

Quanto pior, melhor: Estudo da utilização da análise por envoltória de dados em modelos de análise de inadimplência/insolvência de empresas

When worst is better: Study of the application of data envelopment analysis in models of insolvency/default analysis

Silvia Pereira de Castro Casa Nova

Doutora em Controladoria e Contabilidade pela Universidade de São Paulo

Professora do Programa de Pós-Graduação em Controladoria e Contabilidade da Universidade de São Paulo

Endereço: Avenida Professor Luciano Gualberto, n 908 - Sala 214 - Prédio FEA3

CEP: 05508-010 – São Paulo/SP – Brasil

E-mail: silvianova@usp.br

Telefone: (11) 3091-5820 - Ramal 156

Artigo recebido em 04/05/2012. Revisado por pares em 01/10/2012. Reformulado em 20/12/2012. Recomendado para publicação em 08/02/2013 por Sandra Rolim Ensslin (Editora Científica). Publicado em 30/04/2013.

Resumo

O objetivo foi construir modelos de análise de inadimplência/insolvência aliando Análise por Envoltória de Dados (DEA) e informações contábeis. A pesquisa foi subdividida em etapas: exames de trabalhos anteriores com modelos DEA de previsão de inadimplência/insolvência; composição de amostras de diferentes bancos de dados com informações contábeis de empresas com problemas financeiros e saudáveis; e, proposta de metodologia de aplicação de DEA à análise da inadimplência/insolvência de empresas. Os resultados comparam-se aos obtidos dos modelos de Regressão Logística e Redes Neurais, porém sem indicação de superioridade de algum procedimento.

Palavras-chave: Insolvência. Análise por Envoltória de Dados. Análise de Balanços.

Abstract

This paper aims to develop an insolvency/failure analysis model allying accounting information with Data Envelopment Analysis (DEA). The research was conducted in phases: initially previous studies were examined; then samples were obtained from different databases containing financial information of distressed and healthy companies; and, finally a methodology for application of DEA to failure/insolvency analysis was proposed. The results were comparable to those obtained from logistic regression models and neural network models, without, however, a clear indication of the superiority of any procedure.

Keywords: Insolvency. Data Envelopment Analysis. Balance Sheet Analysis.

1 Introdução

O presente estudo faz parte de uma série de trabalhos dedicados a estudar e apresentar contribuições à avaliação do desempenho econômico das empresas por meio da Análise de Demonstrações Contábeis, associada à Análise por Envoltória de Dados (DEA). Trata-se de uma técnica desenvolvida em 1978 e, ainda hoje, com poucos trabalhos realizados no Brasil. Especificamente teve como objetivo construir, testar e comparar modelos de análise de inadimplência/insolvência aliando análise de balanços e DEA. Diferencia-se pela base de dados utilizada, composta por empresas de médio porte dos setores comercial e industrial.

A pesquisa está subdividida em etapas. Na primeira etapa, com base em levantamento bibliográfico, foram localizados, selecionados e examinados trabalhos que construíram modelos de previsão de inadimplência/insolvência com utilização de DEA. Essa etapa originou um comparativo de estudos que serviu de base para a construção do modelo proposto. A etapa seguinte foi obter uma base de dados com informações contábeis de empresas com problemas financeiros (inadimplentes/insolventes) e com empresas solventes/saudáveis, para um determinado período de tempo e, a partir dela, compor amostras para aplicação do modelo. A janela de tempo ideal seria de cinco anos antes da data da ocorrência do problema. Esse intervalo temporal permitiria um acompanhamento da evolução

(ou involução) da empresa no período, permitindo traçar tendências e verificar os indicadores contábeis que sofrem maior impacto pela deterioração da saúde de empresa. Mas estudos (CASA NOVA, 2002; SANTOS; CASA NOVA, 2005) já debateram as dificuldades de “obter informações de empresas para um período superior a três anos, em função de alterações societárias (aquisições, fusões e cisões) para empresas solventes/saudáveis e para empresas com problemas, da existência de informações divulgadas no período imediatamente anterior à sua ocorrência ou ao início dos problemas financeiros” (CASA NOVA, 2010). Para teste do modelo foi utilizada a amostra de estudo anterior (LAREDO, 1997). As amostras de aplicação foram obtidas das bases de dados da Serasa Experian e utilizadas na proposição de modelo estruturado de Análise por Envoltória de Dados a ser aplicado à análise do processo de inadimplência/insolvência de empresas. Suas vantagens e limitações serão exploradas, comparativamente aos resultados obtidos por outros modelos. Portanto, a questão que se coloca para nortear a pesquisa é: Quais as vantagens e limitações do desenvolvimento de um modelo de análise de inadimplência/insolvência baseado na utilização da Análise por Envoltória de Dados aliada à Análise de Balanços de empresas?

Uma das inovações do estudo foi a utilização da Análise Fatorial para agregação e redução do número de variáveis. O raciocínio que justificava o investimento é o reconhecimento dos efeitos do número de variáveis no poder de discriminação da DEA, já largamente discutido na literatura.

O artigo tem a seguinte estrutura: (1) nesta Introdução, foram apresentadas considerações iniciais sobre o tema, a questão de pesquisa e os objetivos do trabalho; (2) na segunda seção, foi feita a revisão bibliográfica do tema com a apresentação do estado da arte da aplicação da DEA à análise de inadimplência/insolvência, no Brasil e no exterior; (3) a terceira seção se dedicou a discorrer sobre a metodologia utilizada na pesquisa, apresentando as bases de dados e amostras estabelecidas e descrevendo as técnicas aplicadas; (4) a quarta seção demonstrou os resultados da aplicação dessa metodologia às amostras estabelecidas; (5) finalmente foram enfatizadas as conclusões do estudo e indicados caminhos para avanço da pesquisa.

2 Referencial Teórico

O estudo de insolvência/inadimplência de empresas com aplicação de Análise por Envoltória é ainda relativamente recente. Em geral, os trabalhos que se dedicaram ao tema fizeram-no comparativamente aos resultados de modelos mais “tradicionais”, utilizando técnicas como análise discriminante ou regressão logística. Chama atenção o aparente “contra-senso de utilizar uma ferramenta construída para analisar eficiência, no estudo e identificação de empresas com problemas. Mas é exatamente nesse ponto que se encontra o ‘ovo de Colombo’: identificar os melhores em serem os piores” (CASA NOVA, 2010). Assim, a lógica estaria em que, para as empresas com problemas, o indicador de eficiência seria sistematicamente menor do que para as empresas saudáveis, possibilitando, portanto, sua identificação a partir da construção de uma “fronteira de ineficiência”. Possibilitaria ainda identificar padrões de comportamento de empresas com problemas e encontrar indicadores de maior relevância para o diagnóstico.

Fernandez-Castro e Smith (1994), na busca de um modelo de análise do desempenho empresarial, apresentaram o primeiro estudo de DEA aplicado à insolvência. O modelo foi

implementado com a inclusão de seis indicadores financeiros e foi aplicado aos dados contábeis de 27 bases de dados, com o total de 1.423 empresas, sendo sete o número mínimo de empresas por base, e o máximo, 136. Para cada uma, incluía-se uma empresa que faliu no período de janeiro de 1985 a dezembro de 1989 e todas as demais empresas com o mesmo código SIC na base EXSTAT. Ou seja, cada uma das 27 bases de dados continha uma empresa falida e todos os seus pares, no mesmo setor, que sobreviveram até o final do período do estudo. Os dados contábeis eram do ano precedente à falência e incluíam indicadores contábeis como produto, sendo necessário que respeitassem o critério de quanto maior, melhor. Como as empresas operavam em condições de mercados similares, os insumos foram considerados imateriais e iguais para todas. Em suas conclusões, os autores afirmaram que o modelo poderia ser usado para explorar a performance de empresas na iminência de falir, com sucesso moderado em isolar as empresas falidas das sobreviventes e, ainda assim, se empregado em conjunto com outra ferramenta de análise. Poderia ser utilizado para definir metas e como uma ferramenta normativa. Os pontos fracos do enfoque, segundo os autores, foram sua suscetibilidade aos erros (ou flutuações aleatórias), sua inabilidade para capturar considerações de longo prazo e em fazer julgamentos úteis nas observações não usuais (*outliers*). Foi sugerido que pesquisas futuras fossem aplicadas aos trabalhos intertemporais mais recentes em DEA para verificar se é possível traçar uma posição relativa da empresa no tempo (FERNANDEZ-CASTRO; SMITH, 1994, p. 247). Como pontos fortes do modelo, foram citadas a liberdade de suposições sobre a distribuição, a flexibilidade e a independência de qualquer tipo de forma funcional.

No estudo de Emel, Oral, Reisman e Yolalan (2003), foi proposta a aplicação de uma metodologia DEA em dados contábeis de 82 indústrias que compunham a carteira de crédito de um dos maiores bancos comerciais da Turquia. A análise DEA sintetizou o desempenho total das empresas em único escore financeiro de eficiência, batizado pelos autores de escore de credibilidade. Entre as vantagens da utilização da DEA, foram citadas: (1) a conversão de uma multiplicidade de insumos e produtos em um indicador singular de desempenho adimensional; (2) a capacidade de lidar com múltiplos insumos e múltiplos produtos; (3) a definição de uma fronteira composta pelas empresas de melhores práticas (fronteira eficiente) para um dado conjunto; e, (4) a mensuração do grau de eficiência relativa para as demais empresas (não eficientes). No entanto, os autores indicam que, apesar de os resultados do modelo DEA serem consistentes com as análises tradicionais, ofereceram indicações sobre como as empresas em má situação financeira poderiam melhorar sua eficiência financeira. O estudo apresentado considerou empresas de diversos setores. Na opinião dos autores, para desenvolver um modelo mais consistente, a amostra poderia se limitar a um setor específico. Deveria ter-se em mente, no entanto, que se perderia a possibilidade de comparar os escores de eficiência entre empresas de diferentes setores. Finalmente, a metodologia apresentada permitiria ao banco monitorar a exposição de sua carteira de crédito. Seria, ainda, outra ferramenta administrativa para auxiliar em decisões estratégicas e táticas. Finalmente, em países sujeitos à regulação ou ao controle governamental, a análise DEA poderia servir como instrumento de monitoramento das regras de conformidade, inclusive com propósitos consultivos.

Cielen, Peeters e Vanhoof (2004) compararam o desempenho de modelos de programação linear, DEA e *rule induction* (C5.0) em previsão de falência, concluindo que, em termos de acurácia, o modelo DEA superou os demais. O modelo DEA mostrou-se mais consistente e menos sensível à mudança dos fatores de custo. Outra característica analisada

pelos autores foi a facilidade de compreensão dos modelos. Em termos de estabilidade, igualmente o modelo DEA mostrou-se superior. Considerando a facilidade de comunicação dos resultados ao usuário final, no modelo DEA havia, de um lado, uma ligação direta entre as variáveis e os coeficientes utilizados, e, de outro, os resultados. A conclusão dos autores foi que, para uma amostra pequena com informação quantitativa apenas, o modelo DEA superou os demais em termos de acurácia, custo, desenvolvimento e compreensibilidade.

Paradi, Asmild e Simak (2004) apresentaram o conceito de fronteira DEA de pior prática (*worst practice DEA*), com objetivo de identificar as empresas de pior desempenho e posicioná-las na fronteira. Na opinião dos autores, o conceito era particularmente relevante para a aplicação que desenvolveram em avaliação de risco de crédito. O estudo apresentou ainda a técnica de camadas (*layering technique*) como alternativa ao enfoque do ponto de corte. Segundo os autores, os modelos DEA normais classificam as empresas na fronteira como unidades eficientes, e as “envolvidas” pela fronteira como ineficientes. O modelo DEA de piores práticas utilizaria a mesma formulação, mas, em vez de apontar as empresas de bom desempenho, seu objetivo seria identificar as empresas de desempenho ruim. Isso seria conseguido selecionando variáveis que refletissem falhas de desempenho, ou seja, a estratégia seria selecionar variáveis de produtos que refletissem má utilização dos recursos ou resultados indesejáveis como, por exemplo, capital de giro, reclamações em garantia, maus devedores etc. As conclusões foram que a introdução dos modelos de piores práticas em combinação com a técnica de camadas conduziu a taxas de acertos maiores e a resultados mais influenciados pela amostra do que o método do ponto de corte. O enfoque de camadas implicou ainda maior flexibilidade, permitindo incorporar considerações de atitude frente ao risco e às estratégias de preço baseados no risco.

Premachandra, Chen e Watson (2011) utilizaram um modelo DEA aditivo de supereficiência desenvolvendo um novo índice de avaliação baseado em duas fronteiras para prever o sucesso e a falência corporativos. Segundo os autores, esse enfoque, quando aplicado a uma amostra aleatória de 1.001 empresas, com 50 empresas falidas selecionadas da base de dados de falências de Altman e 901 empresas sadias, considerando o período de 1991-2004, em diversos setores. Os resultados indicaram que a DEA é relativamente frágil em prever as falências em comparação às empresas sadias, e o indicador proposto diminuiu essa fragilidade, dando ao decisor várias opções em relação aos níveis de precisão na predição de empresas falidas, não falidas e total.

Finalmente, Liu, Lu, Lu e Lin (*in press*) realizaram um levantamento sobre as aplicações da DEA, englobando os artigos indexados na *Web of Science*, de 1978 até agosto de 2010. Os resultados demonstraram que 63% dos artigos utilizaram dados empíricos, o percentual restante foram trabalhos puramente metodológicos. Os setores mais estudados, perfazendo 41% do total dos artigos com aplicações, foram: bancário, saúde, agricultura e agropecuária, transportes e educação.

No Brasil, Almeida e Milioni (2000) apresentaram um modelo de *credit scoring* baseado em DEA. A proposta foi construir fronteiras de solvência e insolvência a partir de tomadores de crédito que, no passado, se revelaram bons e maus pagadores. O modelo foi estruturado nas seguintes etapas: (1) Determinação das variáveis de *input* e *output*.; (2) Determinação da fronteira de insolvência: foi desenvolvido um modelo DEA de maximização para ser aplicado aos dados das empresas insolventes, previamente selecionadas pelos autores, e definir uma fronteira DEA superior. A eficiência igual a 1 foi o limite superior dessas firmas; (3) Determinação da fronteira de solvência: foi utilizado um modelo DEA de

minimização para, com base nas informações das empresas solventes, definir uma fronteira DEA inferior para as empresas solventes. A eficiência igual a 1 foi o limite inferior das empresas solventes; e, (4) Unificação dos escores: com a utilização simultânea das duas fronteiras envoltórias foram definidas quatro regiões.

Onusic (2004) propôs a utilização conjunta das técnicas de Análise por Envoltória de Dados e Regressão Logística como forma de minimizar os erros de classificação das empresas solventes e insolventes em uma determinada base de dados. Para isso foram aplicados modelos de previsão de insolvência utilizando Regressão Logística (RL) e a Análise por Envoltória de Dados em uma base de dados fornecida pela Serasa relativa aos períodos de 1995 a 2001, composta por 122 empresas (61 solventes e 61 insolventes, pareadas em porte e setor às empresas solventes).

Onusic, Casa Nova e Viana (2004) apresentaram uma comparação entre os resultados da utilização de DEA e da regressão logística em modelos de previsão de insolvência. Concluíram que o modelo DEA foi capaz de discriminar com maior grau de acerto as empresas insolventes. Assim, segundo as autoras, “uma vez que o custo de classificação incorreta de empresas insolventes é maior que o das empresas solventes, o fato de o modelo DEA mostrar um desempenho para este grupo sugere um interesse de aplicação da técnica para a análise de insolvência” (2004, p. 30).

Casa Nova e Onusic (2005) apresentaram um mapeamento dos estudos sobre a utilização de DEA na análise de insolvência, com o objetivo de identificar o estágio do estudo, no Brasil e no exterior. Concluíram que o tema era bastante recente e ainda inexplorado, quando considerados apenas artigos publicados (não foram considerados trabalhos apresentados em encontros e congressos), constituindo-se em um tema promissor para pesquisas.

Onusic, Casa Nova e Almeida (2007) desenvolveram um modelo de previsão de insolvência utilizando DEA, para uma amostra de empresas da base Fipecafi-Exame, para o período de 1995-2003. A partir da seleção de 10 empresas insolventes, sortearam cinco empresas solventes entre as de melhor desempenho do setor, de porte semelhante, totalizando assim 50 empresas saudáveis. Obtiveram os dados contábeis das empresas para o período de três anos anteriores à falência ou à concordata. As classificações obtidas pela DEA foram confrontadas com a situação real da empresa após o período. Segundo as autoras, o modelo DEA foi capaz de discriminar, com bom grau de acerto, empresas solventes e insolventes (90% das empresas insolventes foram corretamente classificadas).

Macedo e Cavalcante (2009) utilizaram DEA para a análise da performance de agências bancárias. Os autores afirmam pretender aplicar “a metodologia DEA a indicadores de perspectivas balanceadas (pesos iguais) do *Balanced Scorecard (BSC)*, com o objetivo de definir unidades eficientes e ineficientes” (2009, p.88), englobando perspectivas de desempenho financeiro, operacional, de clientes e de pessoal. Concluem que a DEA foi capaz de analisar de forma multicriterial o desempenho das unidades (agências bancárias), consolidando as diversas perspectivas em um indicador único e propondo, dessa forma, uma nova percepção sobre as unidades de negócios. Como continuidade da discussão, em estudos futuros, sugerem que se utilize, por exemplo, a técnica Delphi para considerar a opinião dos gestores das agências na definição e seleção de variáveis.

Colauto, Pinheiro e Pinheiro (2009) desenvolvem um trabalho em que pretendem realizar estimativas que permitam estabelecer relações significativas entre os indicadores financeiros e a insolvência empresarial, considerando empresas espanholas presentes na Bolsa

e não pertencentes aos setores financeiro e de seguros. Os resultados encontrados indicam, segundo os autores, que vários indicadores contábeis conseguem discriminar as empresas entre solventes e insolventes com acerto de 95% para o ano imediatamente anterior.

Oliveira, Macedo e Corrar (2011) estudaram o desempenho dos oito maiores bancos nacionais do setor de varejo utilizando DEA. Os resultados mostraram que apenas uma das instituições financeiras foi eficiente nos três períodos sob análise (primeiro e segundo semestres de 2008 e primeiro semestre de 2009). Unidades que fizeram alterações societárias relevantes mantiveram ou melhoraram os índices de desempenho.

3 Procedimentos Metodológicos

O método de pesquisa utilizado constituiu-se de etapas. Após a revisão de literatura, foram buscadas amostras de teste para construção do modelo. Posteriormente, foram definidas as bases de dados a partir das quais se selecionaram as amostras para estudo. Com as amostras definidas, foram efetuadas análises descritivas dos indicadores e, se necessário, as transformações das variáveis para aplicação das técnicas estatísticas e matemáticas. Fez-se ainda Análise Fatorial (AF) por Componentes Principais dos indicadores. Para construção dos modelos foram utilizadas Análise Discriminante com Regressão Logística, DEA e Redes Neurais. Finalmente, os resultados de cada um dos modelos foram comparados.

Entre as dificuldades da definição das amostras de estudo estão: confiabilidade dos dados; disponibilidade dos dados para um número adequado de empresas; e, disponibilidade dos dados para um grupo de empresas em um período de tempo. Para contornar essas dificuldades, utilizou-se, inicialmente, a amostra de estudo anterior (LAREDO, 1997). Com base nessa amostra, foram rodados diversos testes e analisados os resultados preliminares.

Com os testes preliminares concluídos, buscou-se uma base de dados que atendessem aos pontos citados acima. Os problemas de confiabilidade e disponibilidade de dados para um número adequado de empresas foram resolvidos pela utilização das informações contidas nas bases de dados da Serasa Experian. As informações dessas bases são consideradas confiáveis, pois as informações são submetidas a uma série de exames de consistências e verificações por analistas antes de serem incluídas. Das bases de dados da Serasa Experian foram retiradas duas amostras, para indústria e comércio, com empresas de pequeno e médio porte. As amostras assim constituídas foram utilizadas para desenvolvimento dos modelos.

As informações sobre as amostras utilizadas no estudo estão apresentadas na Figura 1.

Figura 1 – Amostras utilizadas no estudo

Amostra	Descrição	Características	Utilização
LRD	Amostra de empresas selecionadas da base de dados da Serasa para o estudo de Laredo (1997), composta de 61 empresas sendo 25 empresas insolventes, 36 empresas solventes e 25 empresas avaliadas como excelentes pelos analistas da Serasa (critério de julgamento).	18 indicadores n ₁ =25 (insolventes) n ₂ =36 (solventes) n ₃ =25 (excelentes) 1996 e 1º. Sem./1997	Estudo e testes preliminares

Amostra	Descrição	Características	Utilização
SRS	Amostra de empresas fornecida pela Serasa Experian composta por 4.800 empresas, dos setores de indústria, comércio e serviços, de pequeno e médio porte.	14 indicadores <u>Indústria</u> n ₁ =116 n ₂ =143 <u>Comércio</u> n ₁ =166 n ₂ =194 2001 a 2003	Desenvolvimento

Fonte: Elaborado pela autora.

Para construção dos modelos com base em DEA, foi utilizado o algoritmo com retornos variáveis à escala (BANKER; CHARNES; COOPER, 1984), que permite, para as unidades avaliadas, que o acréscimo em uma unidade de insumo possa gerar um acréscimo não proporcional no volume de produtos. Os indicadores com valores negativos foram transformados somando uma constante (ALI; SEIFORD, 1990; PASTOR, 1996; KNOX LOVELL; PASTOR, 1995).

As variáveis foram classificadas como insumo ou produto, considerando seu efeito no escore de eficiência da empresa. Como regra, temos a seguinte classificação: para produtos, quanto maior, melhor; para insumos, quanto maior, pior. No entanto, vale ressaltar que nem sempre essa classificação é simples em se tratando de indicadores contábeis. Ela fica ainda mais difícil se considerados os indicadores agregados por meio da Análise Fatorial.

A partir do modelo DEA com a maior taxa de acerto global, foi construído um modelo DEA Inversa (também chamada de fronteira invertida) e calculado o escore de eficiência composta. O modelo DEA Inversa consiste na inversão dos *inputs* e *outputs* e na seleção da orientação oposta ao modelo original. Assim, se o modelo original considera orientação à minimização de *inputs*, o modelo DEA Inversa utilizará orientação à maximização de *outputs*. A eficiência composta é calculada considerando-se:

$$Eficiência\ Composta = \frac{Eficiência\ Normal + (1 - Eficiência\ Invertida)}{2},$$

Para determinar o ponto de corte nos modelos DEA, foi utilizado um método empírico (BARTH, 2004, p. 14-15), que busca maximizar a taxa de acerto global para a amostra.

Para cada uma das amostras, foi utilizado um conjunto de indicadores contábeis específico. Em cada amostra, foram estudados pontos extremos ou influentes e efetuadas as transformações de variáveis, quando necessário, considerando as técnicas específicas utilizadas.

Para a amostra LRD, os indicadores, suas fórmulas de cálculo, descrição e classificação para DEA são apresentadas na Figura 2.

Figura 2 – Indicadores contábeis utilizados – Estudo de Laredo (1997) – LRD

Indicador	Abrev.	Cálculo	Classificação DEA
Recursos Bancários sobre Ativo Circulante	CBAC	$\frac{\text{Recursos Bancários de Curto Prazo}}{\text{Ativo Circulante}}$	Insumo
Capital bancário sobre Ativo Total	CBAT	$\frac{\text{Recursos Bancários Totais}}{\text{Ativo Total}}$	Insumo

Indicador	Abrev.	Cálculo	Classificação DEA
Capital bancário sobre Passivo Circulante	CBPC	$\frac{\text{Recursos Bancários de Curto Prazo}}{\text{Passivo Circulante}}$	Insumo
Capital bancário sobre Patrimônio Líquido	CBPL	$\frac{\text{Recursos Bancários Totais}}{\text{Patrimônio Líquido}}$	Insumo
Exigível sobre o Patrimônio Líquido	EPL	$\frac{\text{Exigível Total}}{\text{Patrimônio Líquido}}$	Insumo
Giro do Ativo Total	GIAT	$\frac{\text{Vendas}}{\text{Ativo Total}}$	Produto
Giro de Estoques	IGE	$\frac{\text{Custos das Mercadorias Vendidas}}{\text{Estoque}}$	Produto
Liquidez Corrente	LC	$\frac{\text{Ativo Circulante}}{\text{Passivo Circulante}}$	Produto
Liquidez Geral	LG	$\frac{\text{A. Circulante} + \text{Real. a LP}}{\text{P. Circulante} + \text{Exig. a LP}}$	Produto
Liquidez Seca	LS	$\frac{\text{Ativo Circulante} - \text{Estoques}}{\text{Passivo Circulante}}$	Produto
Margem Bruta	MB	$\frac{\text{Lucro Bruto}}{\text{Vendas}}$	Produto
Margem Líquida	ML	$\frac{\text{Lucro Líquido}}{\text{Vendas}}$	Produto
Margem Operacional	MO	$\frac{\text{Lucro Operacional}}{\text{Vendas}}$	Produto
Prazo Médio de Cobrança	PMC	$\frac{\text{Duplicatas a Receber}}{\text{Vendas médias por dia}}$	Insumo
Prazo Médio de Pagamento	PMP	$\frac{\text{Fornecedores}}{\text{Compras médias por dia}}$	Produto
Participação de terceiros	PT	$\frac{\text{Exigível Total}}{\text{Ativo Total}}$	Insumo
Return on Investment	ROI	$\frac{\text{Lucro Líquido após IR}}{\text{Ativo Total}}$	Produto
Retorno sobre o Patrimônio Líquido	RPL	$\frac{\text{Lucro Líquido após IR}}{\text{PL - Lucro Líquido após IR}}$	Produto

Fonte: Laredo (1997) baseado em Gitman (1987).

Para as empresas constantes das amostras fornecidas pela Serasa Experian (SRS), o conjunto de variáveis financeiras é apresentado na Figura 3.

Figura 3 – Variáveis – Serasa Experian – SRS

Descrição	Abrev.	Cálculo, se aplicável	Classificação DEA
Endividamento - Balanço no ano	ENDV	$(\text{PC} + \text{ELP})/\text{PL} \times 100$	Insumo
Imobilização do Patrimônio Líquido – no ano	IM_PL	$\text{AP}/\text{PL} \times 100$	Insumo
Liquidez Corrente – no ano	LCO	AC/PC	Produto
Liquidez Seca – no ano	LSE	$(\text{AC} - \text{estoques})/\text{PC}$	Produto
Liquidez Geral – no ano	LGE	$\text{AC} + \text{RLP}/\text{PC} + \text{ELP}$	Produto
Ciclo Financeiro base prazos médios – no ano	CF	$\text{CF} = \text{PMRV} + \text{PMRE} - \text{PMPC}$	Insumo
Rentabilidade de Vendas – no ano	REVD	Resultado Líquido do Exercício/Receita Líquida de Vendas x 100	Produto

Descrição	Abrev.	Cálculo, se aplicável	Classificação DEA
Giro do Ativo – no ano	GIAT	Receita Líquida de Vendas/Ativo Total	Produto
Rentabilidade do Patrimônio Líquido – no ano	REPL	Resultado Líquido do Exercício/ Patrimônio Líquido x 100	Produto
NCG / Ativo Cíclico – no ano	NCG	(ACO – PCO)/Ativo Cíclico	Insumo
Saldo de Tesouraria / Necessidade de Capital de Giro – no ano	STCG	(ACF – PCF)/NCG	Indefinido
Saldo de Tesouraria / (Faturamento líquido + impostos) * 100 – no ano	STFL	(ACF – PCF)/FL x 100	Produto
Margem da Atividade – no ano	MGAT	Resultado da Atividade (Operacional)/Receita Líquida de Vendas x 100	Produto
Margem Bruta – no ano	MGBR	Resultado Bruto/Receita Líquida de Vendas x 100	Produto

Fonte: Glossário dos índices econômico-financeiros (Serasa, 2005).

Para estudo do impacto da agregação de variáveis, foi utilizada a Análise Fatorial por Componentes Principais com rotação Varimax. O raciocínio que justificava o investimento é o reconhecimento dos efeitos do número de variáveis no poder de discriminação da DEA, já largamente discutido na literatura. Os indicadores foram submetidos à AF de forma global e, quando havia um número relevante de indicadores em análise, por famílias. As famílias foram definidas conforme agrupamento de indicadores contábeis usual na literatura que será detalhado adiante.

O desenvolvimento dos modelos considerou as questões levantadas na literatura examinada, os resultados dos testes preliminares e as limitações dos dados disponíveis, após a obtenção das amostras para estudo. Para processamento do modelo DEA, foi utilizado o *software* Frontier Analyst®, cedido por Banxia® Software Ltd., Glasgow, Escócia, com capacidade ilimitada para avaliação de unidades e com possibilidade de inclusão de 32 variáveis de *input* ou *output*. Para os estudos estatísticos foram utilizados os programas Excel®, da Microsoft®, e SPSS® 12.0 for Windows, da ©The Apache Software Foundation, disponíveis na FGV e USP. Para as Redes Neurais Artificiais foi utilizado o *Statistica Neural Network*.

4 Apresentação e Discussão dos Resultados

4.1. Resultados dos testes preliminares – Amostra de Laredo (1997) – LRD

O objetivo dos testes preliminares foi desenvolver, testar e comparar os resultados de modelos de análise de insolvência de empresas utilizando Análise Discriminante com Regressão Logística e Regressão Linear Múltipla (do estudo original de Laredo em 1997) e DEA.

Uma das inovações do estudo foi a utilização da Análise Fatorial para agregação e redução do número de variáveis. Conforme anteriormente discutido, para aplicar Análise Fatorial “por famílias”, os indicadores financeiros foram agrupados como é usual na literatura contábil. Assim, por exemplo, os indicadores de Liquidez Corrente, Liquidez Seca e Liquidez

Geral, compuseram a “família” denominada Liquidez, enquanto os indicadores de Retorno sobre o Investimento (ROI) e Retorno sobre o Patrimônio Líquido (ROE), a “família” Rentabilidade.

Para a amostra de teste, os resultados da Análise Fatorial estão apresentados na Figura 4. Cada grupo foi denominado de “família”, como demonstra a primeira coluna, que inclui ainda os indicadores que compõem a família, a codificação dos fatores resultantes e a classificação considerada para análise DEA, como insumo ou produto.

Figura 4 – Famílias de Indicadores Contábeis – LRD

Família	Indicadores	Fatores	Classificação DEA
Liquidez	Liquidez Corrente, Liquidez Geral e Liquidez Seca	XLIQ_1	Produto
Endividamento	Participação do Capital de Terceiros, Exigível sobre o PL, Capital Bancário sobre Ativo Total, Capital Bancário sobre Ativo Circulante, Capital Bancário sobre Passivo Circulante, Capital Bancário sobre o PL	XENDIV_1 XENDIV_2 XENDIV_3	Insumo
Atividade	Prazo Médio de Cobrança, Prazo Médio de Pagamento, Giro dos Estoques e Giro do Ativo	XATIV_1 (GIAT, PMC) XATIV_2 (IGE, PMP)	Produto
Lucratividade	Margem Bruta, Margem Líquida e Margem Operacional	XLCR_1 XLCR_2	Produto
Rentabilidade	ROI e ROE	XRNT_1	Produto

Fonte: Dados da pesquisa.

Na aplicação da Análise Fatorial “Global”, os indicadores foram considerados sem distinção entre os grupos ou “famílias”. Ou seja, realizou-se a Análise Fatorial para todo o conjunto de indicadores. A construção do modelo DEA considerou os fatores conforme apresentado na Figura 5, que mostra, para cada fator, os indicadores que mais correlação apresentam com o fator, sua interpretação e a classificação para DEA (insumo ou produto).

Figura 5 – Fatores – Análise Fatorial Global – LRD

Fator	Indicadores	Interpretação	Classificação
Fator 1	Liquidez Corrente, Liquidez Geral e Liquidez Seca	LIQUIDEZ	Produto
Fator 2	Participação do Capital de Terceiros, ROI e ROE	ALAVANCAGEM E RETORNO	Produto
Fator 3	Capital Bancário sobre Ativo Total, Capital Bancário sobre Ativo Circulante e Capital Bancário sobre Passivo Circulante	ENDIVIDAMENTO	Insumo
Fator 4	Prazo Médio de Cobrança, Margem Líquida e Margem Operacional	ATIVIDADE E LUCRATIVIDADE	Insumo
Fator 5	Exigível sobre o PL, Capital Bancário sobre o PL	GARANTIA AO CAPITAL DE TERCEIROS	Produto
Fator 6	Prazo Médio de Pagamento e Giro do Ativo	ATIVIDADE	Produto
Fator 7	Margem Bruta	LUCRATIVIDADE	Produto
Fator 8	Giro dos Estoques	ATIVIDADE	Produto

Fonte: Dados da pesquisa.

Para classificação do fator como produto ou insumo foram analisados indicadores que o compõem e sua relação com a situação das empresas (falidas ou não falidas).

Na pesquisa original (LAREDO, 1997), foram desenvolvidos dois casos, considerando no Caso 1, empresas insolventes (Tipo C) e empresas solventes (Tipo B); e no Caso 2, empresas insolventes (Tipo C) e empresas não insolventes (Tipos A e B). A função discriminante foi definida em cada caso utilizando os tipos de empresas conforme indicado, sendo depois utilizada para classificar as empresas que não foram incluídas na amostra de desenvolvimento (no Caso 1, empresas Tipo A). Os resultados obtidos e as variáveis incluídas na pesquisa original, utilizando Regressão Logística e Múltipla, são apresentados na Figura 6, que reportam as taxas de acerto global.

Figura 6 - Resultados e Variáveis do Estudo Original

	Regressão Logística	Regressão Múltipla
Taxa de acerto	Caso 1: 81% Caso 2: 80%	Caso 1: 79% Caso 2: 79%
Variáveis incluídas	Caso 1: TCBPC, TCBPL, TIGE, TLG, TPMC Caso 2: TLC, TPT, TCBPC, TCBPL	Caso 1: TCBPC, TCBPL, TIGE, TLG, TPMC Caso 2: TLC, TPT, TCBPC, TCBPL

Fonte: Laredo (1997)

Nos testes preliminares foram desenvolvidos modelos DEA considerando variáveis selecionadas com base na análise de correlação entre as variáveis e entre estas e os escores de eficiência considerando o procedimento Stepwise DEA (KITTELSEN, 1993) - Modelo DEA 1; com as variáveis agregadas por meio da Análise Fatorial, com os indicadores tendo sido agrupados previamente em famílias de indicadores; as famílias consideraram as classificações usais da literatura financeiro-contábil, ou seja, liquidez, endividamento etc. - Modelo DEA 2; com variáveis agregadas por meio da Análise Fatorial Global; ou seja, sem que os indicadores fossem previamente agrupados em famílias - Modelo DEA 3; considerando as variáveis incluídas nos modelos originais de Regressão Logística e Regressão Múltipla - Modelo DEA 4.

Para todas as situações foram considerados três casos, que diferiam quanto à inclusão de cada tipo de empresas: Caso com a Amostra total; Caso confrontando Empresas com excelente saúde x insolventes; e Caso com Empresas solventes x insolventes. O raciocínio é testar o nível de discriminação conforme se altera o pareamento das empresas que compõem a amostra.

Os resultados obtidos com a construção de um Modelo DEA 1, considerando as variáveis originais, selecionadas pelo Stepwise DEA, foram resumidos na Tabela 1 para os casos , e .

O ponto de corte utilizando a amostra completa (Caso - n=86) é o mais baixo (e=27,17), com taxa de acerto global de 79%. No entanto, a taxa de acerto máxima foi obtida no Caso (82%), ao confrontarem-se as empresas A e C (Excelente x Insolvente), resultado semelhante ao obtido por Laredo (1997) no Caso 2 (vide Figura 6). Ressalte-se ainda que, quando, adicionalmente, foram confrontadas as empresas B e C (Caso), foram obtidos resultados bastante razoáveis (e=51,62 e Taxa de acerto global=74%). É interessante analisar ainda a efetividade do modelo em isolar as empresas insolventes (Tipo C), que variou de 40% (Caso), a 68% (Caso). Consistentemente o modelo foi mais efetivo em classificar empresas solventes do que empresas insolventes.

Tabela 1 – Resultados Modelo DEA 1 – Variáveis Originais – LRD

	Situação Real						
	Caso (Amostra total)			Caso (Excelente x Insolvente)		Caso (Solvente x Insolvente)	
Variáveis – Produto	LG, PMP, GIAT, MOt, ROIt						
Variáveis – Insumo	CBAT, CBPC, PMC						
Tipo de empresa	C	B	A	C	A	C	B
Como Insolventes	40%	8%	4%	68%	4%	60%	17%
Como Solventes	60%	92%	96%	32%	96%	40%	83%
N	86			50		61	
Ponto de Corte	e=27,17			e=46,74		e=51,62	
Taxa Acerto Global	79%			82%		74%	
% e=100	34%			36%		41%	
% e<50	38%			36%		33%	

Legenda: t após a variável indica a adição de uma constante para eliminação de valores negativos.

Fonte: Dados da pesquisa.

Os resultados da construção de um Modelo DEA 2, considerando os indicadores originais agrupados em famílias por meio da Análise Fatorial, foram resumidos e confrontados na Tabela 2. Para selecionar os fatores a serem incluídos no modelo DEA, foram utilizados os resultados da RL.

Tabela 2 – Resultados Modelo DEA 2 – Análise Fatorial por Família – LRD

	Situação Real						
	Caso			Caso		Caso	
Produto	XLIQ1, XATV1, XATV2, XLCR1, XLCR2, XRNT1						
Insumo	XENDV1, XENDV2, XENDV3						
Tipo de empresa	C	B	A	C	A	C	B
Como Insolventes	16%	9%	0%	72%	8%	32%	17%
Como Solventes	84%	91%	100%	28%	92%	68%	83%
N	85			50		60	
Ponto de Corte	e=66,68			e=97,77		e=79,24	
Taxa Acerto Global	64%			82%		62%	
% e=100	46%			54%		48%	
% e<50	1%			0%		2%	

Legenda: X antes a variável indica a utilização dos escores fatoriais das variáveis originais agrupadas

Fonte: Dados da pesquisa.

Confirmaram-se os resultados anteriores: a maior taxa de acerto global é obtida no Caso 1, que confronta empresas falidas e empresas com excelente saúde financeira, o que auxilia em uma maior discriminação. Nos outros Casos (2 e 3), os resultados são inferiores aos obtidos com os modelos de Regressão Logística e Múltipla no estudo original (LAREDO, 1997).

Os resultados da construção de um Modelo DEA 3, considerando os indicadores originais agrupados por fatores, foram resumidos na Tabela 3.

Tabela 3 – Resultados Modelo DEA 3 – Escores Fatoriais – AF Global – LRD

	Situação Real						
	Caso 1			Caso 2		Caso 3	
Produto	Fator 1, 2, 6, 7 e 8						
Insumo	Fator 3, 4 e 5						
Tipo de empresa	C	B	A	C	A	C	B
Como Insolventes	32%	83%	0%	60%	16%	72%	47%
Como Solventes	68%	17%	100%	40%	84%	28%	53%
N	85			50		60	
Ponto de Corte	e=38,16			e=68,95		e=82,68	
Taxa Acerto Global	73%			72%		60%	
% e=100	28%			42%		33%	
% e<50	33%			20%		3%	

Fonte: Dados da pesquisa.

Diferente dos modelos anteriores, os resultados em cada Caso (1, 2 e 3) são bastante próximos, com ligeira vantagem para o primeiro caso. Ressalte-se, no entanto, que são inferiores aos obtidos pelo modelo de Regressão Logística e Múltipla do estudo original. A efetividade do modelo em classificar empresas insolventes no Caso 1 foi de 72% das empresas insolventes.

Para o Modelo DEA 4, considerando as variáveis incluídas nos modelos originais de Regressão Logística e Regressão Múltipla, os resultados comparativos foram resumidos nas Tabelas 4 e 5. Foram incluídas as variáveis consideradas nos Casos 1 e 2 do estudo original (LAREDO, 1997), respectivamente.

Analisando a Tabela 4, conclui-se que o melhor desempenho foi obtido no Caso 1 que confronta as empresas C e A (insolventes x excelentes). As taxas de acerto são inferiores às dos modelos de Regressão Logística e Múltipla para o Caso 1 (81% e 79%). A taxa de acerto para o Caso 2 (empresas C x B) foi de 57%. Se todas as empresas fossem classificadas como boas, a taxa de acerto seria de 59% – proporção de empresas boas na amostra – superior à taxa de acerto do modelo.

Tabela 4 – Resultados Modelo DEA 4 (Caso 1) – LRD

	Situação Real						
	Caso			Caso		Caso	
Produto	LG e IGE						
Insumo	CBPL, CBPC, PMC						
Tipo de empresa	C	B	A	C	A	C	B
Como Insolventes	72%	42%	12%	74%	16%	76%	56%
Como Solventes	28%	58%	88%	24%	84%	24%	44%
N	86			50		61	
Ponto de Corte	9,93			11,45		27,65	
Taxa Acerto Global	71%			80%		57%	
% e=100	9%			12%		20%	
% e<50	83%			80%		69%	

Fonte: Dados da pesquisa.

Os resultados apresentados na Tabela 5 confirmam essa análise: o melhor desempenho é obtido no Caso (88% de taxa global de acerto) que supera os modelos de Regressão Logística e Múltipla para o Caso 2 (79%). O pior desempenho é mostrado pelo Caso (empresas C x B) que confronta empresas com menor grau de discriminação em sua performance. Pode indicar maior influência desse fator do que, por exemplo, do número de empresas na amostra, um ponto bastante discutido na literatura sobre DEA. Assim, a se julgar por esses resultados, a escolha de empresas e variáveis que representem características diferenciadoras apresenta maior influência do que o tamanho da amostra.

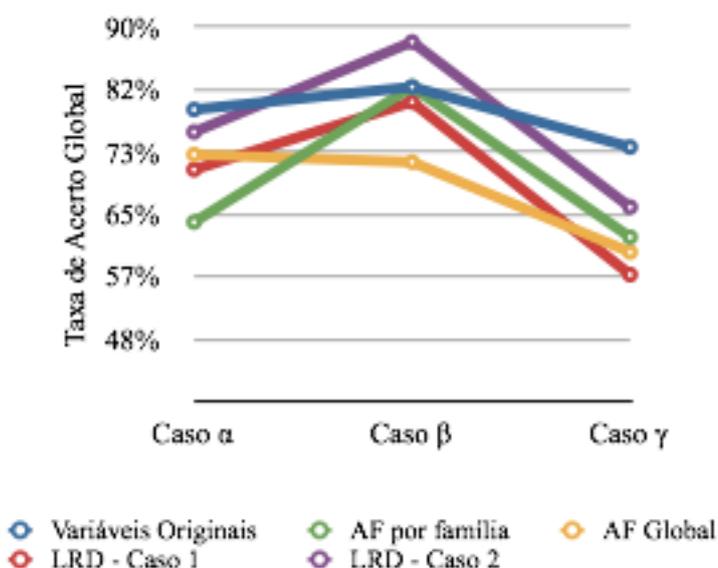
Tabela 5 – Resultados Modelo DEA 4 (Caso 2) – LRD

	Situação Real						
	Caso			Caso		Caso	
Produto	LC						
Insumo	PT, CBPC e CBPL						
Tipo de empresa	C	B	A	C	A	C	B
Como Insolventes	24%	6%	0%	88%	12%	24%	6%
Como Solventes	76%	94%	100%	12%	88%	76%	94%
N	86			50		61	
Ponto de Corte	1,71			10,71		1,71	
Taxa Acerto Global	76%			88%		66%	
% e=100	7%			10%		8%	
% e<50	90%			84%		92%	

Fonte: Dados da pesquisa.

Para resumir, o Gráfico 1 mostra os resultados comparativos de cada modelo. O poder de discriminação do Caso α , que confronta empresas insolventes e excelentes, implica resultados superiores em todos os modelos. No entanto, o modelo de resultados mais estáveis, independente do tipo de amostra e do número de empresas, é o modelo com variáveis originais. O modelo com resultado superior é o modelo com as variáveis do estudo original, considerando o Caso 2, e amostra com empresas excelentes *versus* insolventes.

Gráfico 1 – Comparativo de Resultados - Amostra de Laredo (1997)



Fonte: Dados da pesquisa.

Os testes preliminares confirmam resultados obtidos em estudos anteriores. Demonstram, como no estudo de Onusic, Casa Nova e Viana (2004), que o modelo DEA foi capaz de discriminar com maior grau de acerto as empresas insolventes. Quanto ao raciocínio de testar se o nível de discriminação aumenta conforme se altera o pareamento das empresas que compõem a amostra, ficou claro que as amostras que confrontam empresas falidas e empresas com excelente saúde financeira têm, em qualquer dos modelos testados, maior poder de discriminação.

4.2 Resultados para a amostra de empresas do banco de dados da Serasa – SRS

Para a amostra da Serasa, a Análise Fatorial foi aplicada em cada uma das duas subamostras (empresas industriais de médio porte e empresas comerciais de médio porte), em cada um dos anos em estudo. Os resultados da Análise Fatorial foram consistentes para os três anos considerados. Os cinco componentes resultantes em cada ano apresentam 71,6%, 70,3% e 73,7%, respectivamente, da variância total, em cada ano. Os escores fatoriais de cada componente foram retidos na análise dando origem a cinco novas variáveis (FAC1,

FAC2, FAC3, FAC4 e FAC5). Os resultados em cada ano são consistentes e o significado subjacente dos fatores é o que se segue:

- O Fator 1 (FAC1) consolida os indicadores de liquidez e endividamento da empresa, bem como a necessidade de capital de giro. É uma variável que se relaciona com a capacidade de pagamento e de financiamento da empresa.
- O Fator 2 (FAC2) apresenta Rentabilidade e margem nos anos de 2001 e 2002, enquanto que para o ano de 2003 estão presentes os indicadores de Saldo de Tesouraria, indicativos da capacidade da empresa financiar suas necessidades de capital de giro com recursos próprios.
- No Fator 3 (FAC3) rentabilidade e margem estão presentes no ano de 2003, enquanto que para os anos de 2002 e 2001 estão presentes os indicadores de Saldo de Tesouraria, indicativos da capacidade de a empresa financiar suas necessidades de capital de giro com recursos próprios.
- O Fator 4 (FAC4) reflete os indicadores de giro e de ciclo financeiro relacionados, respectivamente, ao nível de atividade da empresa e a suas características de recebimento e pagamento.
- O Fator 5 (FAC5) em todos os anos apresenta os indicadores de retorno sobre patrimônio líquido.

Para as empresas do setor de Comércio, repetiu-se o procedimento. Foram retidos quatro fatores que representam, em cada ano, 71,56%, 70,46% e 72,19%, respectivamente, da variância total. Os resultados foram consistentes entre os anos, demonstrando a solidez da análise. Os escores de cada fator foram retidos, gerando quatro variáveis: FAC1, FAC2, FAC3 e FAC4. Pela análise os fatores têm o seguinte significado:

- O Fator 1 (FAC1) representa a situação financeira das empresas, congregando indicadores de capacidade de pagamento, endividamento e estrutura.
- O Fator 2 (FAC2) inclui as variáveis Margem Bruta, Giro e Ciclo Financeiro, tendo relação com a operação da empresa.
- O Fator 3 (FAC3) indica a rentabilidade e lucratividade das empresas refletindo, portanto, a situação econômica da empresa.
- O Fator 4 (FAC4) está relacionado com a situação financeira das empresas, consolidando os indicadores compostos pelo Saldo de Tesouraria.

Os resultados obtidos com DEA foram confrontados com os resultados obtidos para as mesmas amostras em modelos de previsão de insolvência/inadimplência utilizando outras técnicas (Análise Discriminante com Regressão Logística e Redes Neurais). Assim, foi estabelecido o grau de acerto comparativo do modelo DEA. Os erros de classificação (empresas inadimplentes/insolventes classificadas como saudáveis e vice-versa) foram estudados em detalhe na busca de determinar os fatores que contribuíram para sua classificação.

4.2.1. Amostra de Indústria de Médio Porte – SRS – Indústria

4.2.1. 1. Resultados da Regressão Logística

Procedeu-se então à Regressão Logística (RL), com o método *Forward Stepwise Wald*, utilizando indicadores (originais e transformados) e escores fatoriais obtidos pela AF. O procedimento foi repetido para cada um dos anos da amostra.

Os resultados da RL consideraram o método *Stepwise Forward Wald*, com as seguintes características, para o ano de 2003, os indicadores originais ou transformados – inclusão de Liquidez Seca (Ln) e de Saldo de Tesouraria sobre Faturamento Líquido (Ln); os escores fatoriais – inclusão dos fatores 1, 2 e 4 (no caso dos anos de 2002 e 2001; em 2003, o fator 2 foi substituído pelo fator 3). Os resultados dos modelos com as variáveis originais transformadas e escores fatoriais em cada ano estão resumidos na Tabela 6.

Tabela 6 – Resultados da RL – SRS – Indústria

		Indicadores Originais			Escores Fatoriais		
		Taxas de Acerto			Taxas de Acerto		
Ano		2001	2002	2003	2001	2002	2003
Situação	Ruim	41%	53%	54%	43%	53%	57%
	Boa	79%	71%	75%	77%	76%	76%
Acerto Global		62%	63%	66%	62%	65%	68%

Fonte: Dados da pesquisa.

4.2.1.2. Resultados da Análise por Envoltória de Dados

Foram construídos três modelos DEA considerando: (1) os indicadores originais retidos pela RL; (2) os escores fatoriais incluídos na RL; e, (3) as variáveis indicadas pelo procedimento *Stepwise DEA*. Assim, resultaram os seguintes modelos: Modelo DEA 1 (Variáveis originais) com Liquidez Seca e Saldo de Tesouraria sobre Faturamento Líquido (*outputs*) e de Endividamento (*input*); Modelo DEA 2 (Fatores) com inclusão dos Fatores 1, 2 (*outputs*) e 4 (*input* - no caso dos anos de 2002 e 2001; em 2003 o fator 2 foi substituído pelo fator 3; Modelo DEA 3 (*Stepwise DEA*) composto por Liquidez Geral, Giro, Retorno sobre o PL, Retorno sobre Vendas, Margem da Atividade, Margem Bruta (*outputs*); Endividamento, Imobilização do PL (*inputs*). Os resultados obtidos em cada um dos modelos estão resumidos na Tabela 7, que mostra os percentuais de acerto de classificação, para cada modelo, dado o ponto de corte.

Tabela 7 – Resultados da DEA – SRS – Indústria

		Taxas de Acerto								
Modelo		DEA 1			DEA 2			DEA 3		
Ano		2001	2002	2003	2001	2002	2003	2001	2002	2003
Variáveis		LS, STFL (<i>outputs</i>) ENDIV (<i>input</i>)			FAC1 e FAC2 ou 3 (<i>outputs</i>) FAC4 (<i>input</i>)			LG, Giro, REPL, REVD, MGAT, MGBR (<i>outputs</i>) ENDIV e IMPL (<i>inputs</i>)		
Situação	Ruim	23%	3%	17%	16%	36%	44%	53%	22%	16%
	Boa	85%	98%	90%	93%	79%	74%	63%	85%	92%
Acerto Global		58%	55%	57%	58%	60%	61%	59%	57%	58%
Ponto de Corte		e=2,17	e=0,54	e=1,77	e=51,94	e=56,6	e=64,88	e=19,19	e=9,22	E=5,38
% e=100		2%	2%	2%	4%	3%	4%	8%	11%	13%
% e<0		95%	95%	95%	7%	12%	6%	83%	78%	76%

Fonte: Dados da pesquisa.

O modelo DEA 2 (cargas fatoriais) mostrou os melhores resultados considerando-se a taxa de acerto global. Os resultados são, no entanto, satisfatórios apenas para o ano de 2003. Isso porque, se considerarmos que a amostra total (n=259) contava com 55,2% de empresas boas (n=143), o ganho informativo é de apenas 8%, para o ano de 2002, e de 5%, em 2001. O erro tipo II (classificar como boas empresas ruins) é, em todos os anos, relevante, variando de 97,4% a 46,5%, respectivamente.

4.2.1.3. Resultados da DEA Inversa

Para construção do modelo DEA Inversa, foi utilizado o modelo DEA 2 (com cargas fatoriais), por ter apresentado os melhores resultados comparativos. As variáveis de *input* e *output* foram invertidas, bem como a orientação do modelo. Depois de obtidos os escores de eficiência, foi calculada a eficiência composta para cada empresa. A Tabela 8 apresenta as taxas de acerto para a classificação das empresas considerando os escores de eficiência composta.

Tabela 8 – Resultados da DEA Inversa – SRS – Indústria

		Taxas de Acerto		
		2001	2002	2003
Variáveis		FAC4 (output) FAC1 e FAC2 ou 3 (inputs)		
Situação	Ruim	52%	54%	50%
	Boa	59%	70%	73%
Acerto Global		56%	63%	63%
Ponto de Corte		e=63,02	e=61,84	e=62,08

Fonte: Dados da pesquisa.

Os resultados foram consistentes com os obtidos pelos modelos anteriores, aumentando ligeiramente as taxas de acerto global, com melhora relevante no desempenho de classificação correta entre as empresas ruins. Porém, entre as empresas boas, as taxas de acerto em todos os anos foram reduzidas.

4.2.1.4. Resultados das Redes Neurais Artificiais

Para as aplicações apresentadas a seguir, foram utilizadas redes Perceptron multicamadas. A base de dados foi subdividida nas amostras de desenvolvimento (n=130), treinamento (n=69) e teste (n=60). As taxas de acerto são apresentadas na Tabela 9, para cada uma das subamostras, em cada um dos anos.

Tabela 9 – Resultados das Redes Neurais Artificiais – SRS – Indústria

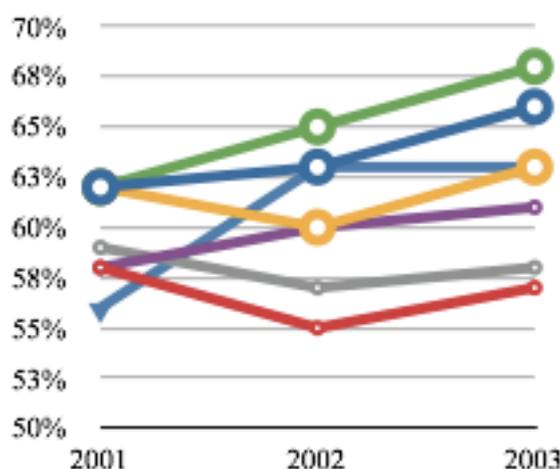
		Taxas de Acerto								
		2001			2002			2003		
Ano		Desenv.	Trein.	Teste	Desenv.	Trein.	Teste	Desenv.	Trein.	Teste
Amostras										

Variáveis		REVD, MGAT, MGBR, L_ENDV, L_IMPL, L_LCO, L_LSE, L_LGE, L_GIAT, CF, NCG, STFL, REPL	L_LCO, L_LSE, L_LGE, L_GIAT, NCG, REPL	L_IMPL, L_GIAT, STFL, REPL						
Situação	Ruim	65%	61%	61%	64%	52%	68%	60%	79%	65%
	Boa	68%	56%	62%	67%	58%	52%	69%	64%	62%
	Total	67%	58%	62%	65%	55%	60%	65%	64%	63%

Fonte: Dados da pesquisa.

O Gráfico 2 apresenta um comparativo dos resultados para cada modelo apresentado, para o período de estudo. O modelo com resultados mais efetivos foi o de regressão logística com escores fatoriais. Entre os modelos DEA, destacou-se o modelo inverso, com resultados estáveis nos anos de 2002 e 2003. Os modelos DEA 1, DEA 3 e Redes Neurais apresentaram taxas de acerto piores no ano intermediário, contrariamente ao esperado.

Gráfico 2 – Comparativo de Resultados - Amostra Serasa – Indústria



◆ RL - Indicadores
 ◆ RL - Fatores
 ◆ Redes Neurais
 ◆ DEA 1
◆ DEA 2
 ◆ DEA 3
 ◆ DEA Inversa

Fonte: Dados da pesquisa.

4.2.2. Empresas comerciais de médio porte – SRS – Comércio

4.2.2.1. Resultados da Regressão Logística

Na RL, foi utilizado o método *Stepwise Forward Wald*, com os dados considerados para as empresas industriais, ou seja, indicadores originais ou transformados quando necessário – inclusão de Margem da Atividade, Retorno sobre Vendas, Saldo de Tesouraria sobre NCG, Saldo de Tesouraria sobre Faturamento Líquido e Giro do Ativo (Ln); Escores Fatoriais – foram incluídos os fatores FAC1 e FAC4, para os anos de 2003 e 2002. No ano de 2001, além dos fatores citados anteriormente, foi considerado o FAC3. As taxas de acerto são apresentadas na Tabela 10.

Tabela 10 – Resultados da RL – SRS – Comércio

Ano	Situatção	Variáveis Originais			Escores Fatoriais		
		2001	2002	2003	2001	2002	2003
	Ruim	46%	56%	58%	36%	46%	52%
	Boa	76%	76%	76%	74%	75%	75%
	Acerto Global	62%	67%	68%	56%	61%	64%

Fonte: Dados da pesquisa.

As taxas de acerto vão se deteriorando a cada ano e não acrescentam informação à decisão, pois, se todas as empresas fossem classificadas como boas (n=194), a taxa de acerto seria de 54%.

4.2.2.2. Resultados da Análise por Envoltória de Dados

Foram construídos os seguintes modelos DEA: Modelo DEA 1 (Variáveis Originais selecionadas pela RL): inclusão dos indicadores de Margem da Atividade, Retorno sobre Vendas, Saldo de Tesouraria sobre NCG, Saldo de Tesouraria sobre Faturamento Líquido e Giro do Ativo; Modelo DEA 2 (Escores Fatoriais): considerados os fatores 1 e 4 (*outputs*) e fator 3 (*input*); Modelo DEA 3 (Stepwise DEA): Retorno sobre Vendas, Liquidez Geral, Saldo de Tesouraria sobre Faturamento Líquido, Saldo de Tesouraria sobre NCG, Giro do Ativo (*outputs*) e Endividamento (*inputs*). As Tabelas 11 e 12 apresentam o resumo estatístico dos resultados obtidos em cada modelo.

A Tabela 11 resume os resultados obtidos com a utilização dos modelos DEA para classificação das empresas em boas e ruins.

Tabela 11 – Resultados da DEA – SRS – Comércio

Modelo	DEA 1			DEA 2			DEA 3			
Ano	2001	2002	2003	2001	2002	2003	2001	2002	2003	
Variáveis	MGAT, REVD, STCG, STFL, GIAT (outputs) ENDIV (input)			FAC1 e FAC4 (outputs) FAC3 (input)			REVD, LGE, STFL, STCG, GIAT (outputs) ENDIV (inputs)			
Situatção	Ruim	27%	27%	32%	79%	68%	49%	23%	42%	46%
	Boa	83%	83%	77%	50%	53%	73%	85%	71%	66%
Acerto Global		57%	57%	56%	63%	60%	62%	56%	58%	58%
Ponto de Corte		e=0,3	e=0,3	e=0,4	e=26,4	e=40,5	e=29,1	e=0,3	e=0,5	e=0,6
% e=100		4%	3%	2%	1%	2%	2%	4%	3%	2%
% e<50		94%	96%	98%	96%	83%	80%	94%	96%	98%

Fonte: Dados da pesquisa.

O modelo DEA 2, que considerou as cargas fatoriais, mostrou o melhor desempenho em todos os anos, com taxas de acerto global superiores a 60%. Foi também o modelo com menor percentual de empresas classificadas como eficientes (e=100). Se for considerado que,

na amostra total (n=360), há 194 empresas boas (53,9%), tem-se um ganho informacional de 15%, 11% e 17%, respectivamente, em 2003, 2002 e 2001.

4.2.2.3. Resultados da DEA Inversa – SRS – Comércio

O modelo DEA Inversa considerou a inversão de *inputs* e *outputs* do modelo DEA 2, com cargas fatoriais. Posteriormente, foi calculada a eficiência composta. A Tabela 12 mostra o resumo obtido da análise dos resultados.

Tabela 12 – Resultados da DEA Inversa – SRS – Comércio

		Taxas de Acerto		
Ano		2001	2002	2003
Variáveis		FAC3 (output) FAC1 e FAC4 (inputs)		
Situação	Ruim	64%	69%	54%
	Boa	64%	60%	68%
Acerto Global		64%	54%	61%
Ponto de Corte		e=38,3	e=53,9	e=45,0

Fonte: Dados da pesquisa.

4.2.2.4. Resultados das Redes Neurais Artificiais – SRS – Comércio

Foram repetidos os procedimentos adotados para a base de dados das empresas industriais. As taxas de acerto para cada subamostra, em cada ano, estão resumidos na Tabela 13.

Tabela 13 – Resultados das Redes Neurais Artificiais – SRS – Comércio

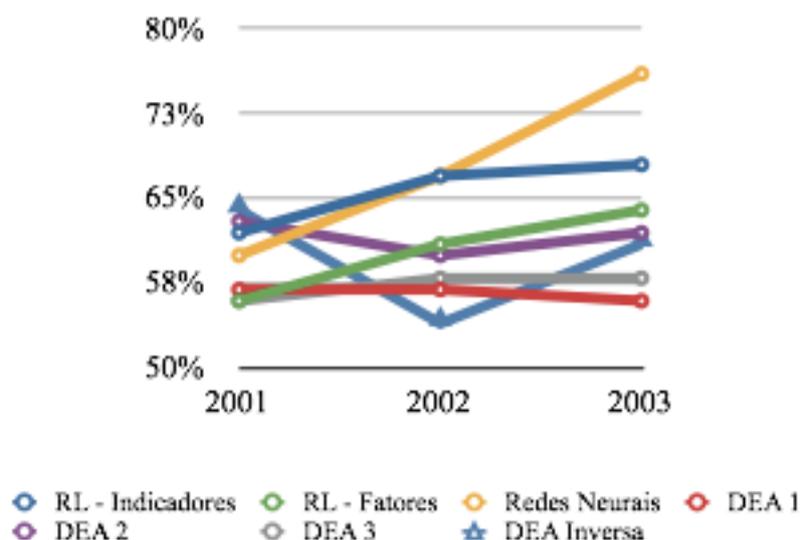
Ano		2001			2002			2003		
Amostras		Desenv.	Trein.	Teste	Desenv.	Trein.	Teste	Desenv.	Trein.	Teste
Variáveis		MGAT			MGAT, REVD, REPL, CF, NCG, STCG, STFL, L_MGBR, L_IMPL, L_ENDV, L_GIAT, L_LCO, L_LSE, L_LGE			MGAT, REVD, REPL, CF, NCG_2003, STFL, L_IMPL, L_GIAT20, L_LCO, L_LGE		
Situação	Ruim	58%	64%	58%	64%	73%	67%	63%	64%	76
	Boa	59%	57%	61%	60%	78%	67%	64%	71%	63%
	Total	59%	60%	60%	62%	75%	67%	63%	68%	68%

Fonte: Dados da pesquisa.

Foram alcançados resultados comparáveis aos demais métodos (RL e DEA), porém com maior equilíbrio entre a classificação correta entre empresas boas e ruins. Ressalte-se o fato de que, em 2001, foi selecionada apenas uma variável (MGAT) para compor o modelo.

O Gráfico 3 apresenta o resultado comparativo dos modelos desenvolvidos. Destacou-se claramente o modelo construído com redes neurais. O modelo DEA Inversa teve uma evolução de resultados contraditória, com a maior taxa de acerto no primeiro ano do período.

Gráfico 3 – Comparativo de Resultados - Amostra Serasa - Comércio



Fonte: Dados da pesquisa.

O modelo de regressão logística com fatores igualmente se destacou, sendo o terceiro com melhores resultados.

5 Considerações Finais

O estudo se concentrou em identificar as vantagens e limitações na utilização da Análise por Envoltória de Dados para análise de insolvência/inadimplência. Para tanto, foram pesquisados, na literatura, outros trabalhos que tenham examinado o tema no Brasil e no exterior. Os trabalhos encontrados foram resumidos, sendo apresentados sua metodologia, contribuições, limitações e resultados. Verificou-se que o tema é ainda recente (data de 1994 o primeiro trabalho publicado), sendo especificamente em DEA aplicado à análise de insolvência/inadimplência. Em geral, os resultados das pesquisas anteriores demonstraram que os modelos DEA obtiveram resultados consistentes ou superiores às análises tradicionais, mas com indicações de melhoria do desempenho financeiro. Foram ainda propostos modelos para aplicação específica para análise de insolvência, como os modelos de piores práticas e a utilização da técnica de camadas com taxas de acerto maiores que os tradicionais com utilização de análise discriminante.

A metodologia proposta incluiu uma etapa de testes preliminares, com amostra de estudo anterior (LAREDO, 1997). Foram estabelecidas diversas amostras a partir das bases de dados da Serasa. Essas amostras consideraram subamostras dos setores de indústria e comércio, com três anos antes da ocorrência dos problemas. Foi procedido então ao tratamento de dados, com identificação de outliers, transformação de variáveis e análise de medidas de estatística descritiva. Os indicadores foram ainda submetidos à Análise Fatorial, para redução de dimensionalidade e estudo dos efeitos da utilização das cargas fatoriais como variáveis em DEA. Como procedimento para definição das variáveis

a serem incluídas nos modelos, foram utilizados, alternadamente, o procedimento Stepwise Forward da Regressão Logística e o procedimento Stepwise DEA. Finalmente, foram construídos modelos DEA considerando as variáveis originais e as cargas fatoriais. Para os modelos com melhores resultados comparativos, foi construída a fronteira invertida e calculada a eficiência composta. Os resultados foram comparados com os obtidos com a Regressão Logística e Redes Neurais.

Como conclusões, os resultados dos modelos DEA foram comparáveis aos obtidos dos modelos de Regressão Logística e Redes Neurais para as amostras Serasa setoriais, subdivididas em comércio e indústria. Os resultados dos diversos modelos DEA desenvolvidos foram contraditórios, sem uma indicação clara de superioridade de algum modelo ou procedimento. A proposta de utilização da Análise Fatorial apresentou-se promissora, particularmente quando associada à Regressão Logística. Por fim, a utilização da fronteira invertida, com cálculo da eficiência composta, manteve as taxas de acerto global praticamente inalteradas, com ligeira melhora do acerto de classificação das empresas ruins.

Uma das limitações do presente estudo é o fato de se concentrar apenas ao estudo de uma base de dados (SERASA) e de englobar empresas de diferentes setores, agrupadas em empresas industriais e comerciais. A literatura contábil tem demonstrado que os indicadores podem ter comportamento diferenciado quando consideradas empresas de diferentes setores de atuação. Esse ponto pode ter influenciado os resultados do estudo.

A pesquisa internacional confirma as fragilidades da DEA quando utilizada na predição de falência/insolvência. No entanto, apesar de largamente explorada no exterior, a temática permanece incipiente no Brasil. Poucos autores têm se dedicado ao estudo do fenômeno apesar de sua relevância.

Estudos futuros podem se endereçar a aplicações em determinados setores, com ênfase em área que recentemente enfrentaram um processo de consolidação, ou tentar englobar períodos de análise mais longos. A confirmação da utilidade da Análise Fatorial é também uma interessante área de estudo. Setores altamente regulados são particularmente interessantes para os pesquisadores, uma vez que a qualidade da informação é maior e, muitas vezes, existem informações de natureza operacional ao lado de informações econômico-financeiras.

Referências

- ALI, Agha Iqbal; SEIFORD, Lawrence M. Translation invariance in Data Envelopment Analysis. **Operations Research Letters**, v. 9, n. 6, p. 403-405, 1990.
- ALMEIDA, Henry Rossi; MILIONI, Armando Zeferino. Análise de Envoltória de Dados na decisão de concessão de crédito. In: Simpósio Brasileiro de Pesquisa Operacional, XXXII, 2000, Viçosa. **Anais ... SIMPOI, 2000**. CD-ROM.
- BARTH, Nelson Lerner. **Inadimplência: construção de modelos de previsão**. São Paulo: Nobel, 2004.
- BANKER, R. D.; CHARNES, A.; COOPER, W. W. Some Models for estimating technical and scale inefficiencies in Data Envelopment Analysis. **Management Science**, v. 30, n. 9, p. 1078-1092, Sept. 1984.

CASA NOVA, Silvia Pereira de Castro. **Utilização da Análise por Envoltória de Dados (DEA) na Análise de Demonstrações Contábeis**. 2002. 318 f. Tese (Doutorado em Controladoria e Contabilidade) - Programa de Pós-Graduação em Controladoria e Contabilidade, Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, 2002.

CASA NOVA, Silvia Pereira de Castro. Bons em ser Ruins: A utilização da Análise por Envoltória de Dados (DEA) em modelos de análise de inadimplência/insolvência de empresas. In: Encontro Anual da Associação de Pesquisa e Pós-Graduação em Administração, ENANPAD 2010, 2010, Rio de Janeiro. **Anais ... ENANPAD, 2010**. CD-ROM.

CASA NOVA, Silvia Pereira de Castro; ONUSIC, Luciana Massaro. Mapeamento de estudos sobre a utilização de Análise por Envoltória de Dados (DEA) na análise de insolvência. **Contabilidade, Gestão e Governança**, v. 8, n. 2, p. 207-238, 2005.

CIELEN, A.; PEETERS, L.; VANHOOF, K. Bankruptcy prediction using a data envelopment analysis. **European Journal of Operational Research**, v. 154, n. 2, p. 526-532, 2004.

COLAUTO, Romualdo Douglas; PINHEIRO, Laura Edith Taboada; PINHEIRO, Juliano Lima. Información Contable en la Predicción de Insolvencia: estudio inferencial univariante aplicado a empresas españolas. **Revista Contemporânea de Contabilidade**, UFSC, Florianópolis, v.1, n.12, p. 151-170, jul./dez., 2009.

EMEL, A. B.; ORAL, M.; REISMAN, A.; YOLALAN, R. A Credit score approach for the commercial banking sector. **Socio-Planning Sciences**, v. 37, n. 2, p. 103-123, 2003.

FERNANDEZ-CASTRO, A.; SMITH, P. Towards a general non-parametric model of corporate performance. **Omega – International Journal of Management Science**, v. 22, n. 3, p. 237-249, 1994.

FIGUEIREDO, Denise Santos de; MELLO, João Carlos Correia Baptista Soares de. **Avaliação de eficiência de lojas do varejo supermercadista**. Relatórios de Pesquisa em Engenharia de Produção da UFF. Mestrado em Engenharia de Produção. Universidade Federal Fluminense, v. 4, p. 1-11, 2004. Disponível em: http://www.producao.uff.br/conteudo/rpep/volume42004/RelPesq_V4_2004_21.pdf.

HAIR JR., J. F. ANDERSON, R.E. TATHAM, R.L. BLACK, W.C. **Multivariate Data Analysis**. Englewood Cliffs, N.J.: Prentice Hall, 1998.

KITTELSEN, Sverre A.C. Stepwise **DEA**: Choosing variables for measuring technical efficiency in Norwegian electricity distribution. SNF, Foundation for research in economics and business administration, Oslo, 1993.

KNOX LOVELL, C. A.; PASTOR, J. T. Units invariant and translation invariance DEA models. **Operations Research Letters**, v. 18, n. 3, p. 147-151, 1995.

LAREDO, André. **Modelos de previsão de insolvência baseados na análise de índices financeiros**. 1997. 113f. il. Projeto de conclusão de curso (Graduação em Administração) - Escola Superior de Propaganda e Marketing, 1997.

LIU, John S.; LU, Louis Y.Y.; LU, Wen-Min; LIN, Bruce J.Y. A Survey of DEA Applications. **Omega – International Journal of Management Science**, v. 41, n. 5, p. 893-902, 2013.

MACEDO, Marcelo Alvaro da Silva; CAVALCANTE, Glaydson Teixeira. Performance de agências bancárias: aplicando DEA a múltiplas perspectivas do desempenho. **Revista Contemporânea de Contabilidade**, UFSC, Florianópolis, v. 1, n. 12, p. 87-108, jul./dez. 2009.

OLIVEIRA, V. H.; MACEDO, M. A. S.; CORRAR, L. J. Estudo do desempenho dos maiores bancos de varejo no Brasil por meio da Análise por Envoltória de Dados (DEA). **RIC - Revista de Informação Contábil**. v. 5, n. 2, p. 01-20, abr./jun. 2011.

ONUSIC, Luciana Massaro. **A utilização conjunta das técnicas análise por envoltória de dados e regressão logística na previsão de insolvência de empresas: um estudo exploratório**. 2004. 123 f. Dissertação (Mestrado em Administração) - Programa de Pós-Graduação em Administração, Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, 2004.

ONUSIC, Luciana Massaro; CASA NOVA, Silvia Pereira de Castro; ALMEIDA, Fernando Carvalho de. Modelos de previsão de insolvência utilizando a Análise Envoltória de Dados: aplicação a empresas brasileiras. **RAC - Revista de Administração Contemporânea**, v. 11, n. 2, p. 77-97, 2007.

ONUSIC, Luciana Massaro; CASA NOVA, Silvia Pereira de Castro; VIANA, Adriana Backx Noronha. Comparação de resultados de utilização de Análise por Envoltória de Dados e Regressão Logística em modelos de previsão de insolvência: um estudo aplicado a empresas brasileiras. **FACEF Pesquisa**, v. 7, n. 1, 2004.

PARADI, J. C. et al. Applications of DEA to measure the efficiency of software production at two large Canadian banks. **Annals of Operations Research**, v. 73, p. 91-115, 1997.

PARADI, J.; ASMILD, M.; SIMAK, P. Using DEA and worst practice DEA in credit risk evaluation. **Journal of Productivity Analysis**, v. 21, n. 2, p. 153-165, 2004.

PASTOR, J. T. Translation invariance in data envelopment analysis: a generalization. **Annals of Operations Research**, v. 66, p. 93-112, 1996.

PREMACHANDRA, I.M.; CHEN, Yao; WATSON, John. DEA as a tool for predicting corporate failure and success: A case of bankruptcy assessment. **Omega – International Journal of Management Science**, v. 39, n. 6, p. 620-626, 2011.

SANTOS, Arioaldo; CASA NOVA, Silvia Pereira de Castro. Proposta de um Modelo estruturado de análise de demonstrações contábeis. **RAE- eletrônica** - v. 4, n. 1, Art. 8, jan./jun. 2005 (www.rae.com.br/eletronica).