neurais constituem sistemas computacionais empregados para tentar imitar o funcionamento do cérebro humano por meio da emulação de uma rede de neurônios interligados que usam os mesmos dados empregados nas técnicas econométricas, mas chegam a um modelo de decisão por meio de implementações alternativas de um método de tentativa e erro. Os Sistemas Especialistas são sistemas computacionais de apoio à tomada de decisões, baseados em julgamentos inferenciais e dedutivos a respeito de um crédito (Caouette, Altman e Narayanan, 1998).

Pelas características descritas, algumas ilações positivas podem ser extraídas da aplicação dos modelos econométricos: tendem a reduzir os índices de inadimplência, ao restringir a subjetividade que influencia a decisão; contribuem para o aprimoramento da função do analista, afastando-o de seus respectivos vícios; evitam opiniões e julgamentos pessoais; agilizam o processo de concessão do crédito e diminuem os custos administrativos em várias áreas das organizações (Minussi, 2001).

Estudos sobre Inadimplência

Uma síntese dos estudos realizados nessa área foi relatada por Silva (1997). Segundo o autor, o primeiro estudo a respeito de inadimplência teria sido realizado por Paul J. Fitz Patrick, em 1932, comparando 19 empresas falidas na década de 20, com outras 19 empresas bem-sucedidas. Muitos outros estudos foram elaborados e a confecção de modelo de previsão de inadimplência obteve um salto qualitativo com o trabalho de Altman, em 1968. Ele construiu seu modelo, denominado *Z-Score*, a partir do uso dos instrumentos estatísticos mais sofisticados para a época, como a análise discriminante múltipla. Nesse estudo, foi utilizada uma amostra de 66 empresas (33 solventes e 33 insolventes). O percentual de acerto no desenvolvimento do modelo atingiu 95%; no entanto, na amostra de validação, sua precisão caiu para 91%. Esses percentuais indicam a proporção de empresas que foram corretamente classificadas como solventes ou insolventes por meio da equação discriminante identificada pelo referido modelo. A função discriminante final identificada por Altman foi composta pelas seguintes variáveis (vide Anexo 1, Lista de Abreviaturas):

- . $(AC - PC)/AT$ (sinal +).
- . Lucros Retidos/ AT (+).
- . Lucros antes dos juros e impostos/ AT (+).

- . Valor de mercado das ações/Exigível total (+).
- . Vendas/AT (+).

Altman, Haldeman e Narayanan (apud Caouette, Altman e Narayanan, 1998) apresentaram um novo modelo denominado ZETA, que não demonstrou ser mais eficaz em relação ao anterior de Altman, considerando os índices de precisão.

Hekanaho et al. (1998) utilizaram uma base de dados chamada *Compustat* para estabelecerem uma interessante comparação entre os métodos *Logistic Regression Analysis*, *Neural Networks* e *Rule-Based Learning*. Embora as técnicas, ao atingirem o máximo de 77,7%, não tenham registrado percentuais de acertos significativos, os autores conseguiram comprovar que a atualização das informações e a expressividade da amostra são dois quesitos importantes na eficácia dos modelos, que, em termos de precisão, demonstraram-se equivalentes entre si.

No Brasil, alguns estudos foram realizados a partir da década de 70, entre os quais se destacam: Altman, Baydia e Dias (1977), Kanitz (1978), Elizabetsky (apud Silva, 1983), Matias (apud Silva, 1983), Silva (1983), Kasznar (1986) e Santos (1996). Estes trabalhos seguiram a linha precursora de Altman, aplicando o método estatístico de análise discriminante. O trabalho realizado por Silva (1983) foi o que utilizou a amostragem mais representativa. Neste estudo, o autor apresenta uma classificação de empresas com vistas à concessão de crédito. Os indicadores selecionados no modelo foram:

- . Duplicatas Descontadas/Duplicatas a Receber(sinal -).
- . Estoques/(CMV)(+).
- . Fornecedores/Vendas(-).
- . Estoque Médio/CMV x 360(-).
- . Lucro Operacional + Despesa Financeira/(Ativo Médio – Investimentos Médios)(+).
- . Capital de Terceiros/(Lucro Líquido + 0,1 Imobilizado Médio – Saldo da Correção Monetária)(-).

A amostra foi composta por 77 empresas industriais e o percentual de acerto global foi de 89,61%.

METODOLOGIA

Amostra Pesquisada

A população de pesquisa foi constituída por clientes de determinado banco, por ele classificados como empresas do setor industrial. Por questões estratégicas e de sigilo bancário não podem ser divulgados os nomes dos clientes e da instituição financeira escolhida para a realização deste trabalho. A amostragem utilizada foi a não probabilística, por julgamento, isto é, de cada grupo de empresas, solventes e insolventes, foram selecionadas aquelas que possuíam os maiores limites de crédito e as maiores inadimplências e/ou pendências, respectivamente⁽¹⁾.

A amostra inicial era composta de 400 empresas, 200 solventes e 200 insolventes. No entanto, devido à ausência de informações contábeis foram excluídas 77 empresas⁽²⁾. Dessa forma, a amostra utilizada no estudo foi de 323 empresas. O grupo de solventes ficou composto por 168 empresas industriais, ordenadas por possuírem os maiores limites de crédito e que estavam operando normalmente com o banco em abril de 2000. Já o grupo das insolventes ficou formado por 155 empresas industriais, que possuíam pendências ou se encontravam inadimplentes, naquela data.

Seleção dos Indicadores

A seleção dos 49 indicadores embasou-se, inicialmente, na consideração daqueles índices tradicionalmente utilizados em análise de balanços, que indicam a estrutura de capital, alavancagem financeira, lucratividade e liquidez. Importante contribuição foi dada pelos cocientes oriundos do balanço reclassificado, conforme estudo elaborado por Fleuriet, Kehdy e Blanc (1980) sobre a análise dinâmica de capital de giro. A experiência dos autores na área também foi fator decisivo na seleção dos indicadores, pois foram contemplados índices e estabelecidas relações entre contas ou grupos de contas, que ainda não integram a literatura de análise de balanços.

Antes de iniciar o processo de modelagem, uma matriz de correlação entre os 49 indicadores foi visualmente analisada e os pares de variáveis que estavam altamente correlacionados foram assinalados. Não é desejável que exista alta correlação entre as variáveis independentes, pois o modelo de regressão é sensível à colinearidade entre as variáveis (Hair et al., 1998). A inclusão de variáveis altamente correlacionadas pode ocasionar estimativas extremamente exageradas dos coeficientes de regressão (Hosmer e Lemeshow, 1989). Desta forma, quatro indicadores foram excluídos, tendo em vista a sua alta correlação verificada com outros indicadores (Altman apud Caouette, Altman e Narayanan, 1998). A relação dos indicadores selecionados está listada no Anexo 2. Foram necessários alguns

ajustes nos cálculos dos indicadores, para evitar ou minimizar distorções, resultantes de relações matemáticas. Estas modificações seguiram as regras estabelecidas por Silva (1983).

Técnicas Utilizadas para Classificação e Discriminação de Grupos

Em muitas situações práticas, pesquisadores desejam separar duas classes de objetos ou alocar um novo objeto em uma dessas classes. De todas as técnicas existentes para esta finalidade, a **Análise Discriminante** e a **Regressão Logística** são as mais encontradas na literatura. Ambas as técnicas se enquadram na classe de métodos estatísticos multivariados de dependência, pois relacionam um conjunto de variáveis independentes com uma variável dependente categórica (Sharma, 1996; Hair et al., 1998; Morgan e Griego, 1998).

De forma geral, as técnicas de discriminação procuram encontrar uma função ou conjunto de funções que discrimine os grupos definidos pela variável categórica, visando a minimizar erros de classificação. Em contexto no qual o conjunto de variáveis independentes possui comportamento probabilístico de normalidade multivariada, a análise discriminante é adequada, porque minimiza os erros de classificação (Sharma, 1996; Hair et al., 1998). Portanto a suposição de normalidade multivariada é necessária para que os resultados da análise discriminante sejam satisfatórios.

Segundo Hosmer e Lemeshow (1989), a técnica de regressão logística tornou-se um método padrão de análise de regressão para variáveis medidas de forma dicotômica, especialmente nas áreas das ciências da saúde. O mesmo modelo pode ser utilizado com enfoque discriminatório, conforme descrevem Krzanowski (1988) e McLachlan (1992). Esses autores argumentam que o modelo logístico de discriminação pode ser utilizado de forma bem mais geral, pois não faz suposições quanto à forma funcional das variáveis independentes, e o número de parâmetros envolvidos no processo de estimação provavelmente será menor.

Comparando as duas técnicas, Krzanowski (1988) diz que é consenso geral que a discriminação logística deve ser preferida, quando as distribuições são claramente não-normais. A mesma afirmação é sustentada por Press e Wilson (1978). Hair et al. (1998) apontam uma lista de motivos que levariam o pesquisador a optar pela regressão logística:

- . não é necessário supor normalidade multivariada;
- . é uma técnica mais genérica e mais robusta, pois sua aplicação é apropriada em grande variedade de situações;

. é uma técnica similar à regressão linear múltipla.

Na regressão logística, a probabilidade de ocorrência de um evento pode ser estimada diretamente. No caso da variável dependente Y assumir apenas dois possíveis estados (1 ou 0) e haver um conjunto de p variáveis independentes $X_1, X_2, \dots, X_p$, o modelo de regressão logística pode ser escrito da seguinte forma:

$$P(Y = 1) = \frac{1}{1 + e^{-g(x)}}$$

onde,

$$g(x) = B_0 + B_1 X_1 + \dots + B_p X_p$$

Os coeficientes $B_0, B_1, \dots, B_p$ são estimados a partir do conjunto de dados, pelo método da máxima verossimilhança, que encontra uma combinação de coeficientes que maximiza a probabilidade de a amostra ter sido observada (Hosmer e Lemeshow, 1989). Considerando certa combinação de coeficientes $B_0, B_1, \dots, B_p$ e variando os valores de X , observa-se que a curva logística tem comportamento probabilístico no formato da letra S , o que é característica da regressão logística. Esse formato dá à regressão logística alto grau de generalidade, aliada a aspectos muito desejáveis:

a) Quando $g(x) \rightarrow +\infty$, então $P(Y = 1) \rightarrow 1$

b) Quando $g(x) \rightarrow -\infty$, então $P(Y = 1) \rightarrow 0$

Assim como podemos estimar diretamente a probabilidade de ocorrência de um evento, podemos estimar a probabilidade de não ocorrência por diferença:

$$P(Y = 0) = 1 - P(Y = 1)$$

Ao utilizarmos a regressão logística, a principal suposição é a de que o logaritmo da razão entre as probabilidades de ocorrência e não ocorrência do evento é linear:

$$\frac{P(Y = 1)}{P(Y = 0)} = e^{B_0 + B_1 X_1 + \dots + B_p X_p}$$

e, por consequência,

$$\ln \left[\frac{P(Y = 1)}{P(Y = 0)} \right] = B_0 + B_1 X_1 + \dots + B_p X_p$$

Por essa razão, ao interpretar os coeficientes da regressão logística, opta-se pela interpretação de e^B e não diretamente de B . Contudo, quando se utiliza o modelo logístico do ponto de vista de discriminação entre grupos, não há grande interesse na interpretação dos coeficientes (Garson, 2000).

Para utilizar o modelo de regressão logística para discriminação de dois grupos, a regra de classificação é a seguinte:

- . se $P(Y=1) > 0,5$ então classifica-se $Y=1$;
- . em caso contrário classifica-se $Y=0$.

Portanto o modelo logístico foi utilizado para estimar, a partir do conhecimento de uma série de indicadores, a probabilidade de a empresa ser enquadrada no grupo das solventes. A variável dependente (Y) indica se a empresa é solvente ($=1$) ou insolvente ($=0$) e a série de indicadores ($X_1, \dots, X_p$) constitui o conjunto de variáveis independentes.

Para obter-se boa estimativa da eficiência classificatória do modelo, a amostra foi separada em duas partes: uma utilizada para estimação do modelo, e outra para testar a eficiência da classificação - *holdout sample* (Hair et al., 1998). A amostra utilizada para estimação, também chamada de amostra de treinamento, contou com 194 casos. Esse número de casos corresponde a, aproximadamente, 60% do conjunto total de empresas, conforme sugerido por Hair et al. (1998). O processo de escolha foi realizado por meio do gerador de números aleatórios do *software* SPSS. Utilizando 194 casos, eliminando as variáveis altamente correlacionadas e optando pelo método *stepwise*, que busca a melhor combinação de variáveis independentes, foi realizado o processo de estimação dos coeficientes do modelo logístico.

ANÁLISE DOS DADOS

O modelo final estimado por meio da regressão logística foi o seguinte:

$$P(\text{Empresa} = \text{Solvente}) = \frac{1}{1 + e^{-g(x)}}$$

$$g(x) = 4,4728 - 1,659X_1 - 1,2182X_2 + 4,1434X_3 + 6,1519X_4 - 1,885X_5$$

onde,

X1 = (PC/PL)/MEDIANA DO SETOR

X2 = IOG/VL

X3 = T/VL

X4 = ESTOQUES/CMV

X5 = OBRIGAÇÕES TRIBUTÁRIAS E PREVIDENCIÁRIAS/VMM

Tabela 1: Estimativas dos Coeficientes Beta e Testes de Significância

Variável	B	S.E.	Wald	df	Sig	R	Exp(B)
X1	-1,659	0,4411	14,142	1	0,0002	-0,2126	0,1903
X2	-1,2182	0,4162	8,5667	1	0,0034	-0,1564	0,2957
X3	4,1434	0,9154	20,4882	1	0,0000	0,2624	63,0149
X4	6,1519	1,279	23,1338	1	0,0000	0,2805	469,6059
X5	-1,885	0,4729	15,8908	1	0,0001	-0,2274	0,1518
Constant	4,4728	0,7765	33,1785	1	0,0000		

O modelo foi altamente significativo (χ^2 ; 5gl = 215,9; Significância = 0,0000) e a medida de *Nagelkerke* foi de ,887, quando o máximo possível de ser atingido é +1. Isso indica grande poder de previsão do modelo, como pode ser verificado na Tabela 2.

Tabela 2: Classificação das Empresas pelo Modelo Econométrico Obtido

Grupo observado	Grupo predito		Total	% de Acerto
	Insolventes	Solventes		
Insolventes	86	7	93	92,47
Solventes	3	98	101	97,03
Total	89	105	194	94,85

Pode-se estimar, portanto, que o percentual máximo de acerto do modelo logístico anteriormente apresentado é de 94,85%. Para ter-se melhor estimativa da eficiência classificatória do modelo, foi utilizado o modelo para classificar as 129 empresas que compõem a amostra de teste (*holdout sample*). A Tabela 3 apresenta a classificação obtida.

**Tabela 3: Tabela de Classificação Obtida na Amostra de Teste
(Holdout Sample)**

Grupo observado	Grupo predito		Total	% de acerto
	Insolventes	Solventes		
Insolventes	60	2	62	96,77
Solventes	0	67	67	100,00
Total	60	69	129	98,45

Deve-se mencionar que este trabalho seguiu rigorosamente as recomendações de Hair et al. (1998) de que é essencial realizar uma divisão da amostra para validação da pesquisa, a fim de evitar qualquer tipo de tendenciosidade. Os autores observam que esta divisão não pode comprometer quantitativamente as amostras, quando relacionadas ao número de variáveis independentes utilizadas nos estudos. A observância destas indicações confere a este trabalho maior robustez em relação a outros estudos que trabalharam com amostras menores e/ou com um número de indicadores inferior. Entende-se que a eficácia de um modelo de crédito somente será comprovada após ampla e minuciosa validação. Conforme pode ser visto na Tabela 3, os resultados apresentados na amostra de validação (precisão de 98,45%) atestam excelente capacidade de previsão do modelo.

No intuito de estabelecer comparações, utilizaremos os dois principais trabalhos referidos neste artigo (Silva, 1983; Altman apud Caouette, Altman e Narayanan, 1998). Altman (apud Caouette, Altman e Narayanan, 1998) reconhece que resultados baseados em amostragem de desenvolvimento acarretam distorções, requerendo, portanto, uma amostra secundária. No entanto os autores utilizam em seu trabalho a mesma amostra, apenas efetuando testes de cinco réplicas do método sugerido, estabelecendo subconjuntos da amostragem original. O resultado médio destes cinco testes atingiu o percentual de precisão de 93,5%. Por sua vez, Silva (1983) não realizou a separação da amostra, para testagem do modelo, definindo-o apenas com a amostra de treinamento. O percentual de acerto do modelo denominado Modelo Pereira para Empresas Industriais foi de 89,61%.

Considera-se que outro item importante na avaliação destes trabalhos, está focado no exame das variáveis selecionadas em cada modelo. O modelo de Altman apresenta uma característica interessante. Dos cinco índices que compõem a equação, quatro retratam a relação de contas ou grupo de contas com o ativo total no denominador. O modelo claramente valoriza variáveis de rentabilidade e apresenta, em segundo plano, um equilíbrio entre indicadores de liquidez e estrutura financeira. Silva (1983) afirma que sua abordagem se volta para os aspectos de crescimento das empresas, para a eficiência da administração dos estoques e de duplicatas a receber, para a capacidade de geração de recursos da empresa e para sua rentabilidade operacional. A afirmação do autor pode ser confirmada

perfeitamente nas variáveis selecionadas em seu modelo, onde se destacam três índices de liquidez, sendo que em dois deles consta a conta estoques no numerador.

O perfil dos indicadores selecionados na função matemática deste estudo difere sobremaneira dos trabalhos anteriormente referidos. Consta-se a forte influência de variáveis integrantes do chamado Modelo Fleuriet (descreve um modelo de análise dinâmica da situação financeira do capital de giro das empresas, reclassificando contas de acordo com a natureza das operações), pois das seis variáveis da equação, duas (x_2 e x_3) representam o cerne deste modelo (Fleuriet, Kehdy e Blanc, 1980). Em relação aos demais índices, deve ser destacada a inclusão de um cociente de liquidez (X_5) relacionada com obrigações tributárias e previdenciárias que, até o momento, não consta na literatura de análise de balanços. Em relação a X_4 e X_1 , estes são índices oriundos da chamada análise tradicional de balanços. O primeiro é um indicador de estrutura financeira, em cuja apuração se utilizou a relação com a respectiva mediana setorial. O segundo é cociente de liquidez.

Contribuição de cada Variável para o Resultado

Hair et al. (1998) e Garson (2000), ao descreverem as principais características e suposições da regressão logística, afirmam que a medida de Wald é a mais indicada para se aferir a contribuição para o resultado de cada variável independente, exceto para situações em que os coeficientes sejam extremamente altos. Nesses casos, recomendam como um método alternativo para assegurar a importância relativa de cada variável independente, a utilização da medida 'R', que é uma função de *Wald statistic* e o número do grau de liberdade da variável. Neste estudo, as duas medidas apresentaram a mesma ordem de grandeza, ao definir a contribuição de cada variável. De acordo com este *ranking* as variáveis da equação são analisadas a seguir.

. Estoques/CMV- Variável X_4 .

Médias: grupo solvente= 0,20; grupo insolvente= 0,40.

O setor industrial tem sido objeto das maiores alterações no que tange à cadeia produtiva, incorporando vários métodos logísticos que invariavelmente conduzem, cada vez mais, à redução ou eliminação dos estoques. Estes devem ser vistos hoje como uma aplicação de recursos. Diante dos altos custos financeiros no Brasil, este cociente de rotatividade, quando elevado, influencia sobremaneira a rentabilidade e, principalmente, a liquidez de uma empresa. Uma questão que requer muita atenção dos analistas é quando as empresas começam a ter dificuldades financeiras e apresentam estoques muito elevados. Isto pode

significar uma política estratégica norteada em premissas preditivas de mercado, mas de razoável grau de risco, ou então a formação de um estoque que permita à empresa produzir por um determinado tempo, caso ela se encontre diante de uma iminente insolvência. Quando esta situação é identificada pelo mercado, a oferta regular de crédito, normalmente, é afetada. Pelas questões apontadas, percebe-se a coerência da seleção deste índice como elemento discriminante, fato que se pode confirmar no desempenho tão assimétrico entre os grupos de empresas.

. Saldo de Tesouraria/Venda Líquida - Variável X3.

Médias: grupo solvente= -0,01; grupo insolvente= -2,90.

Este indicador é a essência do denominado Modelo Fleuriet, que descreve, de forma mais dinâmica do que a análise tradicional de balanços, a situação financeira de uma empresa. O saldo de tesouraria sinaliza a política financeira da empresa. Se for positivo, representa disponibilidade de recursos que garantem a liquidez de curtíssimo prazo do empreendimento. Entretanto, quando for negativo, é importante estabelecer uma relação com o nível de vendas, pois a expressividade do índice pode evidenciar dificuldades financeiras iminentes, em especial por ocasião da manutenção de saldos negativos de tesouraria sucessivos e crescentes. Os dados amostrais refletem claramente a maior dificuldade das empresas insolventes de financiarem-se operacionalmente, utilizando-se, então, mais intensamente de fontes erráticas.

. Obrigações Tributárias e Previdenciárias/Venda Média Mensal - Variável X5.

Médias: grupo solvente= 0,49; grupo insolvente= 3,20.

Embora ainda não tenha indicação bibliográfica, este cociente atesta uma medida de liquidez. A confusa e onerosa política fiscal brasileira, associada aos problemas conjunturais econômicos, está determinando o crescimento dos tributos em relação ao faturamento das empresas, cada vez mais comprometendo as margens operacionais. Esta situação tem sido um elemento indutor para o não reconhecimento contábil dessas obrigações, por parte de um número muito expressivo de empresas. Diante disso, também é verdadeiro que tem aumentado o número de autuações fiscais, que resultam num reconhecimento não espontâneo desses compromissos, levando as empresas a efetuarem parcelamentos dessas obrigações e, por consequência, aumentando sensivelmente os valores contabilizados nessa rubrica. Assim, este indicador ganhou importância para diferenciar empresas saudáveis daquelas problemáticas, conforme pode ser constatado nas médias dos grupos.

. (PC/PL)/Mediana do Setor - Variável X1.

Médias: grupo solvente= 0,80; grupo insolvente= 5,20.

Este é um indicador de estrutura financeira que representa o endividamento de curto prazo. A interpretação dos resultados apurados neste índice comprova que a maior utilização de capitais de terceiros por uma empresa, direciona-a para a insolvência. É muito comum no Brasil, que empresas financiem seus projetos com linhas de crédito que têm características e prazos incompatíveis com seus fluxos de caixa, levando-as, dessa forma, à utilização mais pronunciada de recursos de curto prazo. Assim, incorrem em maiores custos financeiros e, por consequência, afetam sua liquidez e rentabilidade. Esta variável não está utilizada em sua forma absoluta, mais em relação ao índice do setor no qual a empresa está situada.

. IOG/Venda Líquida - Variável X2.

Médias: grupo solvente= 0,80; grupo insolvente= 3,50.

O IOG, quando positivo, refere-se ao investimento líquido de curto prazo. Numa situação estática, representa aqueles recursos necessários à manutenção do atual nível da atividade operacional da empresa, via de regra obtidos por meio de financiamentos de fontes onerosas, de prazos equivalentes ou mesmo de longo prazo. Na avaliação de Braga e Veiga (1995), este elemento guarda proporcionalidade ao ciclo financeiro e ao volume de vendas praticado. Os resultados das amostras comprovam as maiores dificuldades das empresas insolventes para financiarem suas operações, por meio das fontes naturais ligadas a sua atividade.

CONCLUSÕES

Conforme explicitado no início deste artigo, as pesquisas bibliográficas e a realidade do mercado indicam constantes inovações nas técnicas para aferir o risco de mercado. Os métodos para avaliação dos riscos de crédito, entretanto, permanecem com aparência e procedimentos artesanais, não obstante os estudos divulgados por estudiosos da área, a partir do final da década de 60. Diante desta premissa, este artigo, já na sua concepção, buscou diferenciar-se dos trabalhos até então divulgados, no intuito de alcançar resultados que possam contribuir para a modernização do processo de avaliação do crédito. Dessa forma, escolheu-se uma instituição financeira, que possui um banco de dados mais amplo e consistente do que outras fontes passíveis de utilização. Isto representou uma

diferenciação no perfil e no tamanho da amostra comparativamente aos trabalhos anteriores. A partir dessa base de dados diferenciada, foram testadas as técnicas empregadas em outros estudos, não se obtendo resultados que pudessem ser considerados significativamente melhores. Tornava-se necessário, então, testar novas técnicas. Surge, assim, a regressão logística, como o expoente diferencial em relação aos outros estudos, situação comprovada no elevado nível de precisão do modelo. O sucesso do modelo também pode ser atribuído à seleção inovadora dos indicadores, à contemplação de um enfoque mais dinâmico do balanço, tendo como referência o Modelo Fleuriet, à incorporação de novos indicadores na literatura de análise de balanços e, ainda, à utilização de comparações de indicadores com medianas setoriais de um banco de dados que contém aproximadamente 10.000 empresas. Desta forma, é importante destacar duas informações que referendam sua diferenciação comparativamente aos demais estudos:

- . índice de precisão: 98,45%, visto que da amostra *holdout* de 129 empresas utilizadas para testar o modelo, 127 foram classificadas corretamente;
- . o perfil da equação: duas variáveis derivadas da análise dinâmica do capital de giro (T/VL e IOG/VL), um novo indicador (Obrigações Tributárias e Previdenciárias/VMM), um índice produzido pela relação com a mediana setorial ((PC/PL)/Mediana) e um cociente da análise tradicional de balanços (Estoques/CMV).

NOTAS

¹ Pendências: atraso nos pagamentos junto ao banco superior a 60 dias; inadimplência: valores inscritos em créditos em liquidação de janeiro de 1999 a abril de 2000.

² Balanços referentes ao exercício de 1999 ou 1998 não fornecidos ou incompletos.

REFERÊNCIAS BIBLIOGRÁFICAS

ALTMAN, E. I.;
BAYDIA, T;
RIBEIRO D. L.

Assessing potential financial problems for firms in Brazil.
Pontifícia Universidade Católica

do Rio de Janeiro, 1977.
Documento de Trabalho n. 8.

BERNSTEIN, P. L.
Desafio aos deuses. Rio de Janeiro: Campus, 1997.

CAOINETTE, J. B.;

ALTMAN, E. I.;

NARAYANAN, P.

Managing credit risk. New York: John Wiley & Sons, 1998.

FLEURIET, M.;

KEHDY, R.;

BLANC. G.

A dinâmica financeira das empresas brasileiras. Belo Horizonte: Fundação Dom Cabral, 1980.

GARSON, D. G.

PA 765 statnotes: an online textbook. Disponível em: <<http://www.z.chass.ncsu.edu/garson/pa765/logistic.htm>> Acesso em: 10 nov. 2000.

HAIR, J. F. et al.

Multivariate data analysis. 5. ed. New Jersey: Prentice-Hall, 1998.

HEKANAHU, J. et al.

Analysing bankruptcy with data with multiple methods. Turku: American Association for Artificial Intelligence, 1998.

HOSMER, D.;

LEMESHUW, S.

Applied logistic regression. New York: John Wiley & Sons, 1989.

IUDÍCIBUS, S.

Análise de balanços. São Paulo: Atlas, 1998.

KANITZ, S. C.

Como prever falências. 3. ed. São Paulo: McGraw-Hill, 1978.

KASZNAR, I. K.

Falências e concordatas de empresas: modelos teóricos e estudos empíricos – 1978-1982. [S.l.], 1986. Dissertação (Mestrado) - Fundação Getúlio Vargas.

KRZANOWSKY, W. J.

Principles of multivariate analysis. Oxford: Clarendon Press, 1988.

MARQUES J. A. V. C.;

BRAGA, R.

Análise dinâmica do capital de giro: o modelo de Fleuriet. **Revista de Administração de Empresas**, v. 35, n. 3, maio/jun. 1995.

MCLACHLAN, G.

Discriminant analysis and statistical pattern recognition. New York: John Wiley & Sons, 1992.

MINUSSI, J. A.

Um modelo preditivo de solvência utilizando regressão logística. São Leopoldo, 2001. Dissertação (Mestrado em Administração) - Centro de Ciências Econômicas, PUC-Rio/Unisinos.

MORGAN, G. A.;
GRIEGO, O. V.

Easy use and interpretation of SPSS for Windows: answering research questions with statistics. New Jersey: Lawrence Erlbaum, 1998.

PRESS, J.;
WILSON, S.

Choosing between logistic regression and discriminant analysis. **Journal of the American Statistical Association**, v. 73, n. 364, p. 699-705, 1978.

SANTOS, S. C. DOS.

Um modelo de análise discriminante múltipla para

previsão de inadimplência em empresa. Rio de Janeiro, 1996. Dissertação (Mestrado em Administração) - Departamento de Administração, Pontifícia Universidade Católica do Rio de Janeiro.

SHARMA, S.

Applied multivariate techniques. New York: John Wiley & Sons, 1996.

SILVA, J. P.

Administração de crédito e previsão de insolvência. São Paulo: Atlas, 1983.

Gestão e análise de risco de crédito. São Paulo: Atlas, 1997.

ANEXO 1: LISTA DE ABREVIATURAS

AC= ativo circulante

PC= passivo circulante

AT= ativo total

CMV= custo das mercadorias vendidas

PL= patrimônio líquido

IOG= investimento operacional em giro

T= saldo de tesouraria

VL= vendas líquidas

VMM= venda média mensal

ANEXO 2: RELAÇÃO DOS INDICADORES UTILIZADOS NO ESTUDO

1. LIQUIDEZ CORRENTE/MEDIANA SETOR
2. ENDIVIDAMENTO CURTO PRAZO/MEDIANA SETOR
3. ENDIVIDAMENTO/MEDIANA SETOR
4. IMOBILIZAÇÕES/MEDIANA SETOR
5. ENDIVIDAMENTO BANCÁRIO/MEDIANA SETOR
6. PRAZO MÉDIO ESTOCAGEM/MEDIANA SETOR
7. PRAZO MÉDIO RECEBIMENTO VENDAS/MEDIANA SETOR
8. PRAZO MÉDIO PAGAMENTO FORNECEDORES/MEDIANA SETOR
9. CICLO OPERACIONAL/MEDIANA SETOR
10. CICLO FINANCEIRO (EM DIAS)
11. GIRO ATIVO/MEDIANA SETOR
12. LUCRATIVIDADE OPERACIONAL/MEDIANA SETOR
13. RENTABILIDADE CAPITAL PRÓPRIO/MEDIANA SETOR
14. TAXA RETORNO INVESTIMENTO/MEDIANA SETOR
15. NÍVEL DESCONTO DUPLICATA/MEDIANA SETOR
16. CAPITAL CIRCULANTE LÍQUIDO/VENDA LÍQUIDA
17. INVESTIMENTO OPERACIONAL EM GIRO/VENDA LÍQUIDA
18. AUTO FINANCIAMENTO/VENDA LÍQUIDA
19. TESOURARIA/VENDA LÍQUIDA
20. INVEST. OPER. EM GIRO/RESULTADO OPER. BRUTO
21. LUCRO BRUTO OPERACIONAL/VENDA LÍQUIDA
22. LUCRO BRUTO OPERACIONAL/ATIVO OPERACIONAL
23. (ATIVO CIRCULANTE-PASSIVO CIRCULANTE)/ATIVO TOTAL
24. LUCROS RETIDOS/ATIVO TOTAL
25. LUCRO ANTES IR/ATIVO TOTAL
26. VENDAS/ATIVO TOTAL
27. LUCRO LÍQUIDO/VENDAS
28. CONTAS A RECEBER/ATIVO TOTAL
29. PASSIVO CIRCULANTE/ATIVO TOTAL
30. ESTOQUES/ATIVO TOTAL
31. PATRIMÔNIO LÍQUIDO/ATIVO TOTAL
32. LUCRO OPERACIONAL/LUCRO BRUTO

33. ESTOQUES/CUSTO MERCADORIA VENDIDA
34. FORNECEDORES/VENDAS
35. (LUCRO OPER. +DESPESA FINANCEIRA)/ATIVO OPER.
36. ESTOQUES/(VENDAS-LUCRO BRUTO)
37. CUSTO MERCADORIA VENDIDA/VENDA LÍQUIDA
38. RECEITA FINANCEIRA/DESPESA FINANCEIRA
39. LUCRO LÍQUIDO/DESPESA FINANCEIRA
40. DESPESA ADMINIST. E COM VENDA/VENDA LÍQUIDA
41. DESPESA FINANCEIRA/VENDA LÍQUIDA
42. (OBRIG. FIS. E TRIBUT. +PARCEL. TRIB.)/VENDA MENSAL
43. DESPESA FINANCEIRA/PASSIVO TOTAL
44. DESPESA FINANCEIRA/LUCRO BRUTO
45. DÍVIDAS BANCÁRIAS/ATIVO CIRCULANTE