Insper Instituto de Ensino e Pesquisa
Programa de Mestrado Profissional em Economia

Vagner Aparecido Pedro Junior

MODELO DE PREVISÃO DE INSOLVÊNCIA PARA GRANDES EMPRESAS: UMA ABORDAGEM COM GEE

São Paulo

2011

Vagner Aparecido Pedro Junior

Modelo de previsão de insolvência para grandes empresas: Uma abordagem com GEE

Dissertação apresentada ao Programa de Mestrado Profissional em Economia do Insper Instituto de Ensino e Pesquisa, como parte dos requisitos para a obtenção do título de Mestre em Economia.

Área de concentração: Finanças e Microeconomia Aplicadas
Orientador: Prof. Dr. Rinaldo Artes – Insper

São Paulo

2011

Pedro Junior, Vagner Aparecido

Modelo de previsão de insolvência para grandes empresas: Uma abordagem com GEE/ Vagner Aparecido Pedro Junior; orientador: Rinaldo Artes;
– São Paulo: Insper, 2011. 47 f.

Dissertação (Mestrado – Programa de Mestrado Profissional em Economia. Área de concentração: Finanças e Microeconomia Aplicadas) – Insper Instituto de Ensino e Pesquisa.

FOLHA DE APROVAÇÃO

Vagner Aparecido Pedro Junior

Modelo de previsão de insolvência para grandes empresas

Dissertação apresentada ao Programa de Mestrado Profissional em Economia do Insper Intituto de Ensino e Pesquisa, como requisito parcial para obtenção do título de Mestre em Economia.

Área de concentração: Finanças e Microeconomia Aplicadas

Aprovado em: Julho/2011

Banca Examinadora

Prof. Dr. Rinaldo Artes (orientador)

Orientador

Instituição: Insper

Prof. Dr. Naercio Aquino Menezes Filho

Instituição: Insper

Prof. Dr. Nilton Deodoro Moreira Cardoso Junior

AGRADECIMENTOS

Agradeço a Deus por sempre iluminar os meus caminhos.

Agradeço aos meus pais e irmãos por todo o apoio e por sempre acreditarem no meu potencial.

Ao professor Dr. Rinaldo Artes, pela paciência, confiança e pela contribuição acadêmica, fundamental para a qualidade deste trabalho.

Aos meus colegas de trabalho, agradeço pelo apoio e contribuições de forma direta ou indireta ao trabalho.

À minha equipe, um agradecimento especial, tanto pela motivação como pela presença em minha defesa.

Por fim, à minha namorada por ter suportado esta fase da minha vida, por sempre me motivar e ser a companheira perfeita que é.

RESUMO

Pedro Junior, Vagner Aparecido. **Modelo de previsão de insolvência para grandes empresas: Uma abordagem com GEE.** 2011. 47 f. Dissertação (Mestrado) – Insper Instituto de Ensino e Pesquisa, São Paulo, 2011.

Este trabalho aborda a modelagem de previsão de insolvência para grandes empresas. É sugerida uma *proxy* de insolvência*:* apresentar Patrimônio Líquido negativo por quatro trimestres consecutivos. São propostas diferentes formas de amostragem e diferentes técnicas estatísticas. Primeiramente, a regressão logística condicional, recomendada para amostras emparelhadas. Depois, a metodologia GEE, uma extensão dos modelos GLM, que procura acomodar a correlação longitudinal dos dados. Com essas técnicas, procura-se identificar índices econômico-financeiros e variáveis macroeconômicas que auxiliem na previsão de insolvência, gerando uma ferramenta para mensurar a qualidade de empresas de capital aberto.

Palavras-chave: Insolvência, Patrimônio Líquido, GEE, macroeconômica, regressão logística condicional.

ABSTRACT

Pedro Junior, Vagner Aparecido. **Predictive model of insolvency for large companies: A GEE approach.** 2011. 47 f. Dissertação (Mestrado) – Insper Instituto de Ensino e Pesquisa, São Paulo, 2011.

This paper investigates the predictive modeling of insolvency for large companies. It suggests a proxy for insolvency: submit negative equity for four consecutive quarters. Are proposed different ways of sampling and different statistical techniques. First, conditional logistic regression, recommended for paired samples. Then, the GEE methodology, an extension of GLM models, which accommodates the correlation of longitudinal data. With these techniques, seeks to identify economic-financial indicators and macroeconomic variables that aid in the prediction of insolvency, generating a tool to measure the quality of publicly traded companies.

Keywords: Insolvency, Equity, GEE, macroeconomic, conditional logistic regression.

Sumário

1. Introdução 9
2. Revisão da Literatura 10
2.1 Modelos de previsão de falência e insolvência 10
2.2 Definição de insolvência 12
3. Metodologia 16
3.1 Regressão Logística 16
3.2 Regressão Logística Condicional 17
3.3 Equações de estimação generalizadas (GEE) 18
3.3.1 Correlação uniforme (*Exchangeable correlation*) 20
3.3.2 Correlação Auto-regressiva (*Autoregressive correlation*) 20
3.3.3 Correlação não estruturada (*Unstructured correlation*) 21
4. Base de Dados 21
4.1 Descrição da variável resposta 21
4.2 Descrição das variáveis explicativas 22
4.3 Descrição dos dados 23
4.3.1 Amostra emparelhada 24
4.3.2 Visão carteira 27
5. Resultados 28
5.1 Resultados da amostra emparelhada 29
5.1.1 Um ano antes do evento de insolvência 29
5.1.2 Dois anos antes do evento de insolvência 32
5.2 Resultados da visão carteira 34
6. Conclusão 40
APÊNDICE A 42
Descrição dos índices econômico-financeiros 43
APÊNDICE B 44
Relação das empresas solventes e insolventes para a primeira amostra 44
APÊNDICE C 45
Regressão Logística Condicional: um ano antes do evento. 45
Regressão Logística Condicional: dois anos antes do evento. 45
APÊNDICE D 46
APÊNDICE E 47
Referências Bibliográficas 48

1. Introdução

Com a publicação do novo acordo de Basiléia II, as instituições financeiras são encorajadas a mensurar internamente os riscos de suas atividades, entre eles, o risco de crédito. Além de ser um requerimento regulatório, o novo Acordo é um grande guia para o gerenciamento de risco, que incentiva aos bancos calcular, com base em modelos internos, o capital necessário para garantir a solvência do Sistema Financeiro.

As implicações deste acordo podem variar de acordo com o estado de desenvolvimento de cada instituição. Bancos com uma melhor estrutura estão com melhores condições de adotar as abordagens mais avançadas, já que estas exigem um sistema de informações robusto e seguro, uma cultura de gestão de risco bastante difundida e capacidade de programar controles internos e adequação de modelos de crédito já existentes.

O incentivo que as instituições financeiras têm em se adequar aos requisitos de Basiléia II é o de melhorar a precificação de seus produtos de crédito, uma vez que, atendendo aos requisitos do acordo, ganharão em competitividade, alocando seu capital de forma mais precisa do que seu concorrente que não conseguir se adequar aos modelos de rating interno.

Para as exposições classificadas nas demais categorias não varejo, onde se enquadram as grandes empresas, o BACEN (Portal do BACEN, 2011) em seu comunicado N° 18.365, permite definir como descumprimento (*default* para o restante do estudo), entre outras ocorrências, quando “a instituição financeira, por iniciativa própria e independentemente de exigência regulamentar, reconhece contabilmente deterioração significativa da qualidade do crédito”.

A motivação deste trabalho é encontrar índices econômico-financeiros e variáveis macroeconômicas, alinhada a técnicas estatísticas, que consigam reconhecer esta piora representativa com antecedência, auxiliando na precificação dos empréstimos e a mensuração do risco de crédito para fins de alocação de capital.

2. Revisão da Literatura

2.1 Modelos de previsão de falência e insolvência

Em meados da década de cinqüenta nos Estados Unidos, surgiu o interesse por desenvolver modelos de risco de crédito (Andrade, 2004). No final da década de sessenta, foi publicado um dos primeiros estudos sobre modelagem estatística de crédito para classificar o risco das empresas, desenvolvido por Edward Altman (Altman, 1968). Seu modelo media a capacidade financeira das empresas por meio de índices financeiros e econômicos com o intuito de prever o evento de falência em um determinado horizonte de tempo. Com uma amostra de 66 empresas, sendo 33 classificadas como falidas entre 1946 e 1965, ele utilizou a metodologia de análise discriminante multivariada para desenvolver um modelo conhecido como zeta-escore. O modelo mostrou uma alta capacidade de previsão, classificando 95% da amostra em questão corretamente. Estes resultados incentivaram diversos outros estudos subseqüentes.

No começo da década de oitenta, Ohlson (1980) realizou um estudo com o mesmo objetivo de previsão do evento de falência. Sua amostra era constituída de empresas industriais com ações negociadas em bolsa ou balcão entre 1970 e 1976. Um diferencial do estudo para a época foi o seu tamanho amostral. Utilizou-se de 2.163 empresas, sendo que destas, 105 faliram no período analisado.

Além disto, em substituição à análise discriminante, Ohlson utilizou um modelo de regressão logística para a previsão de falência. Ohlson (1980) estimou três modelos onde alterava o horizonte de tempo para o evento de interesse (falência ou concordata):

- Modelo 1: período de 1 ano;
- Modelo 2: período de 2 anos, caso não tenha falido no primeiro ano;
- Modelo 3: período de 1 ou 2 anos.

O primeiro modelo foi o que apresentou os melhores resultados, até corroborando uma análise prévia que indicava que o tempo médio até a falência era em torno de 13 meses. Por fim, ele sugeriu que a amostra fosse repartida em duas:

uma amostra de desenvolvimento utilizada para estimar os parâmetros, e outra para aplicar o modelo obtido e validar os resultados.

Seguindo no mesmo objetivo, Casey e Bartezak (1985) investigaram se dados de fluxo de caixa operacional melhorariam o desempenho dos modelos de previsão de insolvência. Para uma amostra de 60 empresas falidas e 230 empresas não falidas que possuíam a mesma classificação industrial das empresas falidas entre 1971 e 1982, foram utilizadas as seguintes variáveis além das tradicionais de regime de competência:

- Fluxo de caixa operacional;
- Fluxo de caixa operacional / passivo circulante;
- Fluxo de caixa operacional / exigibilidades.

Os resultados obtidos foram que estas novas variáveis não melhoraram o desempenho do modelo, ao menos para esta amostra em questão.

Em todos os trabalhos anteriores, o principal objetivo era discriminar as empresas boas das empresas ruins. Westgaard e Wijst (2001) desenvolveram para uma carteira de grandes empresas um modelo para estimar a freqüência esperada de *default* (EDF), onde *default* é o evento de falência. Para isto, eles utilizaram dados cadastrais e contábeis para construir suas variáveis. Com uma amostra de 70.574 empresas norueguesas, das quais 1.989 faliram no período de 1995 a 1999, foi possível separar os dados em duas amostras: Desenvolvimento e validação.

Westgaard e Wijst (2001) afirmam que não seria recomendável utilizar a análise discriminante para modelar a previsão de falência, já que ela não gera necessariamente uma probabilidade como resultado de sua estimação. Bandyopadhyay (2006, p. 258) apresenta a regressão logística como ferramenta para estimação da EDF diretamente, pois, além de gerar valores previstos entre 0 e 1, possui a flexibilidade de incorporar variáveis financeiras e não financeiras.

O modelo desenvolvido por Westgaard e Wijst (2001) mostrou boa capacidade na discriminação entre empresas falidas e não falidas e uma adequada previsão no cálculo da EDF. Vale ressaltar que o estudo foi realizado sobre condições de recuperação do cenário macroeconômico e por isso eles sugerem que sejam realizados estudos similares em um cenário de recessão e que com isso

sejam incorporadas variáveis macroeconômicas para que possam ser capturados efeitos do ciclo econômico.

No cenário nacional, Brito (2005) desenvolveu um modelo para previsão de inadimplência com o conceito de falência também como variável resposta. A população analisada era composta por empresas de capital aberto não financeiras, no período de 1994 e 2004. A amostra compreendeu 60 empresas, sendo 30 falidas e 30 não falidas, sendo estas emparelhadas por setor econômico e porte. A técnica utilizada foi a regressão logística.

Por fim, Akiama (2008) desenvolveu um modelo de risco de crédito para grandes empresas utilizando variáveis contábeis, cadastrais e variáveis macroeconômicas. O grande diferencial de seu estudo em relação aos anteriores foi a utilização de outro critério para definir o *default*: atraso superior a 90 dias de parcela material da dívida. Para que isso fosse possível, ela teve acesso aos dados de atraso de operações de crédito no Sistema Financeiro Nacional, provenientes do Sistema de informações de Crédito do Banco Central.

Como amostra de desenvolvimento do seu modelo, Akiama (2008) selecionou as 500 maiores empresas do país (Revista Exame) no período de 2002 a 2005, totalizando 2.000 registros, que correspondem a 638 empresas distintas. Destacando as variáveis macroeconômicas, Akiama (2008) utilizou em sua modelagem, três variáveis que poderiam influenciar no evento de inadimplência:

- Taxa de juros;
- Taxa de inflação;
- Taxa de variação do PIB.

Como resultado final, as variáveis IGPM (índice geral de preços – Mercado) e PIB (variação do PIB real) foram significantes e o modelo apresentou um alto percentual global de eventos corretamente classificados (97,4%).

2.2 Definição de insolvência

Em todos os trabalhos citados anteriormente, a definição do evento *default* foi a ocorrência de concordata ou a decretação de falência ou o atraso da dívida da

empresa com algum credor superior a 90 dias. Para o presente estudo, porém, nenhum destes conceitos poderá ser abordado devido à limitação de informação quanto ao atraso e da baixa quantidade de empresas concordatárias ou falidas no período proposto de análise (2002 – 2010).

Diante de tal dificuldade, alinhada a definição do BACEN no comunicado N° 18.365, foi preciso encontrar alternativa para identificar se uma empresa se encontra em situação de insolvência. Muitas vezes, a palavra insolvente é usada como sinônimo de dificuldade financeira, gerando alguma confusão. Wruck (1990, p. 421) define dificuldade financeira como "uma situação onde o fluxo de caixa é insuficiente para cobrir obrigações correntes. Essas obrigações podem incluir dívidas não pagas a fornecedores e a empregados, prejuízos atuais e potenciais com litígios, e *default*".

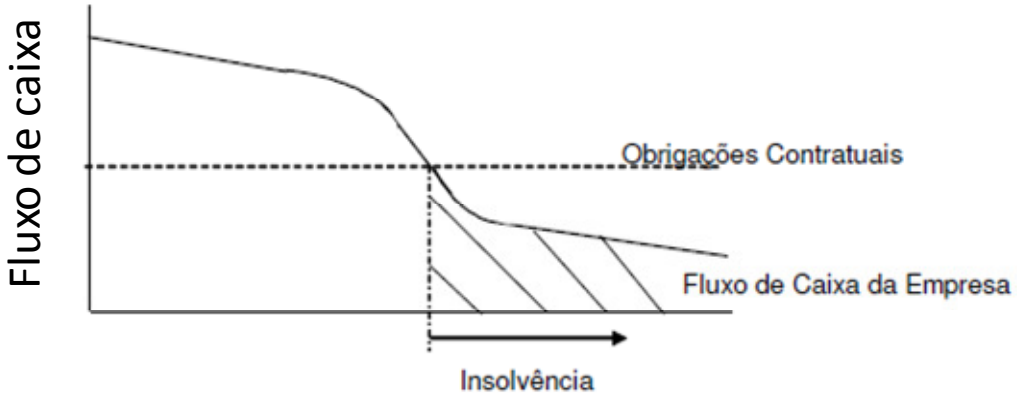
Insolvência pode ser resumida como a incapacidade de pagar dívidas. Em 1983, Altman (1983, p.6), define que é possível separar o conceito de insolvência em duas formas conceitualmente distintas:

- Insolvência baseada em fluxos (flow-based insolvency) ou insolvência técnica (technical insolvency);
- Insolvência baseada em saldos (stock-based insolvency).

Segundo Altman (1983, p.6), a insolvência técnica diz respeito a uma falta de liquidez da empresa que impossibilita o cumprimento de uma obrigação. Sendo assim, a insolvência técnica pode ser uma condição temporária. Já a insolvência baseada em saldos, segundo Wruck (1990, p.421), ocorre quando a firma tem valor econômico negativo, ou seja, **o valor presente dos fluxos futuros é menor que o total de suas obrigações**. Esse conceito de insolvência é mais grave, pois indica segundo Altman (1983) uma situação crônica e não temporária.

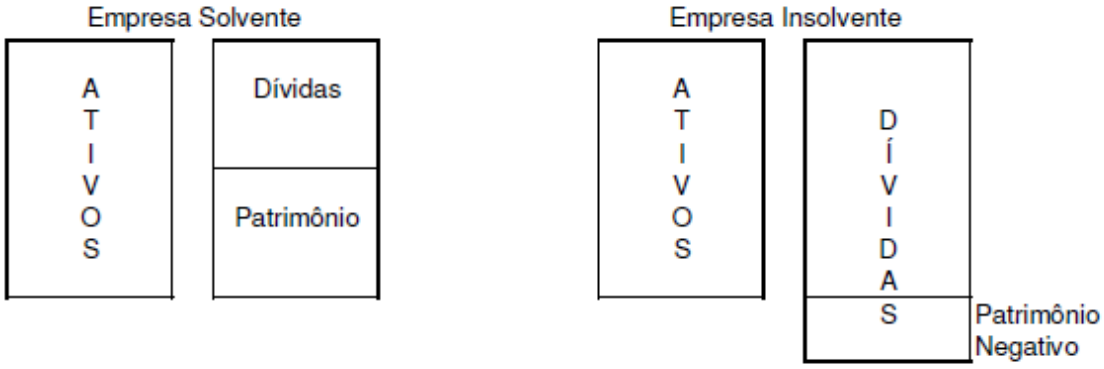
Nas figuras 1 e 2, apresentam-se as diferenças entre os conceitos de insolvência baseada em fluxos e insolvência baseada em saldos. A insolvência associada a fluxos se dá quando os fluxos de caixa são insuficientes para cobrir pagamentos determinados por contrato. Já a insolvência associada a saldos ocorre quando o valor dos ativos de uma empresa é inferior ao valor de suas dívidas, isto é, quando seu **Patrimônio Líquido (PL) é negativo** (Ross, Westerfield & Jaffe, 2002).

Figura 1: Insolvência associada a fluxos



Fonte: extraído de Ercolin (2007), adaptação de Ross, Westerfield & Jaffe (2002)

Figura 2: Insolvência associada a saldos



Fonte: extraído de Ercolin (2007), adaptação de Ross, Westerfield & Jaffe (2002)

Holmes e Haswell (1989) mostram em seus estudos que 90% de pequenas empresas australianas falidas entre 1981 e 1985 tiveram sua morte decorrente da insolvência associada a fluxo ou saldo.

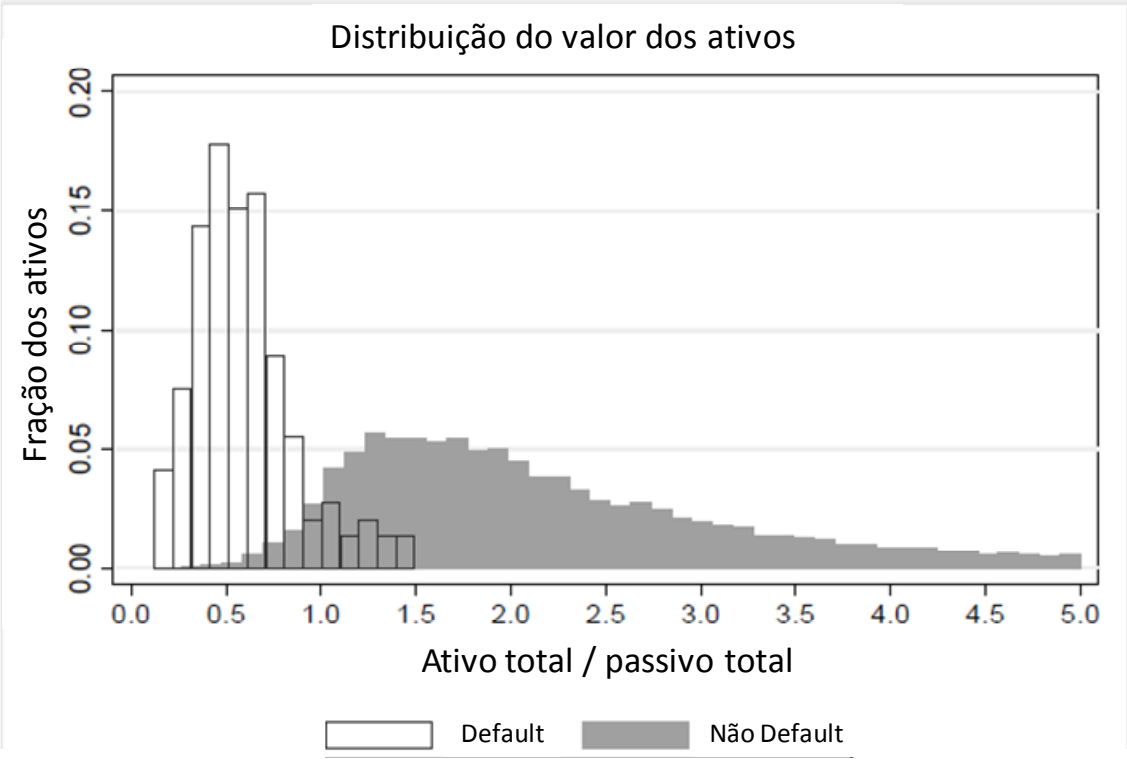
Ainda sobre este tema, Fama e French (1992) e Dichev (1998) afirmam em diferentes estudos que toda empresa com Patrimônio Líquido negativo está em crise financeira e poderá trazer implicações para os investidores. Brown et al (2008) comenta que acadêmicos e profissionais excluem as ações de empresas que possuem PL negativo de suas análises, argumentando que elas possuem um

padrão de risco muito elevado. Ele ainda acrescenta que desde o final de 1980, em torno de 5% das empresas listadas na NYSE/AMEX/Nasdaq reportam PL negativo.

Embora o PL seja amplamente utilizado na avaliação do mercado de ações, a literatura existente em contabilidade e finanças presta pouca atenção a empresas com PL negativo. É uma prática comum omitir empresas com PL negativo em estudos empíricos, por faltar interpretação econômica e dificultar o cálculo de índices financeiros, como *book-to-market* e retorno sobre o patrimônio líquido. Além disso, conforme Fama e French (1992) e Dichev (1998), citado em Ang (2010), "empresas com PL negativo são percebidas como financeiramente em dificuldades e espera-se que elas irão a falência".

A hipótese que suporta esta percepção é a de perdas persistentes, o que sugere que PL negativo é causado pela corrosão do PL que **durante anos foi impactado por resultados negativos acumulados**. Na figura 3, vemos a distribuição da razão entre ativo total e passivo total para empresas de um estudo realizado por Davydenko (2007). As empresas em default estão concentradas abaixo da razão de um, ou seja, patrimônio líquido negativo.

Figura 3: Distribuição do valor dos ativos de empresas em default e não default



Fonte: adaptação de Davydenko (2007)

Outro detalhe importante é que Wruck (1990) alerta para o uso do PL negativo no sentido contábil. Ele indica que uma empresa saudável pode apresentar no seu balanço um ativo total menor que a soma de seus passivos.

Porém, Subramanyam e Venkatachalam (2000) trazem uma importante contribuição para este assunto. Eles documentam que o valor contábil do patrimônio líquido está associado com o valor de mercado em grande parte porque capta as informações das receitas atuais e passadas, sendo assim, o PL no sentido contábil é de avaliação relevante, com particular ênfase sobre a importância relativa do estoque e do fluxo. Ainda acrescentam que, em especial para as empresas com a deterioração da saúde financeira, o valor contábil consegue capturar informações relevantes e os resultados obtidos foram consistentes.

O número crescente de empresas com PL negativo nos mercados financeiros necessitam de maiores investigações. Sob a égide dos artigos citados, e por empresas com PL negativo apresentarem um maior risco de insolvência, define-se para este trabalho, como uma *proxy* para insolvência, **o momento em que a empresa apresentar Patrimônio Líquido negativo por quatro trimestres consecutivos.**

3. Metodologia

3.1 Regressão Logística

A regressão logística (Hosmer e Lemeshow, 2000) é uma técnica muito utilizada para a gestão de risco de crédito, quando o objetivo é explicar a probabilidade de ocorrência de um evento por meio do comportamento de um conjunto de variáveis. A grande vantagem desta metodologia é que seus valores estimados serão limitados entre 0 e 1, atendendo a necessidade do estudo.

O modelo de regressão logística assume a seguinte forma:

$$p_i = \frac{e^{\alpha+\beta^T X_i}}{1+e^{\alpha+\beta^T X_i}}, \qquad (1)$$

onde, e é o número de Euler, p_i é a probabilidade de ocorrência do evento de interesse para o indivíduo i, α é o intercepto, β^T é o vetor de parâmetros (β_1, β_2,...,β_k) e X_i é o vetor de variáveis explicativas do indivíduo i, (i=1, 2, ..., n). Em geral o modelo estimado pelo método de máxima verossimilhança, sendo a função de verossimilhança dada por:

$$L(\beta) = \prod_{i=1}^{n} p_i^{y_i}(1-p_i)^{1-y_i}, \qquad (2)$$

onde, y_i é uma variável indicadora que assume o valor um se o evento de interesse ocorreu para o indivíduo *i.*

3.2 Regressão Logística Condicional

Parte deste estudo utilizou dados pareados (ver Schlesselman, 1982). Nesse caso, dispúnhamos de uma amostra de empresas insolventes (casos) que foram pareadas com empresas solventes, com mesmo perfil (controle) [1].

Uma forma muito utilizada na literatura para emparelhar empresas insolventes e solventes é utilizar o setor econômico e porte da empresa, selecionando empresas do mesmo setor e valor do ativo total próximos para a escolha da empresa que será pareada a um caso. Desta forma, cada empresa insolvente terá uma empresa solvente como par.

A regressão logística condicional (Hosmer e Lemeshow, 2000) assemelha-se muito à regressão logística apresentada em 3.1. Usada com bastante freqüência em estudos de epidemiologia, ela é recomendada para estudos de caso-controle, no qual se utiliza métodos de emparelhamento para construção da amostra final.

Considere x_{i1} o vetor de variáveis explicativas associado à i-ésima empresa insolvente e x_{i0} ao vetor de variáveis explicativas associado à empresa solvente que foi pareada com a empresa i. Seja

$$p_i = \frac{e^{\beta^T x_{i1}}}{e^{\beta^T x_{i1}} + e^{\beta^T x_{i0}}}, \qquad (3)$$

[1] Mais detalhes na seção 4.

em que p_i é a probabilidade de uma empresa com vetor de covariáveis x_{i1} ser a insolvente quando se compara um par de empresas uma com vetor de covariáveis x_{i1} e outra com vetor de covariáveis x_{i0}.

Dividindo-se o numerador e o denominador de (3) por $e^{\beta^T x_{i0}}$, chega-se a

$$p_i = \frac{e^{\beta^T x_i^*}}{1 + e^{\beta^T x_i^*}}$$

sendo $x_i^* = (x_{i1} - x_{i0})$, Desse modo, a função de verossimilhança para o modelo logístico condicional é dada por:

$$L(\beta) = \prod_{i=1}^{n} \frac{e^{\beta^T x_i^*}}{1+e^{\beta^T x_i^*}}, \qquad (4)$$

Note que (4) pode ser reescrita num formato semelhante a (2), da seguinte forma:

$$L(\beta) = \prod_{i=1}^{n} \left[\frac{e^{\beta^T x_i^*}}{1+e^{\beta^T x_i^*}}\right]^{y_i} \left[\frac{1}{1+e^{\beta^T x_i^*}}\right]^{1-y_i},$$

com $y_i = 1$.

3.3 Equações de estimação generalizadas (GEE)

O método dos modelos lineares generalizados (GLM) (Hardin e Hilbe, 2003) é uma ferramenta muito poderosa na análise de dados, já que abrange modelos de regressão com variável resposta na família exponencial de distribuições. Essa família inclui a normal, a binomial (regressão logística e probit, entre outros), Poisson, gama, gaussiana inversa, geométrica, para citar alguns (Hardin e Hilbe, 2003). Caso y seja uma variável aleatória com distribuição na família exponencial e $E(y) = \mu$, então $Var(y) = \phi V(\mu)$, sendo ϕ um parâmetro de dispersão e $V(\mu)$ uma função positiva denominada função de variância.

A teoria usual de GLM assume independência entre as observações. O Método de Equações de Estimação Generalizadas (GEE) foi desenvolvido como um meio para estender a modelagem dos GLM para o caso em que existe uma estrutura

longitudinal nos dados, acomodando assim, situações em que os dados sejam correlacionados (Hardin e Hilbe, 2003).

Admita que o comportamento do indivíduo i seja acompanhado ao longo de n_i momentos. Considere $Y_i=(y_{i1},\ldots, y_{in_i})^T$ o vetor com as n_i observações da variável resposta, $\boldsymbol{x}_{ij}$ o vetor (pX1) de covariáveis observadas para o i-ésimo indivíduo e $\boldsymbol{X}_i=(\boldsymbol{x}_{i1},\ldots, \boldsymbol{x}_{in_i})^T$ a matriz de covariáveis observadas para o i-ésimo indivíduo, i = 1,...,N. Admita também que a relação entre a média da variável resposta, $\boldsymbol{\mu}_i$, e as covariáveis, possa ser expressa por uma função de ligação, dada por (5):

$$g(\boldsymbol{\mu}_{ij}) = \boldsymbol{x}_{ij}{}^T\boldsymbol{\beta} \ , \qquad (5)$$

sendo $g(\boldsymbol{\mu}_{ij})$ uma função monótona, duplamente diferenciável e $\boldsymbol{\beta}$ um vetor paramétrico de dimensão *p*.

Nesta dissertação, $\boldsymbol{\mu}_{ij}$ é a probabilidade de uma empresa apresentar Patrimônio Líquido negativo por quatro trimestres consecutivos e a função de ligação utilizada é o logito $\left(g(\boldsymbol{\mu}_{ij}) = log\left(\frac{\boldsymbol{\mu}_{ij}}{1-\boldsymbol{\mu}_{ij}}\right)\right)$, o que implica na estimação de modelos logísticos para $\boldsymbol{\mu}_{ij}$.

As equações de estimação generalizadas foram desenvolvidas a partir da função escore construída no caso de haver independência entre as observações. Essa função escore é dada por (6).

$$\mathcal{U}(\beta) = \sum_{i=1}^{n}\left(\frac{\partial\mu_i}{\partial\beta}\right)^T V_i{}^{-1}\big(y_i - \mu_i(\beta)\big) = 0 \ , \qquad (6)$$

sendo $\mu_i(\beta) = (\mu_{i1}, \ldots, \mu_{in_i})^T$ e $V_i = D\big(V(\mu_i)\big)$, a matriz diagonal com as funções de variância de cada y_{ij}.

Para incorporar a existência de dependência entre as observações, Liang e Zeger (1986) propuseram a função de estimação dada em (7). Basicamente, eles substituíram a matriz diagonal V_i de (6) por uma matriz de covariância W_i.

$$\Psi(\beta) = \sum_{i=1}^{n}\left(\frac{\partial\mu_i}{\partial\beta}\right)^T W_i{}^{-1}\big(y_i - \mu_i(\beta)\big) = 0 \qquad (7)$$

O trabalho de Liang e Zeger (1986) inspirou-se na teoria de quase-verossimilhança de Wedderburn (1974). Basicamente, a estimação é feita levando-se em conta a distribuição marginal dos dados, o modelo para a média e uma especificação para a matriz W_i que não necessariamente corresponde à verdadeira estrutura de covariância dos dados. Note que essa matriz pode ser escrita como em (8). Nesse caso a matriz $\boldsymbol{R}(\theta)$ é uma matriz de correlação definida por meio de um vetor de parâmetros θ.

$$\mathrm{W_i} = D\big(V(\mu_i)\big)^{1/2}\boldsymbol{R}(\theta)D\big(V(\mu_i)\big)^{1/2} \qquad (8)$$

Liang e Zeger (1986) demonstram que, sob certas condições, as raízes de (7) são estimadores consistentes e assintoticamente normais do parâmetro β, mesmo quando a especificação da matriz de correlação não reflete a verdadeira estrutura de dependência existente nos dados. Por essa razão, $\boldsymbol{R}(\theta)$ é denominada de matriz de correlação de trabalho. Detalhes sobre o processo de estimação podem ser encontrados em Hardin e Hilbe (2003).

Neste estudo, serão utilizadas três especificações diferentes para a matriz de correlação de trabalho. A seguir é apresentada uma breve descrição de cada uma delas seguida de um exemplo de uma matriz de correlação de dimensão 3 x 3.

3.3.1 Correlação uniforme (*Exchangeable correlation*)

É a forma mais simples, que assume uma correlação única para todas as observações. Ela é recomendada quando as medidas repetidas não possuem dependência no tempo e qualquer permutação das medidas repetidas é válida.

$$\boldsymbol{R}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{1} & \boldsymbol{\theta} & \boldsymbol{\theta} \\ \boldsymbol{\theta} & \mathbf{1} & \boldsymbol{\theta} \\ \boldsymbol{\theta} & \boldsymbol{\theta} & \mathbf{1} \end{bmatrix}.$$

3.3.2 Correlação Auto-regressiva (*Autoregressive correlation*)

É recomendada quando se suspeita de um decaimento na correlação quando se compara observações mais distantes no tempo. Neste estudo será utilizada uma estrutura de auto-correlação de ordem 1.

$$\boldsymbol{R}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{1} & \boldsymbol{\theta} & \boldsymbol{\theta}^2 \\ \boldsymbol{\theta} & \mathbf{1} & \boldsymbol{\theta} \\ \boldsymbol{\theta}^2 & \boldsymbol{\theta} & \mathbf{1} \end{bmatrix}$$

3.3.3 Correlação não estruturada (*Unstructured correlation*)

É a mais geral, não impondo qualquer estrutura para a matriz de correlação.

$$\boldsymbol{R}(\boldsymbol{\alpha\theta}) = \begin{bmatrix} \mathbf{1} & \boldsymbol{\theta}_{12} & \boldsymbol{\theta}_{13} \\ \boldsymbol{\theta}_{12} & \mathbf{1} & \boldsymbol{\theta}_{23} \\ \boldsymbol{\theta}_{13} & \boldsymbol{\theta}_{23} & \mathbf{1} \end{bmatrix}$$

4. Base de Dados

A base de dados analisada nesta dissertação compreende empresas brasileiras de capital aberto no período de 2002 a 2010. A seguir, definimos a variável resposta.

4.1 Descrição da variável resposta

Conforme apresentado na revisão da literatura, utiliza-se neste trabalho, como *proxy* de insolvência, o momento em que a empresa apresentar Patrimônio Líquido negativo por quatro trimestres consecutivos.

Em um primeiro momento, foram encontradas 64 empresas que em algum dos balanços trimestrais apresentaram PL negativo. Estas informações foram obtidas da Economática, observando todo o universo de empresas brasileiras de capital aberto no período entre 2002 e 2010. Utilizando o critério de **PL negativo em pelo menos quatro balanços trimestrais consecutivos** esse número reduz-se a 47 empresas[2].

[2] Analisando estas 47 empresas, 37 delas não retornaram ao estado de solvência em todo o período analisado. Das 10 restantes que em algum momento conseguiram reverter a situação de insolvência, cinco empresas continuaram por mais de três anos consecutivos insolventes.

4.2 Descrição das variáveis explicativas

As variáveis utilizadas para este estudo são índices contábeis altamente difundidos na literatura de risco de crédito. Esses indicadores são capazes de descrever a situação econômico-financeira das empresas. Considerando que o evento de insolvência não é um processo que ocorre de forma instantânea, os demonstrativos contábeis possuem informações valiosas para indicar se há sinais de deterioração da empresa em questão.

A fonte das informações dos balanços das empresas é a Economática, onde foram obtidas as variáveis necessárias para a criação dos índices econômico-financeiros. Existe um grande número de índices que são comumente utilizados em modelos de previsão de inadimplência. Para este trabalho, foram selecionados 13 índices de acordo com a sua importância e disponibilidade de dados para sua construção. Esses índices podem ser classificados em seis grupos, que são apresentados na Tabela 1. No Apêndice A, apresenta-se uma breve descrição desses indicadores, assim como a construção dos mesmos, sendo que maiores detalhes podem ser encontrados em Assaf Neto (2002), Matarazzo (2003) e Silva (2003).

Tabela 1: Indicadores econômico-financeiros utilizados

Categoria	Variável	Índice
Alavancagem financeira	X1	Patrimônio líquido sobre exigível total
	X2	Endividamento total
	X3	Endividamento de curto prazo
Capacidade de pagamento	X4	Liquidez geral
	X5	Liquidez corrente
	X6	Liquidez seca
	X7	Capital de giro líquido
Cobertura de juros	X8	Lucro operacional sobre despesas financeiras
Rentabilidade	X9	Retorno sobre o ativo
	X10	Lucros retidos sobre ativo
Recursos próprios	X11	Patrimônio líquido sobre ativo
	X12	Imobilização do PL
Volume de estoque	X13	Estoques sobre ativo

Os índices econômico-financeiros têm o objetivo de identificar efeitos idiossincráticos das empresas que possam levar a situação de insolvência. Para capturar efeitos sistêmicos do período do ciclo econômico, serão analisadas algumas variáveis macroeconômicas.

A Tabela 2 apresenta as variáveis macroeconômicas que serão analisadas neste estudo.

Tabela 2: Descrição das variáveis macroeconômicas

Variável	Macroeconômica	Descrição das variáveis	Fonte
M1	IPCA	Índice Nacional de Preços ao Consumidor Amplo	IBGE
M2	Desemprego	Taxa de desemprego	IBGE
M3	Selic	Taxa de juros - Over / Selic - (% a.m.)	Bacen
M4	Inad_PJ	Inandimplência PJ (over 90) %	Bacen
M5	Prod_Fisica_Ind	Produção industrial - indústria geral - quantum - índice dessaz. (média 2002 = 100)	IBGE
M6	emp_industrial	CNI - emprego industrial índice	IBGE
M7	op_cred_pj	Operações de crédito aos setores público e privado - recursos livres - spread - PJ - (p.p.)	Bacen

4.3 Descrição dos dados

O estudo será divido em duas partes definidas como: **amostra emparelhada** e **visão carteira**. Na primeira parte, a amostra compreenderá empresas insolventes emparelhadas com empresas solventes por setor econômico e porte. A metodologia aplicada será a regressão logística condicional descrita na Seção 3.2.

Na segunda parte do estudo, será aplicada uma outra forma de amostragem. Novamente serão escolhidas empresas insolventes, só que desta vez, ao invés de emparelhadas, serão comparadas com as demais empresas de capital aberto ano a ano durante o período de 2002 a 2010. O objetivo desta análise é verificar a possibilidade de inclusão de variáveis macroeconômicas para previsão de insolvência. O horizonte de tempo de interesse será de um ano antes do evento de *default.*

4.3.1 Amostra emparelhada

Nesta abordagem foram estimados dois modelos: um utilizando informações disponíveis um ano antes do evento de insolvência e outro utilizando informações dois anos antes do evento. Assim, uma das 47 empresas foi excluída por não possuir informação dois anos antes do evento. Com isso, foram consideradas 46 empresas insolventes e 46 solventes. Encontra-se no Apêndice B a relação das empresas, seus setores e o ano que ocorreu o evento de insolvência. O critério para emparelhamento foi o de selecionar empresas solventes do mesmo setor econômico e que possuía ativo total mais próximo no período analisado[3].

As tabelas 3 e 4 apresentam estatísticas descritivas para o período de um ano e dois anos antes do evento de insolvência, separadas para as empresas solventes e insolventes. Na penúltima coluna, é apresentado o sinal esperado da diferença entre médias da amostra das empresas solventes e insolventes. Em todos os indicadores, o sinal apresentado está de acordo com a expectativa. Na última coluna, tem-se o resultado de um teste t-pareado para verificar se as médias das empresas solventes e insolventes são estatisticamente diferentes. Exceto para a variável X13, a hipótese nula de igualdade das médias foi rejeitada ($P<0,05$) tanto quando se utiliza dados um ano como 2 anos antes do evento de insolvência.

[3] Como existiam 10 empresas com setor econômico "Outros" na Economática, optou-se por verificar os resultados obtidos na seção 5.1 excluindo-as e seus respectivos pares. Estes resultados podem ser observados no APÊNDICE C.

Tabela 3: Estatística descritiva e teste t-pareado para diferença entre médias - um ano antes do evento

Variável	Solvente		Insolvente		Sinal Amostra solvente - insolvente	p-valor
	média	desvio padrão	média	desvio padrão		
X1	8.38	25.02	0.24	0.34	+	0.016
X2	0.48	0.26	0.80	0.17	-	0.000
X3	0.26	0.18	0.44	0.23	-	0.000
X4	2.66	3.94	0.68	0.49	+	0.001
X5	2.87	3.37	0.88	0.90	+	0.000
X6	2.31	3.13	0.69	0.88	+	0.001
X7	0.17	0.24	-0.10	0.23	+	0.000
X8	10.23	49.81	-2.90	13.90	+	0.044
X9	0.02	0.05	-0.07	0.10	+	0.000
X10	0.14	0.19	-0.30	0.47	+	0.000
X11	0.51	0.26	0.16	0.13	+	0.000
X12	1.92	5.52	7.64	18.42	-	0.008
X13	0.11	0.10	0.09	0.11	+	0.249

Tabela 4: Estatística descritiva e teste t-pareado para diferença entre médias – dois anos antes do evento

Variável	Solvente		Insolvente		Sinal Amostra solvente - insolvente	p-valor
	média	desvio padrão	média	desvio padrão		
X1	5.41	12.90	0.40	0.36	+	0.006
X2	0.46	0.24	0.74	0.14	-	0.000
X3	0.26	0.17	0.38	0.20	-	0.001
X4	2.01	2.43	0.75	0.51	+	0.001
X5	2.34	2.50	1.03	0.86	+	0.001
X6	1.78	1.99	0.80	0.87	+	0.003
X7	0.17	7.60	-0.03	-1.51	+	0.000
X8	8.76	28.82	-0.68	1.17	+	0.016
X9	0.03	0.06	-0.04	0.07	+	0.000
X10	0.14	0.19	-0.20	0.48	+	0.000
X11	0.52	0.25	0.25	0.15	+	0.000
X12	1.36	1.82	3.26	4.33	-	0.001
X13	0.11	0.11	0.10	0.12	+	0.368

A fim de verificar a existência de uma eventual multicolinearidade entre os indicadores financeiros, calculou-se a matriz de correlação desses indicadores considerando-se os dados um ano e dois anos antes do evento. As tabelas 5 e 6 trazem essas informações. Caso duas variáveis com alta correlação (acima de 80%) entrem no modelo, será verificada a possibilidade de exclusão de uma delas, sem que haja prejuízo significativo no resultado obtido.

Tabela 5: Matriz de correlação – um ano antes do evento

	default	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
default	100%													
X1	-23%	100%												
X2	60%	-50%	100%											
X3	39%	-33%	62%	100%										
X4	-34%	89%	-67%	-40%	100%									
X5	-38%	70%	-62%	-47%	87%	100%								
X6	-34%	69%	-56%	-45%	84%	97%	100%							
X7	-51%	8%	-55%	-53%	30%	49%	41%	100%						
X8	-18%	72%	-33%	-22%	53%	27%	28%	2%	100%					
X9	-49%	11%	-48%	-35%	23%	24%	19%	36%	21%	100%				
X10	-52%	11%	-43%	-33%	21%	21%	18%	38%	15%	41%	100%			
X11	-64%	53%	-92%	-55%	71%	66%	60%	54%	35%	46%	41%	100%		
X12	21%	-7%	26%	10%	-12%	-10%	-7%	-17%	-4%	-8%	-4%	-32%	100%	
X13	-7%	-15%	0%	26%	-5%	2%	-11%	38%	-9%	7%	9%	4%	-13%	100%

Tabela 6: Matriz de correlação – dois anos antes do evento

	default	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
default	100%													
X1	-27%	100%												
X2	58%	-60%	100%											
X3	31%	-39%	57%	100%										
X4	-34%	61%	-69%	-35%	100%									
X5	-33%	32%	-56%	-41%	82%	100%								
X6	-31%	30%	-53%	-44%	79%	97%	100%							
X7	-44%	9%	-54%	-37%	54%	68%	67%	100%						
X8	-23%	68%	-43%	-26%	35%	23%	21%	9%	100%					
X9	-49%	14%	-38%	-18%	35%	38%	33%	45%	18%	100%				
X10	-43%	10%	-40%	-7%	27%	10%	5%	25%	12%	33%	100%			
X11	-55%	61%	-98%	-55%	69%	54%	52%	52%	43%	37%	37%	100%		
X12	28%	-13%	40%	9%	-23%	-18%	-15%	-29%	-10%	-11%	-13%	-49%	100%	
X13	-3%	-16%	4%	40%	3%	7%	-9%	34%	-10%	21%	10%	-2%	-17%	100%

4.3.2 Visão carteira

Como critério para definição da amostra, primeiramente são consideradas todas as 47 empresas insolventes e suas informações contábeis são observadas até um ano antes do evento. Das empresas solventes, são incluídas no estudo todas as

que publicaram informações contábeis em todo final do exercício de cada ano entre 2001 até 2009, resultando em 177 empresas solventes.

Para tentar capturar efeitos do ciclo econômico, as características contábeis destas 177 empresas foram observadas ao longo dos nove anos, uma vez ao ano, sempre no quarto trimestre, que conjuntamente com as observações das empresas insolventes, totaliza uma amostra de 1.766 casos.

Para os índices econômico-financeiros, foi preciso realizar um tratamento nos dados para que valores aberrantes não prejudicassem a análise descritiva e posteriormente a estimação dos parâmetros. Para os valores 5% inferiores e 95% superiores de cada índice, foi adotado o valor truncado do respectivo percentil. Na Tabela 7, segue a matriz de correlação dos indicadores contábeis para a segunda parte do estudo.

Tabela 7: Matriz de correlação – índices econômico-financeiros para a visão carteira

	default	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
default	100%													
X1	-11%	100%												
X2	19%	-87%	100%											
X3	12%	-52%	62%	100%										
X4	-11%	77%	-64%	-20%	100%									
X5	-12%	64%	-59%	-47%	73%	100%								
X6	-11%	60%	-58%	-54%	57%	85%	100%							
X7	-16%	44%	-51%	-31%	65%	82%	67%	100%						
X8	-10%	42%	-40%	-22%	40%	36%	39%	31%	100%					
X9	-21%	24%	-34%	-17%	28%	26%	27%	36%	58%	100%				
X10	-24%	52%	-62%	-38%	42%	38%	36%	43%	33%	47%	100%			
X11	-19%	89%	-95%	-57%	64%	56%	55%	49%	39%	33%	65%	100%		
X12	29%	-39%	43%	-4%	-55%	-38%	-22%	-49%	-23%	-30%	-43%	-52%	100%	
X13	0%	5%	-14%	14%	21%	27%	9%	48%	9%	12%	10%	14%	-17%	100%

5. Resultados

Para este estudo, foi atribuído o valor zero para indicar se a empresa permaneceu solvente e o valor um para indicar se a empresa tornou-se insolvente.

De acordo com essa definição, a interpretação dos parâmetros é feita da seguinte forma:

- Sinal positivo indica que quanto maior o valor da variável, maior é a probabilidade de insolvência da empresa;
- Sinal negativo, de forma inversa, indica que quanto maior o valor da variável, menor é a probabilidade de insolvência da empresa.

Para chegar ao modelo final, o método utilizado foi o *backward*, onde todas as variáveis são testadas em um modelo inicial, depois são retiradas as variáveis com maior valor-p, até que as variáveis restantes sejam todas significantes a 10%.

5.1 Resultados da amostra emparelhada

5.1.1 Um ano antes do evento de insolvência

Considerando os dados para um ano antes do evento de insolvência, o modelo final foi composto por duas variáveis explicativas. As variáveis que foram significantes (P<0,10) foram os índices X5 (Liquidez corrente) e X10 (Lucros retidos sobre ativo).

Apesar de não ser a técnica recomendada, o objetivo de apresentar os resultados da regressão logística que não considera o efeito do emparelhamento (incondicional) é o de comparar os resultados com a regressão logística condicional. No quadro abaixo, podemos verificar o beta, erro-padrão e significância (p-valor) para as duas metodologias.

Tabela 8: Estimadores de parâmetros dos modelos com variáveis explicativas observadas um ano antes do evento

Variável	Índice	Regressão incondicional			Regressão condicional		
		beta	erro-padrão	p-valor	beta	erro-padrão	p-valor
X5	Liquidez corrente	-1.08	0.36	0.003	-1.81	1.08	0.094
X10	Lucros retidos sobre ativo	-10.72	2.64	0.000	-8.71	3.27	0.008
	Constante	1.06	0.49	0.030			

Conforme encontrado na análise descritiva, os coeficientes apresentaram o sinal esperado. As variáveis X5 e X10 possuem coeficientes negativos, indicando que, quanto maior os valores destas variáveis explicativas, menor será a probabilidade da empresa a possuir um PL negativo nos próximos 12 meses.

Para avaliar o ajuste de um modelo logístico, podemos observar a medida Log likelihood e o Pseudo R^2. Em ambas as medidas, quanto maior o valor, mais bem ajustado será o modelo. No quadro abaixo, vemos que as duas medidas são melhores com o uso da regressão logística condicional.

Tabela 9: Medidas para avaliar ajuste dos modelos logísticos

Metodologia	Log likelihood	Pseudo R^2
Regressão incondicional	-28.83	0.55
Regressão condicional	-6.63	0.79

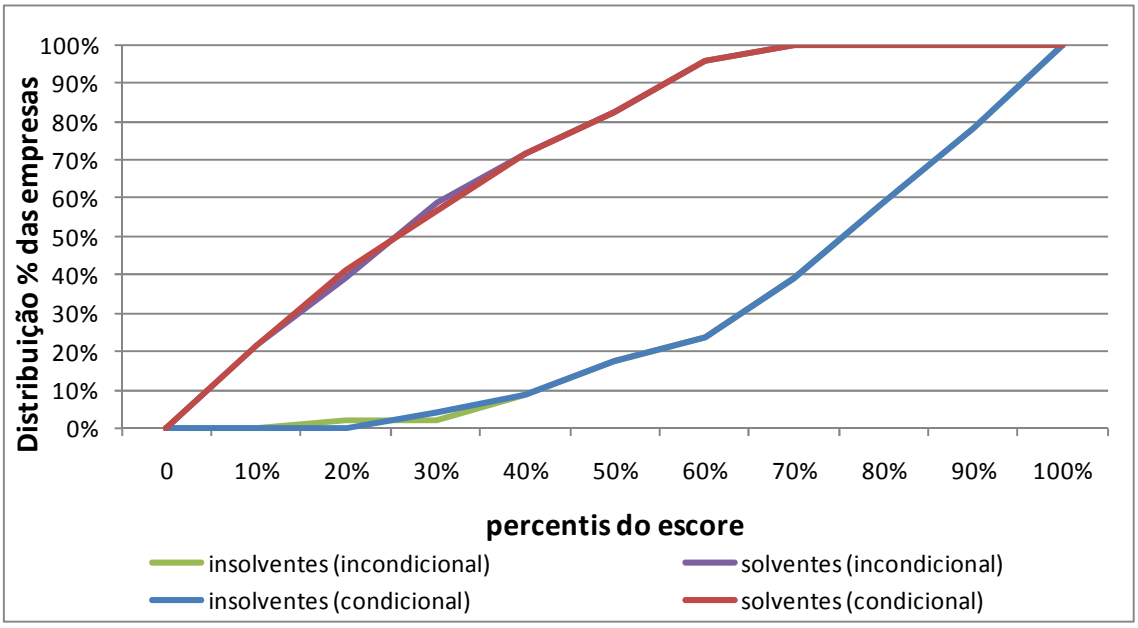
A capacidade de previsão do modelo pode ser avaliada através dos dados da Tabela 10. A partir dos valores previstos pelo modelo, são criados intervalos tendo os decis desses valores como limites. Em cada intervalo, observou-se quantas eram solventes e quantas eram insolventes. O esperado é que as empresas insolventes estejam mais concentradas nos maiores escores, indicando maior probabilidade de insolvência, enquanto as empresas solventes apresentem menor escore, indicando menor probabilidade de insolvência. Para ambos modelos ajustados, é exatamente isto que é observado.

Tabela 10: Distribuição por percentis do escore – modelo condicional com variáveis explicativas observadas um ano antes do evento

percentis	escore máximo	insolventes	solventes	taxa de insolvência
10%	0.00	0	10	0.0%
20%	0.00	0	9	0.0%
30%	0.01	2	7	22.2%
40%	0.11	2	7	22.2%
50%	0.16	4	5	44.4%
60%	0.27	3	6	33.3%
70%	0.42	7	2	77.8%
80%	0.71	9	0	100.0%
90%	0.88	9	0	100.0%
100%	1.00	10	0	100.0%
Total		**46**	**46**	**50.0%**

Além da tabela acima, um indicador muito utilizado tanto na literatura como por consultorias e grandes empresas na avaliação de um bom modelo de risco de crédito é o Kolmogorov-Smirnov (KS). Essa medida corresponde à máxima distância entre a distribuição acumulada das empresas insolventes e da distribuição acumulada das empresas solventes (Lecumberri e Duarte, 2003). Como prática de mercado, um KS acima de 50% indica uma excelente discriminação do modelo. Para as duas metodologias, o KS obtido foi de 73,9%, indicando uma excelente discriminação das empresas solventes e insolventes. A curva da distribuição acumulada de solventes e insolventes dos dois modelos pode ser observada na figura abaixo.

Figura 4: Distribuição acumulada de insolventes e solventes – um ano antes do evento



5.1.2 Dois anos antes do evento de insolvência

Para o segundo horizonte de tempo, o modelo final foi composto por três variáveis explicativas. As variáveis que foram significantes (P<0,05) foram os índices X1 (Patrimônio líquido sobre exigível total), X4 (Liquidez geral), X8 (Lucro operacional sobre despesas financeiras). A Tabela 11 traz um quadro com as estimativas dos parâmetros, erro-padrão e significância (p-valor) para as duas metodologias.

Tabela 11: Estimadores de parâmetros dos modelos com variáveis explicativas observadas dois anos antes do evento

Variável	Índice	Regressão incondicional			Regressão condicional		
		beta	erro-padrão	p-valor	beta	erro-padrão	p-valor
X4	Liquidez geral	1.70	0.79	0.031	3.75	1.81	0.039
X8	Lucro operacional sobre despesas financeiras	-1.24	0.36	0.001	-1.22	0.54	0.023
X1	Patrimônio líquido sobre exigível total	-2.58	0.87	0.003	-4.57	1.99	0.022
	Constante	0.38	0.52	0.465			

Conforme encontrado na análise descritiva, os coeficientes apresentaram sinal conforme o esperado, exceto pelo índice X4 (Liquidez geral). Era esperado

também um coeficiente negativo, mas por estar fortemente correlacionada com as demais variáveis, houve inversão de sinal. As variáveis X1 e X8 possuem coeficientes negativos, indicando que, quanto maior os valores destas variáveis explicativas, menor será a probabilidade da empresa a possuir um PL negativo nos próximos dois anos.

Abaixo, um quadro com as medidas de ajuste dos modelos de regressão logística. Novamente, o modelo estimado por regressão logística condicional apresentou o melhor ajuste.

Tabela 12: Medidas para avaliar ajuste dos modelos logísticos

Metodologia	Log likelihood	Pseudo R2
Regressão incondicional	-35.52	0.44
Regressão condicional	-8.73	0.73

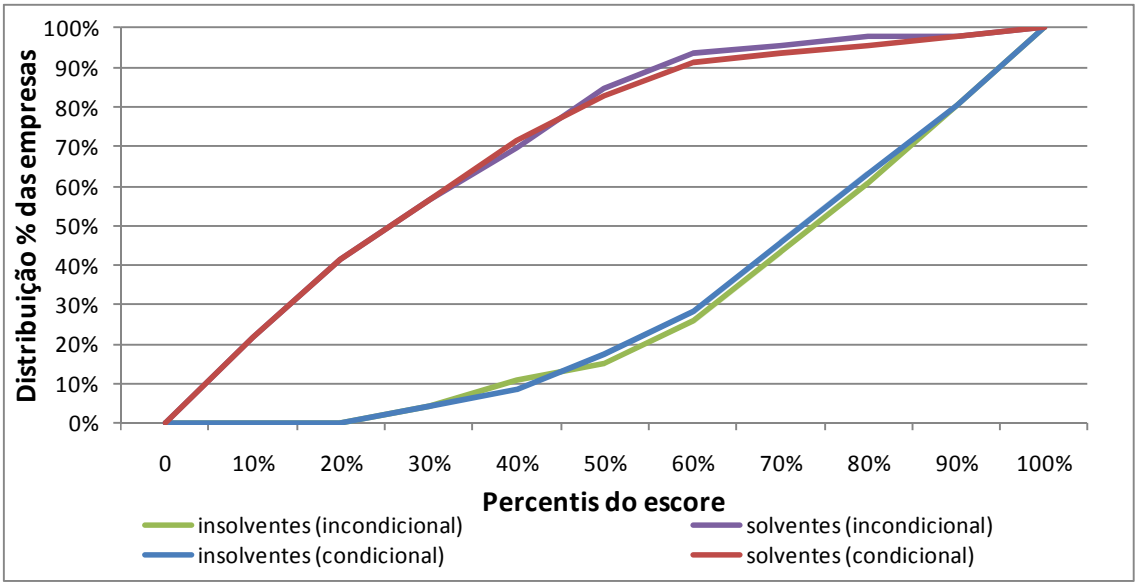
A capacidade de previsão do modelo pode ser avaliada por meio da Tabela 13. De forma análoga ao primeiro horizonte de tempo, foram criados percentis do valor estimado (escore) e acompanhado a distribuição das empresas solventes e insolventes. Novamente, os resultados foram satisfatórios para ambos modelos.

Tabela 13: Distribuição por percentis do escore – modelo condicional com variáveis explicativas observadas dois anos antes do evento

percentis	escore máximo	insolventes	solventes	taxa de insolvência
10%	0.00	0	10	0.0%
20%	0.02	0	9	0.0%
30%	0.15	2	7	22.2%
40%	0.46	2	7	22.2%
50%	0.65	4	5	44.4%
60%	0.78	5	4	55.6%
70%	0.84	8	1	88.9%
80%	0.89	8	1	88.9%
90%	0.96	8	1	88.9%
100%	1.00	9	1	90.0%
0	**0**	**46**	**46**	**50.0%**

A seguir, o gráfico da Curva KS. Para o modelo ajustado por regressão logística incondicional, o KS obtido foi de 72%, já para o modelo ajustado por regressão logística condicional foi de 70%, indicando uma excelente discriminação das empresas solventes e insolventes, por ambos os modelos.

Figura 5: Distribuição acumulada de insolventes e solventes – dois anos antes do evento



5.2 Resultados da visão carteira

Para a segunda amostra, será aplicada a metodologia descrita na Seção 3.3 (Equações de Estimação Generalizadas).

Foram ajustados modelos utilizando-se estrutura de correlação uniforme, auto-regressiva e não estruturada. Após a aplicação do método *backward* de seleção de variáveis, nos três modelos, foram selecionadas (P<0,10) as seguintes variáveis: X9 (Retorno sobre o ativo), X10 (Lucros retidos sobre ativo), X11 (Patrimônio líquido sobre ativo) e X12 (Imobilização do PL).

A Tabela 14 traz as estimativas dos parâmetros desses modelos. Os resultados dos modelos foram muito próximos.

Tabela 14: Estimativas GEE

Variável	Índice	GEE – Uniforme			GEE - Não estruturado			GEE - Auto-regressivo		
		beta	erro-padrão	p-valor	beta	erro-padrão	p-valor	beta	erro-padrão	p-valor
X9	Retorno sobre o ativo	-13.22	3.82	0.001	-13.23	3.81	0.001	-11.71	3.94	0.003
X11	Patrimônio líquido sobre ativo	-3.60	1.88	0.056	-3.60	1.88	0.056	-2.77	1.87	0.138
X10	Lucros retidos sobre ativo	-7.84	2.19	0.000	-7.84	2.19	0.000	-6.55	2.23	0.003
X12	Imobilização do PL	0.38	0.14	0.007	0.38	0.14	0.007	0.38	0.15	0.013
	Constante	-3.64	0.70	0.000	-3.64	0.70	0.000	-3.86	0.73	0.000

As variáveis X9, X10 e X11 apresentaram coeficientes negativos, indicando que, quanto maior os valores destas variáveis explicativas, menor será a probabilidade da empresa a possuir um PL negativo nos próximos 12 meses. Já o índice X12 apresentou coeficiente positivo, indicando que quanto maior o valor assumido por esta variável, maior será a probabilidade da empresa possuir um PL negativo nos próximos 12 meses[4].

Para comparar o ajuste das diferentes metodologias, não podemos observar as mesmas medidas da primeira amostra, já que o ajuste por GEE não é baseado na máxima verossimilhança. Uma saída é utilizar a estatística de Wald (Engle, 1983). Quanto maior o valor, mais bem ajustado será o modelo. No quadro abaixo, vemos a estatística para as quatro abordagens. Aparentemente, o uso da matriz de correlação com estrutura auto-regressiva não apresentou bom ajuste.

Tabela 15: Estatística de Wald para os modelos GEE

Metodologia	Wald
GEE - uniforme	83.74
GEE - Não estruturado	83.76
GEE - Auto-regressivo	65.34

[4] Para verificar se uma possível quebra estrutural na economia em 2008 poderia afetar a estimação deste modelo, no APÊNDICE D encontra-se o mesmo modelo estimado para o período anterior a 2008.

Como os resultados apresentados pelas abordagens uniforme e não estruturada foram semelhantes, não apresentaremos mais resultados referentes à estrutura uniforme. A Tabela 16 traz a distribuição das observações por intervalos definidos por percentis dos valores previstos (escores) das empresas solventes e insolventes.

Tabela 16: Distribuição por percentis do escore

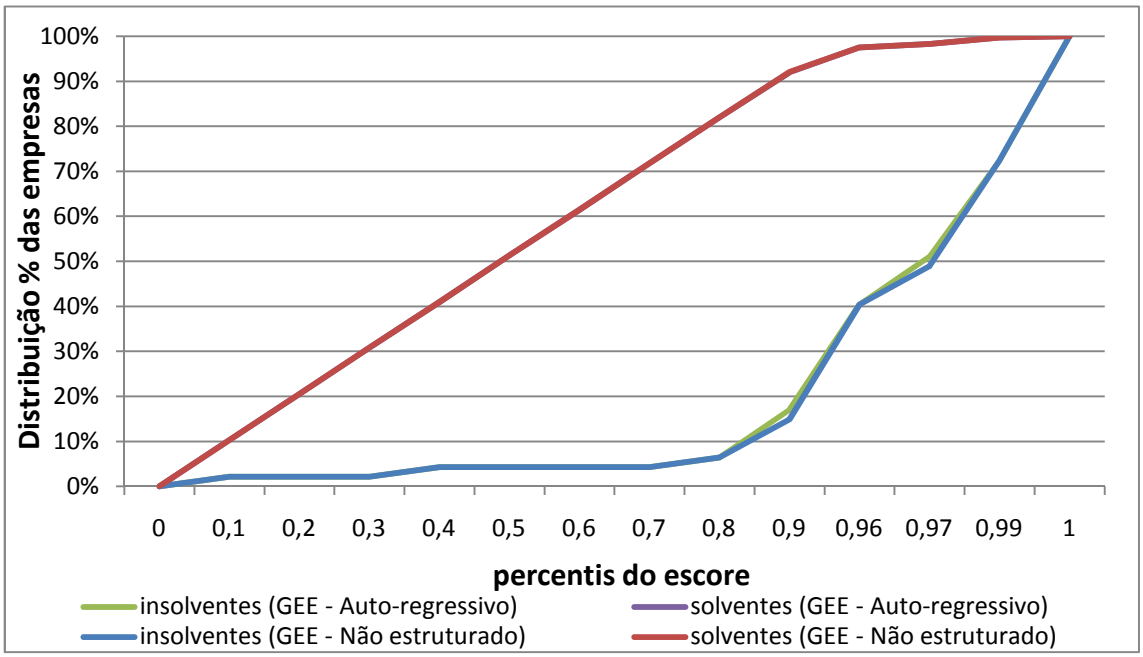
	GEE - Não estruturado				GEE - Auto-regressivo			
percentis	**escore máximo**	**insolventes**	**solventes**	**taxa de insolvência**	**escore máximo**	**insolventes**	**solventes**	**taxa de insolvência**
10%	0.000	1	176	0.6%	0.000	1	176	0.6%
20%	0.000	0	176	0.0%	0.000	0	176	0.0%
30%	0.000	0	177	0.0%	0.001	0	177	0.0%
40%	0.001	1	175	0.6%	0.002	1	175	0.6%
50%	0.002	0	177	0.0%	0.003	0	177	0.0%
60%	0.003	0	176	0.0%	0.005	0	176	0.0%
70%	0.008	0	177	0.0%	0.009	0	177	0.0%
80%	0.017	1	175	0.6%	0.016	1	175	0.6%
90%	0.068	4	173	2.3%	0.058	5	172	2.8%
96%	0.205	12	94	11.3%	0.165	11	95	10.4%
97%	0.265	4	14	22.2%	0.207	5	13	27.8%
99%	0.418	11	24	31.4%	0.325	10	25	28.6%
100%	1.000	13	5	72.2%	1.000	13	5	72.2%
		47	**1719**	**2.66%**		**47**	**1719**	**2.66%**

A Tabela 16, para o modelo ajustado por GEE não estruturado, indica que empresas com escore até 0,017 possuem menos de 1% de chance de ser tornar insolvente no próximo ano. A proporção de empresas insolventes com escore entre 0,017 e 0,068 é de 2,3%. À medida que o escore aumenta, vemos que a proporção de empresas com PL negativo em 12 meses aumenta. Empresas com escore entre 0,418 e 1.000 possuem 72,2% de chance de se tornar insolvente em 12 meses, enquanto a taxa média da amostra é de apenas 2,66% de empresas insolventes.

Para o modelo ajustado por GEE não estruturado, o KS obtido foi de 78%, já para o modelo ajustado por GEE com estrutura auto-regressiva foi de 77%,

indicando uma excelente discriminação das empresas solventes e insolventes para ambas as metodologias.

Figura 6: Distribuição acumulada de insolventes e solventes



Como próximo passo, as sete variáveis macroeconômicas foram incluídas na modelagem, uma de cada vez, devido à alta correlação entre elas. O modelo final é composto pelas quatro variáveis explicativas, uma macroeconômica mais o intercepto. Além das variáveis anteriores, a variável M7 (Operações de crédito aos setores público e privado - recursos livres - spread - PJ - (p.p.)) foi significante ao nível de 10% para o modelo ajustado por GEE não estruturado.

A Tabela 17 apresenta as estimativas obtidas com a inclusão da variável macroeconômica. A variável M7 apresentou coeficiente negativo indicando que, quanto maior o crédito aos setores público e privado com recursos livres à Pessoa Jurídica, menor será a probabilidade das empresas se tornarem insolventes nos próximos 12 meses. As demais variáveis mantiveram os sinais observados na Tabela 14.

Tabela 17: Estimativas GEE

Variável	Índice	GEE - Não estruturado			GEE - Auto-regressivo		
		beta	erro-padrão	p-valor	beta	erro-padrão	p-valor
X9	Retorno sobre o ativo	-14.23	3.85	0.000	-12.54	3.99	0.002
X11	Patrimônio líquido sobre ativo	-3.75	1.87	0.045	-2.94	1.87	0.117
X10	Lucros retidos sobre ativo	-7.02	2.19	0.001	-6.02	2.23	0.007
X12	Imobilização do PL	0.39	0.14	0.006	0.39	0.15	0.011
M7	Op_cred_pj	-26.04	11.33	0.022	-18.00	11.48	0.117
	Constante	0.11	1.73	0.948	-1.24	1.79	0.489

A Tabela 18 compara o ajuste das diferentes metodologias, com e sem a inclusão da variável M7. Vemos que inclusão da variável M7 aumenta a estatística de Wald, melhorando o ajuste dos dois modelos[5].

Tabela 18: Estatística de Wald para os modelos GEE

Metodologia	Wald	
	Sem M7	Com M7
GEE - Não estruturado	83.74	88.23
GEE - Auto-regressivo	65.34	67.55

Além da tabela acima, foi calculada a estatística KS. Para os dois modelos, mesmo após a inclusão da variável macroeconômica, o KS obtido foi de 78%.

O intuito de ter uma variável macroeconômica no modelo é que mesmo que os efeitos idiossincráticos das empresas não se alterem, a conjuntura da economia possa afetar a probabilidade de insolvência das empresas.

Para verificar se os modelos ajustados para esta segunda amostra foram capazes de capturar a taxa de insolvência observada, comparamos as taxas de insolvência observadas com as esperas segundo o modelo, entre 2002 e 2010. Este resultado pode ser visto na Tabela 19 e na Figura 7. Na Tabela 19, é acrescentado o Erro Quadrático Médio (EQM) para ter uma medida resumo comparativa entre os modelos. Todos se mostraram capazes de capturar os efeitos desta amostra

[5] No APÊNDICE E, foi incluída na modelagem uma *dummy* para casa ano analisado, em substituição da variável macroeconômica. Como resultado, verificou-se que a inclusão das *dummies* de ano não alteram os parâmetros dos demais índices econômico-financeiros de forma significativa, além de demonstrar uma menor estatística de WALD.

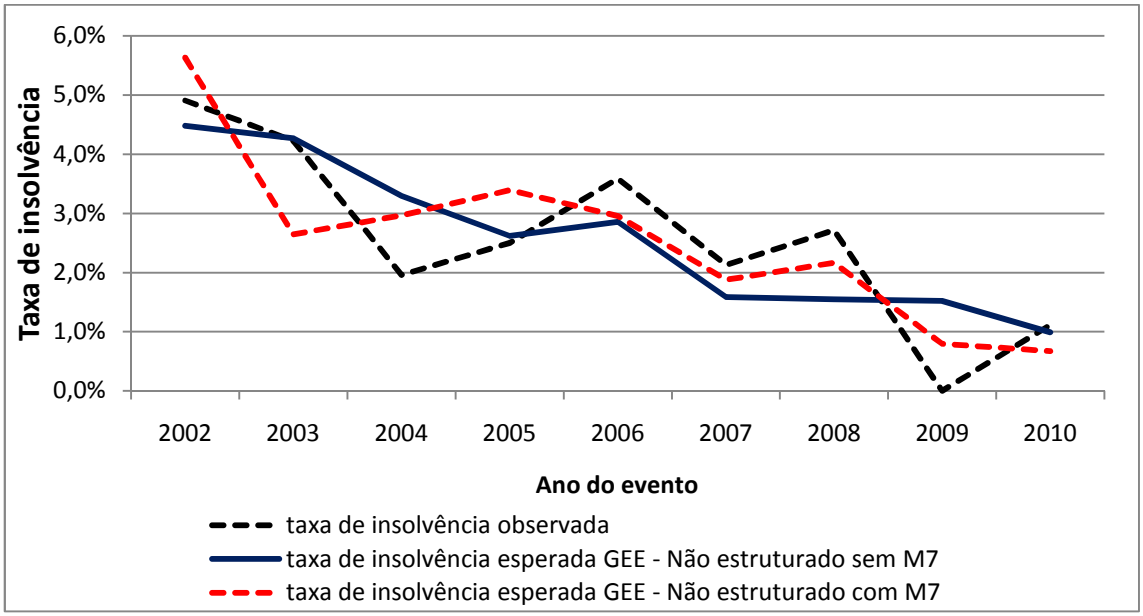
proposta, sendo que o modelo ajustado por GEE auto-regressivo apresentou o menor EQM.

Tabela 19: Taxa de insolvência esperada x observada

ano do evento	solventes	insolventes	Total	taxa de insolvência observada	taxa de insolvência esperada			
					GEE - Não estruturado		GEE - Auto-regressivo	
					sem M7	com M7	sem M7	com M7
2002	213	11	224	4.9%	4.5%	5.6%	4.4%	5.5%
2003	204	9	213	4.2%	4.3%	2.6%	4.0%	3.0%
2004	200	4	204	2.0%	3.3%	3.0%	3.4%	3.0%
2005	195	5	200	2.5%	2.6%	3.4%	2.8%	3.2%
2006	188	7	195	3.6%	2.9%	3.0%	2.7%	2.8%
2007	184	4	188	2.1%	1.6%	1.9%	1.6%	1.9%
2008	179	5	184	2.7%	1.6%	2.2%	1.6%	2.0%
2009	179	0	179	0.0%	1.5%	0.8%	1.6%	1.0%
2010	177	2	179	1.1%	1.0%	0.7%	1.1%	0.7%
Total	**1719**	**47**	**1766**	**2.66%**	**2.66%**	**2.66%**	**2.66%**	**2.66%**
				EQM	**0.065%**	**0.064%**	**0.072%**	**0.057%**

Na Figura 7, observa-se graficamente a comparação da taxa de insolvência observada no estudo com a taxa de insolvência esperada do modelo ajustado por GEE não estruturado, com e sem a presença da variável M7.

Figura 7: Taxa de insolvência observada x esperada



6. Conclusão

Neste trabalho foi sugerida uma *proxy* para insolvência: apresentar PL negativo durante mais de quatro balanços trimestrais consecutivos. Esta definição foi motivada pela pouca quantidade de empresas falidas ou concordatárias no período que compreende entre 2002 e 2010.

Como visto na revisão da literatura, a abordagem por amostras emparelhadas é bastante utilizada no desenvolvimento de modelos de previsão de insolvência. Porém, o emparelhamento em si não é comumente considerado na aplicação da regressão logística. Este estudo apresenta a metodologia de regressão logística condicional, recomendada para estudos com amostras emparelhadas.

Os índices que conseguiram explicar a *proxy* de insolvência com um ano de antecedência ao evento foram os de liquidez corrente (X5) e lucros retidos sobre ativo (X10) enquanto que os índices de Patrimônio líquido sobre exigível total (X1), liquidez geral (X4) e Lucro operacional sobre despesas financeiras (X8) foram os que melhores explicaram a *proxy* de insolvência com dois anos antes do evento. Os resultados encontrados foram satisfatórios, mostrando um alto índice de discriminação entre empresas solventes e insolventes, com KS superior a 70%.

Com o objetivo de incluir variáveis macroeconômicas no modelo de previsão de insolvência, foi sugerida uma abordagem diferente da amostra emparelhada, na qual as empresas são acompanhadas ano a ano, simulando uma carteira de crédito. Para considerar a correlação dos dados longitudinais, a metodologia estudada foi a de Equações de Estimação Generalizadas (GEE). Com índices que resumem a rentabilidade (X9 e X10) e os recursos próprios das empresas (X11 e X12), constatou-se uma adequada discriminação, com KS superior a 77%, uma boa ordenação no escore obtido, além de identificar faixas de escore onde empresas têm em torno de 10%, 20%, 30% ou 70% de probabilidade de insolvência nos próximos 12 meses, servindo como uma boa ferramenta no julgamento do risco destas empresas.

A inclusão da variável macroeconômica “Operações de crédito aos setores público e privado - recursos livres - spread - PJ - (p.p.)” no modelo não apresentou grande melhora na capacidade de previsão de insolvência no período analisado.

Por fim, recomenda-se para estudos futuros, a aplicação das metodologias analisadas neste estudo para carteiras com maior incidência de empresas insolventes, como empresas de menor porte.

APÊNDICE A

O índice X1 avalia a proporção de recursos próprios em relação aos recursos de terceiros na estrutura de capital. Os índices X2 e X3 são indicadores de estrutura que avaliam o grau de endividamento da empresa, X3 mais específico no curto prazo.

Os índices X4, X5 e X6 são indicadores comuns de liquidez, medindo a capacidade de pagamento da empresa. X7 é um indicador de análise dinâmica que avalia a situação financeira da empresa.

O índice X8 é um indicador de cobertura de juros que mede a capacidade da empresa para pagar esses encargos. X9 e X10 são indicadores de rentabilidade, avaliando o resultado gerado pela empresa e o volume de lucros retidos pela empresa em relação ao ativo, respectivamente.

Os índices X11 e X12 avaliam como está sendo utilizada a parcela de recursos próprios da empresa. O primeiro mede a parcela dos ativos que é financiada por recursos próprios, o segundo mede a parcela dos mesmos que está comprometida com o ativo permanente.

Por último, o índice X13 mede o volume de estoques que a empresa tem em relação ao ativo total. De forma resumida, pode-se observar na tabela abaixo a construção destes indicadores de acordo com as contas do balanço patrimonial.

Descrição dos índices econômico-financeiros

Categoria	Variável	Índice	Descrição das variáveis
Alavancagem financeira	X1	Patrimônio líquido sobre exigível total	Patrimônio Líquido / (Passivo circulante + Exigível a longo prazo)
	X2	Endividamento total	(Patrimônio Líquido + Exigível a longo prazo) / Ativo total
	X3	Endividamento de curto prazo	Passico circulante / Ativo total
Capacidade de pagamento	X4	Liquidez geral	(Ativo circulante + Realizável a longo prazo) / (Passivo circulante + Realizável a longo prazo)
	X5	Liquidez corrente	Ativo circulante / Passivo circulante
	X6	Liquidez seca	(Ativo circulante - Estoques) / Passivo circulante
	X7	Capital de giro líquido	(Ativo circulante - Passivo circulante) / Ativo total
Cobertura de juros	X8	Lucro operacional sobre despesas financeiras	Lucros antes dos Juros e Impostos de renda / Despesas financeiras
Rentabilidade	X9	Retorno sobre o ativo	Lucros antes dos Juros e Impostos de renda / Ativo total
	X10	Lucros retidos sobre ativo	(Patrimônio Líquido - Capital Social) / Ativo total
Recursos próprios	X11	Patrimônio líquido sobre ativo	Patrimônio Líquido / Ativo total
	X12	Imobilização do PL	Ativo permanente / Patrimônio Líquido
Volume de estoque	X13	Estoques sobre ativo	Estoques / Ativo total

APÊNDICE B

Relação das empresas solventes e insolventes para a primeira amostra

Insolvente	Setor	*Default*	Solvente
Telet S.A.	Telecomunicações	2002	Tele Nordeste Celular Participacoes S/A
Cia Fabricadora de Peças	Veiculos e peças	2002	Mahle Cofap Aneis S/A
SA Ind e Comercio Chapeco	Alimentos e Beb	2002	SA Fabrica Prods Alimcs Vigor
Companhia Energetica do Maranhao - Cemar	Energia Elétrica	2002	Empresa Energetica de M.S. S.A.
Celg Distribuicao S/A - Celg D	Energia Elétrica	2002	Cia Est Geracao e Trans de En El Ceee-Gt
Wiest S/A	Veiculos e peças	2002	Fras-Le S/A
Inds Arteb S/A	Veiculos e peças	2002	Dana-Albarus S.A. Industria e Comercio
Sauipe S/A	Outros	2002	Dijon S/A
Lark S/A Maq e Equipamentos	Outros	2002	Baumer S/A
Paranapanema S/A	Siderur & Metalur	2002	Arcelormittal Inox Brasil S/A
Douat Cia Textil	Textil	2002	Empresa Nac Com Redito e Part. S/A Encorpar
D F Vasconc SA Opt Mec A Prec	Outros	2002	Grucai Participacoes S/A
Brasil Ferrovias S/A	Transporte Serviç	2002	All - America Latina Logistica S/A
Ligas de Aluminio S/A - Liasa	Minerais não Met	2003	Portobello S/A
Bombril S/A	Química	2003	Petroflex Industria e Comercio S.A.
Hopi Hari S/A	Outros	2003	Technos Relogios S/A
Textil Renauxview S/A	Textil	2003	Döhler S/A
Cimob Participacoes S/A	Construção	2003	Construtora Sultepa S/A
Industrias Micheletto S/A	Siderur & Metalur	2003	Metalgrafica Iguacu S/A
Kuala S.A.	Textil	2003	Companhia Tecidos Santanense
Cia Docas de Imbituba	Transporte Serviç	2003	Santos-Brasil S/A
Recrusul S/A	Veiculos e peças	2003	Bicicletas Monark S/A
Buettner S/A - Industria e Comercio	Textil	2004	Marisol S/A
Teka Tecelagem Kuehnrich S/A	Textil	2004	Santista Textil S.A.
Gazola S/A Industria Metalurgica	Siderur & Metalur	2004	Fibam Cia Industrial
Eldorado	Outros	2005	Blue Tree Hotels & Resorts do Brasil S/A
Manasa Madeireira Nacional SA	Agro e Pesca	2005	Rasip Agro Pastoril S/A
Ceramica Chiarelli S/A	Minerais não Met	2005	Nadir Figueiredo Ind. e Com. S/A
Hoteis Othon S/A	Outros	2006	Petropar S/A
Arthur Lange S/A Ind e Com	Outros	2006	Sondotecnica Engenharia de Solos S/A
Telecom Brasileiras S/A	Telecomunicações	2006	Amazonia Celular S.A.
Sansuy S/A Industria de Plasticos.	Outros	2006	Ideiasnet S/A
Kepler Weber S/A	Siderur & Metalur	2006	Forjas Taurus S/A
Fiacao e Tecelagem Sao Jose S/A	Textil	2006	Karsten S.A.
Fab Tecidos Carlos Renaux S/A	Textil	2006	Cia. de Fiacao e Tecidos Cedro Cachoeira
IGB Eletronica S/A	Eletroeletrônicos	2007	Brasmotor S/A
Biomm S/A	Outros	2007	Polpar S/A
Const Adolpho Lindenberg S/A	Construção	2007	Sergen Servicos Gerais de Engenharia S/A
Docas Investimentos S/A	Outros	2007	Saraiva S/A Livreiros Editores
Livraria do Globo S/A	Comércio	2008	Minasmaquinas S/A
Refinaria de Petroleo de Manguinhos S/A	Petróleo e Gas	2008	Cia. Distrib. de Gas do Rio de Janeiro

Tecblu - Tecelagem Blumenau S/A	Textil	2008	Companhia Industrial Cataguases
Botucatu Textil S/A	Textil	2008	Pettenati S/A Ind Textil
Metalurgica Riosulense S/A	Veiculos e peças	2008	Schulz S/A
Construtora Beter S/A	Construção	2010	Const Lix da Cunha S/A
La Fonte Telecom S/A	Telecomunicações	2010	LF Tel S/A

APÊNDICE C

Regressões retirando as 10 empresas do setor econômico “Outros” definido pela Economática.

Regressão Logística Condicional: um ano antes do evento.

Variável	Regressão condicional Seção 5.1.1			Excluindo setor "Outros"		
	beta	erro-padrão	p-valor	beta	erro-padrão	p-valor
X5	-1.81	1.08	0.094	-1.75	1.06	0.099
X10	-8.71	3.27	0.008	-8.42	3.30	0.011

Regressão Logística Condicional: dois anos antes do evento.

Variável	Regressão condicional Seção 5.1.2			Excluindo setor "Outros"		
	beta	erro-padrão	p-valor	beta	erro-padrão	p-valor
X4	3.75	1.81	0.039	3.45	1.78	0.053
X8	-1.22	0.54	0.023	-1.12	0.54	0.038
X1	-4.57	1.99	0.022	-4.28	1.93	0.026

APÊNDICE D

Comparativo do modelo da Seção 5.2, considerando informações até 2008.

Índice	GEE - Não estruturado			GEE - Não estruturado (até 2008)		
	beta	erro-padrão	p-valor	beta	erro-padrão	p-valor
Retorno sobre o ativo	-13.22	3.82	0.001	-14.39	4.02	0.000
Patrimônio líquido sobre ativo	-3.60	1.88	0.056	-3.80	1.95	0.052
Lucros retidos sobre ativo	-7.84	2.19	0.000	-6.43	2.19	0.003
Imobilização do PL	0.38	0.14	0.007	0.39	0.15	0.011
Constante	-3.64	0.70	0.000	-3.48	0.74	0.000

Metodologia	Wald
GEE - Não estruturado	83.74
GEE - Não estruturado (até 2008)	77.12

APÊNDICE E

Comparativo do modelo da Seção 5.2, incluindo *dummies* de ano em substituição a variável macroeconômica.

Variável	Índice	GEE - Não estruturado			GEE - Não estruturado (*Dummies*)		
		beta	erro-padrão	p-valor	beta	erro-padrão	p-valor
X9	Retorno sobre o ativo	-14.23	3.85	0.000	-14.34	3.92	0.000
X11	Patrimônio líquido sobre ativo	-3.75	1.87	0.045	-3.81	1.89	0.043
X10	Lucros retidos sobre ativo	-7.02	2.19	0.001	-7.44	2.23	0.001
X12	Imobilização do PL	0.39	0.14	0.006	0.40	0.14	0.006
M7	Op_cred_pj	-26.04	11.33	0.022	-	-	-
	Constante	0.11	1.73	0.948	-3.55	0.96	0.000
D2001	*Dummy* de 2001	-	-	-	0.06	0.84	0.939
D2002	*Dummy* de 2002	-	-	-	-0.25	0.85	0.765
D2003	*Dummy* de 2003	-	-	-	-0.77	0.93	0.411
D2004	*Dummy* de 2004	-	-	-	-0.13	0.91	0.885
D2005	*Dummy* de 2005	-	-	-	0.33	0.89	0.705
D2006	*Dummy* de 2006	-	-	-	0.26	0.94	0.782
D2007	*Dummy* de 2007	-	-	-	0.97	0.91	0.287
D2008	*Dummy* de 2008	-	-	-	-24.26	0.92	1.000

Metodologia	Wald
GEE - Não estruturado	88.23
GEE - Não estruturado (Dummies)	84.75

Referências Bibliográficas

AKIAMA, Simone R. **Probabilidade de inadimplência de grandes empresas no sistema financeiro nacional**. São Paulo, 2008 150f. Dissertação (Mestrado em Controladoria e Contabilidade) – Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, São Paulo, 2008.

ALTMAN, Edward. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, **Journal of Finance** 23, 589–609, 1968.

ALTMAN, Edward. **Corporate Financial Distress**: a complete guide to predicting, avoiding, and dealing with bankruptcy. New York: John Wiley and Sons, 1983. 356p.

ANDRADE, Fabio W. M. **Desenvolvimento de modelo de risco de portfólio para carteiras de crédito a pessoas físicas**. São Paulo, 2004. Tese (Doutorado em Administração) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas, São Paulo, 2004.

ANG, Tze Chuan. **Are Firms with Negative Book Equity in Financial Distress?** Australia, 2010. Department of Finance, Level 2, Economics and Commerce Building, University of Melbourne. 2010.

ASSAF NETO, Alexandre. **Estrutura e análise de balanços**. 7. ed. São Paulo: Atlas, 2002.

BABDYOPADHYAY, Arindam. Predicting probability of default of Indian corporate bons: logistic and Z-score model approaches. **Journal of Risk Finance***,* [S.1], v. 7, n.3, p. 255 – 272, 2006

BRITO, Giovani A. S. **Mensuração de risco de portfólio para carteira de crédito a empresas**. São Paulo, 2005. 83 f. Dissertação (Mestrado em Controladoria e Contabilidade) – Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, São Paulo, 2005.

BROWN, Stephen, LAJBCYGIER, Paul and LI, Bob. Going negative: What to do with negative book equity stocks? **Journal of Portfolio Management** 35, 95–102. 2008.

CASEY, Cornelius; BARTCZAK, Norman. Using operation cash flow data to predict financial distress: some extensions. **Journal of Accounting Research**, [S.1], v. 23, n. 1, Spring 1985.

DAVYDENKO, Sergei A. **When do Firms Default?** A Study of the Default Boundary, Canada, 2007. University of Toronto, 2007.

DICHEV, Ilia. Is the risk of bankruptcy a systematic risk? **Journal of Finance** 53, 1131 – 1147. 1998.

ENGLE, Robert F. Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics. In M.D. Intriligator; Z. Griliches. **Handbook of Econometrics**. II. Elsiever. pp. 796–801. 1983.

ERCOLIN, Carlos Alberto, **Fatores financeiros determinantes da mortalidade de micro e pequenas empresas**. São Paulo, 2007. 136 f. Dissertação (Mestrado em Administração) – Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, São Paulo, 2007.

FAMA, Eugene and FRENCH, Kenneth, The cross-section of expected stock returns, **Journal of Finance** 47, 427-465. 1992.

HARDIN, James W.; HILBE, Joseph M. **Generalized Estimating Equations**, 1st ed. Florida: Chapman & Hall. 2003.

HOLMES, S.; HASWELL, S. Estimating the business failure rate: a reappraisal. **Journal of Small Business Management**. V.27, n3, p.68 (7). 1989.

HOSMER, D.W. e LEMESHOW, S. **Applied Logistic Regression**. 2nd ed. New York: Wiley. 2000.

LECUMBERRI, L.F.L.;DUARTE, A.M. Uma metodologia para o gerenciamento de modelos de escoragem em operações de crédito de varejo no Brasil. **Revista de Economia Aplicada**, São Paulo, v.7. n.4, PP. 795-818, 2003.

LIANG, K.Y. and ZEGER, S.L. Longitudinal data analysis using generalized linear models. **Biometrika** 73: 13-22. 1986.

MATARAZZO, Dante C. **Análise financeira de balanços**. 6. ed. São Paulo: Atlas, 2003.

OHLSON, James, Financial ratios and the probabilistic prediction of bankruptcy, **Journal of Accounting Research** 18, 109–131. 1980.

Portal do BACEN. Disponível em: http://www.bcb.gov.br/htms/Normativ/COMUNICADO18365.pdf.

Acesso em 21.05.2011.

ROSS, S. A.; WESTERFIELD, R. W. e JAFFE, J. F. **Administração financeira**: Corporate finance. 2. Ed, São Paulo: Atlas, 2002.

SILVA, José Pereira da. **Gestão e análise de risco de crédito**. 4. ed. São Paulo: Atlas, 2003.

SCHLESSELMAN, J. J. **Case-Control Studies**: Design, Conduct, Analysis. New York: Oxford University Press, 1982.

SUBRAMANYAM, K.R. and VENKATACHALAM, M. **Why is Book Value Relevant for Equity Valuation?** Assets-in-Place vs Past Earnings. University of Southern California, 2000.

WEDDERBURN, R.W.M. Quasi-likelihood functions, generalized linear models, and the Gauss—Newton method. **Biometrika** 61 (3): 439–447. doi:10.1093/biomet/61.3.439, 1974.

WESTGAARD, Sjur; WIJST, Nico van der. Default probabilities in a corporate bank portfolio: A logistic model approach. **European Journal of Operational Research**, [S.1.], v135, p. 338-349, 2001.

WRUCK, K. H. Financial distress, reorganization, and organizational efficiency. **Journal of Financial and Economics**, v. 27, p. 419-444, 1990.