

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

Mateus Cyrineu Munhoz

**AVALIAÇÃO DO RISCO DE CRÉDITO DE EMPRESAS:
UMA ABORDAGEM COM DADOS EM PAINEL**

**São Paulo
2010**

Mateus Cyrineu Munhoz

**Avaliação do risco de crédito de empresas:
uma abordagem com dados em painel**

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 Macroeconomia Aplicadas
Orientador: Prof. Dr. Naercio Aquino Menezes Filho
– Insper

**São Paulo
2010**

Munhoz, Mateus Cyrineu

Avaliação do risco de crédito de empresas: uma abordagem com dados em painel/ Mateus Cyrineu Munhoz; orientador Naercio Aquino Menezes Filho – São Paulo: Insper, 2010.

54 f.

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

1. Ratings de crédito 2. Modelos de resposta ordenada
3. Dados em painel

FOLHA DE APROVAÇÃO

Mateus Cyrineu Munhoz

Avaliação do risco de crédito de empresas: uma abordagem com dados em painel.

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

Área de concentração: Finanças e Macroeconomia Aplicadas

Aprovado em: Julho/2010

Banca Examinadora

Prof. Dr. Naercio Aquino Menezes Filho
Orientador

Instituição: Insper

Assinatura: _____

Prof. Dr. Abraham Laredo Sicsú

Instituição: Fundação Getúlio Vargas

Assinatura: _____

Prof^a.Dra. Andrea Maria Accioly Fonseca Minardi

Instituição: Insper

Assinatura: _____

Dedico esse trabalho a Miguel.

AGRADECIMENTOS

Agradeço a todas as pessoas que de alguma forma, direta ou indiretamente, me ajudaram na produção desse trabalho.

Primeiramente agradeço ao Banco Itaú-Unibanco, que através de seu programa de incentivo a formação de seus colaboradores, me apoiou desde o início nessa jornada que se iniciou em meados de 2008. Em especial a Sergio Werlang e Roberta Geyer por acreditarem que, diante da incerteza, a racionalidade na tomada de decisão é o melhor caminho a ser trilhado.

Agradeço aos meus pais pelo respeito às minhas escolhas e por sempre incentivarem, acreditarem e colocarem a educação como prioridade em nossas vidas. Aos meus irmãos pelo carinho, e por promoverem encontros alegres, musicais, e sempre acompanhados de um bom churrasco no sagrado espaço *Eden pauper et saturabantur*.

Aos colegas da quinta turma do mestrado em economia pelo espírito mosqueteiro nos momentos difíceis e descontraídos que passamos juntos.

Ao professor e orientador Naercio em aceitar a orientação desse trabalho e pela sua sabedoria e didática no ensinamento baseado pela intuição.

Ao mestre Claudio Alvarenga pelo condicionamento rígido e disciplinado no desenvolvimento de modelos. E a Claudia Dornelles pela imensa paciência em me ensinar o passo a passo do desenvolvimento de um modelo de *credit scoring*.

A Frida e Zen pelos momentos alegres e amor incondicional à sua família. A Alice, que veio ao mundo em um especial momento e dando o grande presente que um homem pode receber que é ser chamado de “papá”.

E finalmente a Joana, pelo amor, carinho, amizade, e me apoiar nessa jornada, respeitando minhas ausências nos finais de semanas e madrugadas, e sempre presente em me ouvir e participar sobre a cada nova empolgação e aprendizado.

RESUMO

MUNHOZ, Mateus Cyrineu. **Avaliação do risco de crédito de empresas: uma abordagem com dados em painel.** 2010. 54 f. Dissertação (Mestrado) – Insper Instituto de Ensino e Pesquisa, São Paulo, 2010.

Impulsionada pelo crescimento da demanda por crédito, as instituições financeiras estão abandonando a análise de crédito tradicional e adotando novas técnicas no gerenciamento do porta-fólio baseando-se em modelos quantitativos. Este trabalho estuda as diferentes abordagens para avaliar o risco de crédito das empresas utilizando dados em painel. Foram testados modelo lineares e não-lineares, considerando a variável resposta como ordinal e também como cardinal. Os resultados em todos os modelos foram semelhantes. Os principais fatores determinantes na avaliação do risco das empresas foram: tamanho da empresa, retorno sobre ativos, volatilidade e EBIT / receita líquida.

Palavras-chave: Ratings de crédito; modelos de resposta ordenada; dados em painel

ABSTRACT

MUNHOZ, Mateus Cyrineu. **Assessing the credit risk of companies: a panel data approach.** 2010. 54 f. Dissertation (Mastership) – Insper Instituto de Ensino e Pesquisa, São Paulo, 2010.

Driven by the growth of credit demand, financial institutions are adopting new techniques to manage portfolio: the traditional credit analysis gives place to quantitative models. This paper studies different approaches to assess credit risk of companies using panel data. Linear and nonlinear models were tested considering ordinal and cardinal response variables. The results were similar in all tested models. The main factors in determining credit risk were: company size, return on assets, stock volatility and EBIT/net income.

Keywords: Credit ratings; ordered response models; panel data

LISTA DE TABELAS

Tabela 2 – Análise descritiva das variáveis independentes vs classes de <i>rating</i>34
Tabela 3 - Modelos em <i>Cross-Section</i>35
Tabela 4 - Modelos de Dados em Painel.37
Tabela 5 - Recodificação das Variáveis Resposta.38
Tabela 6 - Modelos Não-Lineares de Efeitos Fixos.39
Tabela 7 - Estatísticas dos coeficientes dos modelos não-lineares de efeitos fixos .	.39

LISTA DE FIGURAS

Figura 1 – Resumo das definições de Rating.	27
Figura 2 – Análise de Classificação da Moody's de uma Empresa Industrial.	28
Figura 3 – Análise gráfica de correlação das variáveis.	34

LISTA DE GRÁFICOS

Gráfico 1 – Taxas de Inadimplência Moody's.	29
Gráfico 2 – Distribuição das empresas por rating.	31
Gráfico 3 – Quantidade de Empresas por Período.	32
Gráfico 4 – Distribuição da Quantidade de Empresas que possuem dados no maior número de períodos possíveis.	32
Gráfico 5 – Distribuição das Mudanças de Ratings da Moody's.	33

SUMÁRIO

1. Introdução.	12
2. Revisão da Literatura.	14
3. Metodologia Econométrica.	18
3.1. Modelos Lineares.	18
3.1.1. Modelos Lineares com Dados em Cross-Section.	18
3.1.2. Modelos Lineares com Dados em Painel.	19
3.2. Modelos Não Lineares com Resposta Binária.	22
3.2.1. Modelo Probit Binário com Dados em Cross-Section.	22
3.2.2. Modelo Probit Binário com Dados em Painel.	23
3.3. Modelos Não Lineares com Resposta Ordenada.	24
3.3.1. Modelo Probit Ordenado com Dados em Cross-Section.	25
3.3.2. Modelo Probit Ordenado com Dados em Painel.	26
4. Descrição dos Dados.	27
4.1. Agências de Rating.	27
4.2. Dados.	29
4.3. Correlação: variáveis independentes versus rating.	33
5. Resultados.	35
6. Conclusão.	41
Referências.	42
Apêndice.	45

1. INTRODUÇÃO

O gerenciamento do risco de crédito é um dos elementos fundamentais no comportamento do mercado financeiro. Diariamente, as decisões financeiras das empresas de negócios, de governos e, em particular, das instituições financeiras são calcadas de instrumentos de gestão de risco.

Quando corremos um risco, apostamos em resultados que serão consequências da decisão tomada, embora não seja certo qual será o resultado dessa decisão. O elemento 'incerteza' nos leva a tomar decisões racionais, e a aceitação do risco consiste em acreditarmos que essa incerteza pode estar a nosso favor.

A essência da administração do risco está em maximizar as áreas onde temos algum controle sobre o resultado, enquanto minimizamos as áreas onde não temos absolutamente nenhum controle sobre o resultado e onde o vínculo entre efeito e causa está oculto de nós (BERNSTEIN, 1997, p. 197).

Algumas vezes, a informação de que dispomos para aplicar as leis da probabilidade é restrita. Até podemos reunir quantidades imensas de informações, mas, diante da incerteza, nunca conseguimos juntar cem por cento de todas elas. Kenneth Arrow, ganhador do prêmio Nobel, sabia dessa limitação:

Para mim, nosso conhecimento do funcionamento das coisas, na sociedade ou na natureza, vem a reboque de nuvens de imprecisão. Grandes males têm se seguido a uma crença na certeza, seja na inevitabilidade histórica, em grandiosos projetos diplomáticos ou em visões extremas na política econômica. No desenvolvimento de políticas com amplos efeitos sobre um indivíduo ou a sociedade é preciso cautela, pois não podemos prever as consequências (Arrow apud Bernstein, 1997).

O reconhecimento da existência da incerteza e a utilização da informação disponível de forma eficiente e ótima são essenciais na tomada de decisão. Não tomamos nenhuma decisão sem referência a um passado que compreendemos com certo grau de certeza e um futuro que não temos nenhum conhecimento certo.

Com o amadurecimento da noção da existência da incerteza e o crescimento temporal das informações disponíveis, cada vez mais se faz necessário o desenvolvimento de metodologias mais sofisticadas no controle e gerenciamento do risco de crédito. Em 1988, o Banco para Compensações Internacionais (BIS¹) determinou algumas regras para dimensionamento de capital em âmbito mundial para cobrir riscos de crédito. Esse foi o primeiro esforço internacional para lidar com a crescente exposição das instituições financeiras a riscos. A evolução dessa regulamentação, desde então, foi bastante rápida.

Um dos produtos principais dessa evolução da regulamentação foi o Novo Acordo de Basileia II, no qual é exigido que instituições financeiras se adequem a uma série de requisitos. Um deles trata especificamente de modelos regulatórios internos para uma alocação de capital adequadamente sensível ao grau de risco envolvido nas posições de um banco. Algumas instituições financeiras, quando avaliam o risco de crédito para empresas de médio porte, se deparam muitas vezes com o desafio de desenvolver modelos robustos, devido a baixa quantidade histórica de inadimplência (*low default portfólio*).

Esta pesquisa tem como objetivo principal estimar um modelo de *credit rating*, ferramenta amplamente utilizada na gestão do risco de crédito. O modelo empírico proposto é baseado em Minardi, Sanvincente e Artes (2007), que sugerem uma metodologia para estimar *credit rating* através de um modelo logito ordenado. Nesse estudo, os autores utilizam uma amostra de dados em *cross-section*. A contribuição dessa pesquisa é avaliar o impacto de um melhor aproveitamento das informações disponíveis utilizando a dimensão temporal dos dados. O modelo econométrico utilizado é uma regressão cuja variável resposta é categórica e ordenada, e a amostra, ao invés de uma *cross section*, é de dados em painel.

¹ BIS = Bank of International Settlement

2. REVISÃO DA LITERATURA

Os primeiros modelos de crédito direcionados para a classificação do risco do devedor começaram a ser desenvolvidos na década de cinquenta nos Estados Unidos para avaliação do crédito ao consumidor (Andrade, 2004). Esse tipo de ferramenta de gestão de risco de crédito ficou conhecido como modelos de *credit scoring* (Lewis, 1992).

Na década seguinte, Edward Altman, um dos pioneiros no estudo de modelos estatísticos de classificação de risco para empresas, desenvolveu um modelo chamado de zeta-escore (Altman, 1968). Utilizando-se da metodologia de análise discriminante múltipla, o zeta-escore é uma medida de solidez financeira da empresa. O propósito do estudo foi analisar índices financeiros e econômicos para prever falência². Os dados utilizados por Altman foram grandes empresas manufatureiras. A função discriminante resultante de seu estudo é:

$$Z = 0,012*X1 + 0,014*X2 + 0,033*X3 + 0,006*X4 + 0,999*X5$$

onde X1 = capital de giro / ativo total

X2 = lucro retido / ativo total

X3 = lucro antes de impostos e taxas / ativo total

X4 = valor de mercado da empresa / dívida total

X5 = vendas / ativo total

Z = índice geral

O ponto de corte do índice geral (escore) que melhor discrimina as firmas que foram à falência das que não foram é 2,675. Dessa forma, empresas com escore acima desse ponto são classificadas como boas e abaixo como ruins. O modelo mostrou-se extremamente acurado classificando corretamente 95% da amostra total. Os resultados encontrados foram extremamente encorajadores.

Com o crescimento da demanda por crédito, as instituições financeiras estão abandonando a tradicional análise de crédito caso a caso e passando a fazer gestão do porta-fólio de crédito baseando-se em modelos estatísticos, econométricos e de

² A insolvência, ou falta de solidez financeira, refere-se a firmas que entraram legalmente em falência.

finanças. A partir da década de noventa modelos financeiros mais estruturados foram desenvolvidos para previsão de inadimplência passando a indústria bancária a desenvolver metodologias próprias para o gerenciamento do risco (Saunders, 2000). Alguns dos modelos de risco de crédito mais conhecidos são: *CreditMetrics*, criado no banco J. P. Morgan (Gupton et al, 1997) no qual são utilizadas metodologias de marcação a mercado, e o KMV, modelo que leva o nome da empresa de consultoria que o criou (KMV, 1993), que utiliza a Teoria de Opções dos modelos de Merton³.

O *CreditMetrics* é um modelo para mensuração do risco de crédito de carteiras relativo a empréstimos ou títulos baseado em valores de mercado. Ele busca avaliar a distribuição de valor do porta-fólio em um determinado período de tempo através da probabilidade de migração do *rating*. Dessa forma, a aplicação do *CreditMetrics* depende do uso de modelos de classificação de risco, que pode ser um *rating* interno da instituição ou *rating* provido por uma agência de classificação de riscos.

O KMV é baseado no modelo de estrutura de capital da firma proposto por Merton (1974), no qual uma empresa é considerada em *default* quando o valor de seus ativos é inferior ao de seus passivos. Sendo uma extensão da abordagem de Black-Scholes, esse modelo é baseado na abordagem de avaliação neutra ao risco (*risk neutral valuation*) e determina a perda da carteira, definida como a diferença entre o valor de mercado e o valor do porta-fólio, gerando a probabilidade de *default* que é chamada de EDF – *Expected Default Frequency*. A aplicação do KMV, assim como o *CreditMetrics*, enfrenta dificuldades no Brasil pois poucas são as empresas negociadas em bolsa e com razoável liquidez.

Os modelos financeiros não são os mais adequados do ponto de vista de alocação de capital regulatório quando, em particular, a gestão de risco está sob a ótica do Novo Acordo de Basiléia II. Alguns modelos possuem limitações de aplicabilidade e outros utilizam de metodologias que não refletem necessariamente um modelo de probabilidade de inadimplência. No contexto de Basiléia II as avaliações de risco precisam ser expressas em termos de probabilidade de

³ Análises comparativas dos modelos de risco de porta-fólio podem ser encontradas em Crouhy *et al.* (2000), Smithson (2003) e Gordy (1998)

inadimplência e precisam incorporar certo grau de conservadorismo a fim de refletir uma gestão de riscos prudente. No caso de carteiras de crédito que não possuem *default*, ou poucos *defaults* ao longo dos anos, o resultado dos modelos pode incorrer em estimativas muito próximas de zero, o que claramente não envolveria um conservadorismo suficiente.

Existem algumas abordagens na literatura para tratar dessa particularidade, sendo a de Pluto & Tasche (2005) a mais popular. Baseado em Vasicek (1997) para estimar a probabilidade de *default* em carteiras com baixa incidência de clientes inadimplentes, utiliza uma técnica que consiste em atribuir probabilidade de *default* por níveis de risco dado um nível de confiança. O método é baseado na quantidade de observações e do número de eventos em *default*, utilizando como premissa a distribuição binomial e considerando a correlação entre os ativos, conforme descrito abaixo.

Para um nível de confiança α escolhido, estima-se a probabilidade de inadimplência que satisfaça a seguinte equação:

$$\int_{-\infty}^{+\infty} \sum_{k=0}^d \binom{N}{k} f(\Phi, Y, \rho) \left[1 - f(\Phi, Y, \rho)^{N-k} \Phi(\Phi) \right] dY = 1 - \alpha$$

onde d é a quantidade de empresas em situação de inadimplência, p é a probabilidade de inadimplência para o nível de risco, ρ sendo a correlação entre os valores dos ativos e

$$f(\Phi, Y, \rho) \equiv p(\text{default} | Y) \equiv N \left(\frac{N^{-1} \Phi_j \sqrt{\rho Y}}{\sqrt{1 - \rho}} \right)$$

Em geral os resultados geram probabilidades superestimadas, onde a probabilidade média estimada do porta-fólio é superior à taxa de inadimplência observada para a carteira. Para corrigir esse viés, Pluto e Tasche (2005) propõem um método de calibração que consiste em escalonar as estimativas ao redor da taxa de inadimplência observada.

Outra abordagem usualmente praticada pelo mercado de risco de crédito é utilizar *ratings* de agências externas. No entanto, muitas empresas, especialmente no caso do Brasil, não são avaliadas por essas agências. Como resposta a essa dificuldade prática, Servigny e Renault (2004) propõem uma alternativa factível para o desenvolvimento de sistemas internos de avaliação de risco que consiste em replicar essas classificações atribuídas pelas empresas de classificação.

Essa abordagem já havia sido utilizada por Kaplan e Urwitz (1979). Os autores utilizaram modelos logito ordenados para estimar *credit ratings* e analisaram variáveis referentes a índices de cobertura, de capitalização, tamanho da firma, variáveis da estabilidade como coeficiente da variabilidade total do recurso e coeficiente da variabilidade do lucro; e variáveis de mercado: coeficiente beta e o resíduo da regressão do modelo de mercado. De acordo com os autores, o risco ou o erro específico da regressão podem ser interpretados como um *proxy* para a habilidade da gerência.

Minardi, Sanvicente e Artes (2007) usaram dados da Económica para obter dados contábeis e informação de mercado de seiscentas e vinte e sete empresas norte americanas do setor industrial e os sites das agências de classificação da Moody's e S&P para capturar seus respectivos *credit ratings*. Foram calculadas e testadas dezesseis variáveis explanatórias, fazendo referência a: tamanho da empresa, índices de solvência, variáveis de desempenho operacional e de estabilidade.

As variáveis seleccionadas para compor o modelo final foram: *logaritmo neperiano do ativo total da empresa, dívida sobre o ativo total, lucro antes dos efeitos financeiros e impostos sobre a renda dividido pela dívida líquida, retorno sobre o ativo total, lucro antes dos efeitos financeiros e impostos sobre a renda dividido pela renda líquida* e, por último, a *volatilidade do ativo*. Os autores encontraram que *ativo total, dívida sobre ativo total, retorno sobre ativo total e volatilidade* foram estatisticamente significantes e os sinais dos coeficientes de acordo com o esperado. As duas variáveis referentes ao *lucro antes dos efeitos financeiros e impostos sobre a renda* não foram significativos a um nível de 5%, e o sinal dos coeficientes contra intuitivos.

A motivação para esse trabalho originou-se da necessidade das instituições financeiras desenvolverem um modelo de avaliação de risco próprio capaz de prever a probabilidade de inadimplência em um contexto de exigências regulatórias e também em situações onde praticamente inexistente o evento de interesse, ou seja, carteiras com baixo índice de inadimplência. A abordagem com dados em painel surgiu de forma natural, acreditando que a inserção de informações temporais pode trazer ganhos nos coeficientes estimados.

3. METODOLOGIA ECONOMÉTRICA

Frijters et al. (2004) e Ferrer-i-Carbonell et al. (2004) investigam a robustez de modelos nos quais a variável resposta é ordenada comparando duas abordagens econométricas: a primeira considerando a variável resposta como sendo cardinal, e a segunda considerando a variável resposta como sendo ordinal. Os autores estimam um modelo de variável latente para os indivíduos através do modelo logito ordenado de efeitos fixos. Este modelo é bastante similar ao modelo logito de efeito fixo desenvolvido por Chamberlain (1980). A principal conclusão é que a suposição de cardinalidade ou ordenação na variável resposta traz pequenas diferenças nos resultados. Faremos agora uma revisão dos principais modelos econométricos relevantes para esse trabalho.

3.1. Modelos Lineares

3.1.1. Modelo Linear com Dados em Cross-Section

Considere uma amostra de n observações. Seja y_j , $j = 1, 2, \dots, n$, observações aleatórias independentes que assumem, para o j -ésimo elemento da amostra, qualquer valor da família dos números reais.

Seja x_j o vetor ($p \times 1$) de variáveis independentes para o indivíduo j . O modelo linear é definido por:

$$y_j = \beta x_j' + u_j$$

sendo β um vetor de parâmetros ($1 \times p$).

O estimador de mínimos quadrados ordinários é:

$$\hat{\beta} = \left(\sum_{j=1}^N x_j' x_j \right)^{-1} \left(\sum_{j=1}^N x_j' y_j \right)$$

3.1.2. Modelo Linear com Dados em Painel

Nos dados em painel uma mesma unidade amostral é observada ao longo do tempo. Ou seja, y_{jt} representa a j -ésima firma no período de tempo t , onde $j = 1, 2, \dots, n$ e $t = 1, 2, \dots, T$. A estrutura de um modelo linear em dados em painel é:

$$y_{jt} = \beta x'_{jt} + v_{jt}$$

onde $v_{jt} = c_j + \varepsilon_t$.

O termo c_j é o efeito específico da firma.

A estimação de β desconsiderando a existência do efeito específico c_j considera os dados como uma grande *cross-section*, ou seja, é obtida pelo estimador de mínimos quadrados ordinários.

Em um modelo tradicional do tipo *cross-section* o efeito fixo c_j não é estimado, e a suposição de sua existência é admitir que os erros v_{jt} e v_{jt-1} são correlacionados, pois ambos os erros $v_{jt} = c_j + \varepsilon_t$ e $v_{jt-1} = c_j + \varepsilon_{t-1}$ possuem um termo comum c_j .

Dessa forma, a estrutura dos dados em painel sugere a utilização de metodologias mais adequadas, pois a hipótese de que o erro em cada período não está correlacionado com as variáveis explicativas no mesmo período é muito forte. Então a principal motivação para usar metodologias de dados em painel é a presença de variáveis omitidas c_j , e desconsiderá-las causa inconsistência e viés nos estimadores. O primeiro questionamento que deve ser feito nessa abordagem com dados em painel é: esses efeitos específicos das empresas são fixos ou aleatórios? Ou seja, os efeitos c_j devem ser tratados como uma variável aleatória ou um parâmetro a ser estimado?

Modelo de Efeitos Aleatórios

Para discutir esses dois modelos é fundamental entender as hipóteses que os sustentam. No modelo de efeitos aleatórios a hipótese principal é

$$E(u_{jt} | x_j, c_j) = 0, t = 1, \dots, T \text{ e}$$

$$E(u_j | x_j) = 0.$$

Já no modelo de efeitos fixos a hipótese principal é apenas

$$E(u_{jt} | x_j, c_j) = 0, t = 1, \dots, T.$$

A diferença está no fato de que o modelo de efeitos aleatórios exige que os efeitos específicos não sejam correlacionados com as variáveis explicativas (efeitos aleatórios) *versus* que são correlacionados (efeitos fixos). Por outro lado, o modelo de efeitos aleatórios é mais eficiente visto que a quantidade de parâmetros a serem estimados é bastante inferior ao de efeitos fixos.

O modelo de Efeitos Aleatórios assume que o erro idiossincrático possui uma variância incondicional constante, isto é, $E(u_{it}^2) = \sigma_u^2, t = 1, 2, \dots, T$ e que os erros idiossincráticos são serialmente não-correlacionados $E(u_{it}, u_{is}) = 0$, para todo $t \neq s$.

Somando estas hipóteses é possível definir $\Omega = E(u_i u_i')$ através de uma estrutura de efeitos aleatórios $\Omega = \sigma_u^2 I_T + \sigma_c^2 J_T J_T'$ onde I_T é uma matriz identidade e $J_T J_T'$ uma matriz unitária de dimensão (T x T).

Dessa forma, o estimador de efeitos aleatórios consiste em estimar as variâncias dos efeitos idiossincráticos e específicos. E o estimador é:

$$\hat{\beta} = \left(\sum_{j=1}^N X_j' \hat{\Omega}^{-1} X_j \right)^{-1} \left(\sum_{j=1}^N X_j' \hat{\Omega}^{-1} y_j \right)$$

Modelo de Efeitos Fixos

A hipótese de exogeneidade estrita do modelo de efeitos aleatórios é bastante forte quando estamos tratando de empresas de grande porte. Dessa forma, justifica-se o modelo considerar os efeitos como sendo individuais.

A estimação do modelo de Efeitos Fixos se dá através da eliminação do efeito específico estimando o modelo da seguinte forma: calcula-se as médias temporais das variáveis de cada firma e subtrai cada observação de sua média temporal. Feito isso o modelo a ser estimado será por essa diferença, ou seja:

$$y_{jt} - \bar{y}_j = \epsilon_{jt} - \bar{x}_j \beta + v_{jt} - \bar{v}_j$$

ou

$$\ddot{y}_{jt} = \ddot{x}_{jt} \beta + \ddot{v}_{jt}$$

Dessa forma, o estimador é:

$$\hat{\beta} = \left(\sum_{j=1}^N \ddot{X}_j' \ddot{X}_j \right)^{-1} \left(\sum_{j=1}^N \ddot{X}_j' \ddot{y}_j \right) = \left(\sum_{j=1}^N \sum_{t=1}^T \ddot{x}_{jt}' \ddot{x}_{jt} \right)^{-1} \left(\sum_{j=1}^N \sum_{t=1}^T \ddot{x}_{jt}' \ddot{y}_{jt} \right)$$

Além das metodologias de Efeitos Fixos e Efeitos Aleatórios há ainda a abordagem de Primeiras Diferenças. Como na metodologia de Efeitos Fixos, esta abordagem também elimina o efeito não-observável c_j . Quando há apenas dois períodos, essas duas abordagens produzem estimativas e inferências idênticas. Quando $T > 2$, que é o caso neste trabalho, o estimador de Efeitos Fixos é mais eficiente sob a hipótese de que v_{jt} são serialmente não-correlacionados, enquanto que o modelo de Primeiras Diferenças é mais eficiente se v_{jt} for um passeio aleatório.

Hausman (1978) propôs um teste baseado na diferença entre as estimativas dos modelos fixos e aleatórios. A estatística do teste é:

$$H = \left(\hat{\epsilon}_{FE} - \hat{\delta}_{RE} \right)' \left[\hat{A} \hat{v} \hat{A}' \left(\hat{\epsilon}_{FE} - \hat{\delta}_{RE} \right) \right]^{-1} \left(\hat{\epsilon}_{FE} - \hat{\delta}_{RE} \right)$$

onde os deltas estimados são os vetores das estimativas de efeitos fixos e aleatórios, FE e RE respectivamente.

Dado que os estimadores de efeitos fixos são consistentes quando c_j e x_{jt} são correlacionados, mas os estimadores de efeitos aleatórios são inconsistentes, uma diferença estatisticamente significativa é interpretada como evidência contra a suposição de Efeitos Aleatórios.

3.2. Modelos Não Lineares de Resposta Binária

3.2.1. Modelo Probit Binário com Dados em Cross-Section

Sejam y_j , $j = 1, 2, \dots, n$, observações aleatórias independentes que assumem, para o j -ésimo elemento da amostra, valor 0 ou 1. Seja x_j o vetor ($p \times 1$) de variáveis independentes para o indivíduo j . O modelo probito é definido por:

$$Pr(y_j = 1 | x_j) = \Phi(\beta x_j')$$

sendo β um vetor de parâmetros ($1 \times p$) e $\Phi(\cdot)$ a função distribuição acumulada de uma variável aleatória normal com média zero e variância unitária.

A função de verossimilhança de β é dada por:

$$L(\beta; y, X) = \prod_{j=1}^n \Phi(\beta x_j')^{y_j} [1 - \Phi(\beta x_j')]^{1-y_j}$$

onde $y = (y_1, y_2, \dots, y_n)$ e $X = (x_1, x_2, \dots, x_n)$.

Conforme Greene (2002) considera-se a variável latente⁴ y^* :

$$y^* = \beta x_j' + \varepsilon_j$$

onde ε_j são independentes com distribuição normal padrão. Assumindo que:

⁴ Variável latente é uma variável que não é diretamente observada.

$$y_j = 1 \Leftrightarrow y^* \geq 0 \text{ e } y_j = 0 \Leftrightarrow y^* < 0$$

temos que:

$$Pr(y_j = 1 | x_j) = Pr(\beta x'_j + \varepsilon_j \geq 0 | x_j) = Pr(\varepsilon_j \leq \beta x'_j | x_j) = \Phi(\beta x'_j)$$

A abordagem da variável latente é útil na estimação da inadimplência, pois a variável resposta é dicotômica. Suponha que y_j representa uma variável qualitativa ordinal com m categorias (cada m corresponde a categoria do *rating*), denominadas R_1, R_2, \dots, R_m . Admitindo, por exemplo, que $y_j \geq R_m$ se inadimplente e $y_j < R_m$ se adimplente, através da inserção da variável latente tem-se o que não é diretamente observado e que expressa a probabilidade de default.

3.2.2. Modelo Logito Binário com Dados em Painel

Modelo Logito Binário de Efeitos Fixos

O modelo de efeitos fixos é:

$$y^*_{jt} = c_j d_{jt} + \beta x'_{jt} + \varepsilon_{jt}, \quad j = 1, \dots, n, t = 1, \dots, T$$

$$y_{jt} = 1 \text{ se } y^*_{jt} > 0, \text{ e } y_{jt} = 0 \text{ caso contrário.}$$

onde d_{jt} é uma variável *dummy* que recebe valor um para o indivíduo j e zero caso contrário. Por questões de conveniência o vetor x_{jt} é redefinido excluindo a constante do modelo.

A função de verossimilhança para modelo de efeitos fixos é:

$$\ln L = \sum_{j=1}^n \sum_{t=1}^T \ln P(y_{jt} | c_j + \beta x'_{jt})$$

onde $P(.)$ é a probabilidade de resposta observada.

No caso do modelo linear a eliminação do efeito específico é simples, pois a estimação dos parâmetros é possível pela transformação dos dados em desvios em relação a média, o *within estimator*. Entretanto, no modelo logito, a não-linearidade da variável dependente inviabiliza a aplicação da solução de desvios com relação à média. Esse problema é chamado de problema do parâmetro incidental⁵. Mas quão sério é esse problema? Heckman e MaCurdy's (1980) fazem um estudo de Monte Carlo com uma amostra de $n=100$ e $T=8$, e o viés encontrado é da ordem de dez por cento.

Para um modelo de escolha binária, Chamberlain (1980) sugere uma abordagem baseada na função de verossimilhança condicional para estimar β . Condicionando a distribuição de $y_{j,t}$ a c_j e a $\sum_{t=1}^T y_{j,t}$, o efeito específico c_j é eliminado do estimador, gerando uma probabilidade que não depende dos parâmetros incidentais c_j fazendo com que o estimador de máxima verossimilhança de β_t seja consistente e assintoticamente normal.

3.3. Modelos Não Lineares de Resposta Ordenada

O modelo de resposta ordenada é um tipo de modelo de resposta multinomial. A variável resposta desse estudo é um *credit rating*, ou seja, um “ranking” que reflete a qualidade creditícia da empresa. Dessa forma, um modelo para estimar *credit rating* se enquadra perfeitamente em um modelo de resposta ordenada. É importante ressaltar que variável resposta com essa característica possui apenas uma medida de ordenação, nada se podendo afirmar sobre a magnitude entre as diferenças de uma escala e outra. Por exemplo, é equivocado afirmar que o “*rating 4*” é duas vezes mais (ou menos) arriscado que o “*rating 2*”.

O modelo de resposta ordenada tem o objetivo de estimar qual a probabilidade do indivíduo pertencer a uma determinada classe. No caso desse

⁵ Ver Neyman e Scott (1948).

estudo, o modelo irá fornecer qual a probabilidade da empresa pertencer, por exemplo, ao *rating Baa2*.

3.3.1. Modelo Probit Ordinal com Dados em Cross-Section

A estimação dos *ratings* usando um modelo probito ordenado é, em geral, considerada uma boa estratégia. Os *ratings* são medidas qualitativas ordinais, e esse método irá determinar o tamanho da diferença entre cada categoria. Esse procedimento foi utilizado por Hu et al. (2002), Bissoondoyal-Bheenick (2005) e Bissoondoyal-Bheenick et al.(2005) na estimação de matrizes de transição para *credit ratings* soberano.

A partir da generalização do modelo $y^* = \beta x'_j + \varepsilon_j$, assume-se que y_j representa uma variável qualitativa ordinal com m categorias (cada m corresponde a categoria do *rating*), denominadas R_1, R_2, \dots, R_m , derivados da variável não observável y . Para cada *rating* existem $m+1$ pontos de cortes, $\mu_0, \mu_1, \mu_2, \dots, \mu_m$, com $\mu_0 = -\infty$, $\mu_m = +\infty$ e $\mu_0 \leq \mu_1 \leq \mu_2 \leq \dots \leq \mu_m$ em que $\mu_{k-1} \leq y_j \leq \mu_k \Leftrightarrow y^*_j \in R_k$ para $1 \leq j \leq N$. X_j é o vetor de variável independente (x_1, \dots, x_k) da empresa j ($x_0 = 1$). Dessa forma:

$$\mu_{k-1} \leq Y_j \leq \mu_k \Leftrightarrow \mu_{k-1} \leq \beta X_j + \varepsilon_j \leq \mu_k \Leftrightarrow \frac{\mu_{k-1} - \beta X_j}{\sigma} < \frac{\varepsilon_j}{\sigma} \leq \frac{\mu_k - \beta X_j}{\sigma}$$

e

$$\Pr(\mu_{k-1} \leq Y_j \leq \mu_k) = \Phi\left(\frac{\mu_k - \beta X_j}{\sigma}\right) - \Phi\left(\frac{\mu_{k-1} - \beta X_j}{\sigma}\right)$$

onde $\Phi(\cdot)$ é a função de distribuição acumulada para uma variável aleatória padronizada. Sem perda de generalidade, assume-se que $\mu_1 = 0$ e $\sigma = 1$. Dessa forma, o modelo a ser estimado é:

$$\Pr(\mu_{k-1} \leq Y_j \leq \mu_k) = \Phi(\mu_k - \beta X_j) - \Phi(\mu_{k-1} - \beta X_j)$$

Será necessário estimar $m+k+1$ parâmetros: $\mu_1, \mu_2, \dots, \mu_{m-1}$ e $\beta_0, \beta_1, \dots, \beta_k$.

$$P(y=0) = Pr(y_j \leq \mu_k) = \Phi(\mu_k - \beta x'_j)$$

$$P(y=1) = Pr(y_j > \mu_k) = 1 - \Phi(\mu_k - \beta x'_j)$$

$$P(y=k) = Pr(\mu_{k-1} < y_j \leq \mu_k) = \Phi(\mu_k - \beta x'_j) - \Phi(\mu_{k-1} - \beta x'_j)$$

onde $\Phi(\cdot)$ é uma distribuição normal⁶. A soma das probabilidades acima é unitária. Logo, os parâmetros μ e β podem ser estimados maximizando a verossimilhança.

3.3.2. Modelo Probita Ordinal com Dados em Painel

Como visto no início da seção anterior o modelo probito ordenado é usado para modelos de escala discreta que representam indicadores de preferência, desempenho, ou nível de risco. Entretanto, devido a natureza dos dados é necessária a consideração da heterogeneidade individual presente.

Modelo Probita Ordinal com Efeitos Fixos

Modelos não lineares incorrem no problema do parâmetro incidental. A abordagem utilizada no modelo binário para contornar esse problema não pode ser diretamente estendida para modelos de painel de resposta ordenada cuja variável dependente tem $p > 2$ possibilidades. Entretanto, é possível fazer combinações das categorias adjacentes tal qual a variável dependente é resumida como uma variável binária, e então usar esse método. Repetindo-se esse método para todas as possíveis combinações das categorias adjacentes, tem-se $p-1$ estimativas dos parâmetros de interesse.

⁶ Outras distribuições, particularmente a logística, poderiam ser utilizadas. As distribuições logística e normal geralmente dão resultados muito similares na prática, Greene (2008) p.832.

4. DESCRIÇÃO DOS DADOS

4.1. Agências de Rating

Atualmente, as duas principais agências de classificação de crédito são Standard & Poor's (S&P) e Moody's. A primeira operando em mais de cinquenta países ao redor do mundo e esta operando principalmente nos Estados Unidos, mas tendo diversas filiais internacionais. A Figura 1 oferece um resumo das definições das classificações de S&P e Moody's para crédito de longo prazo. As categorias são definidas em termos de risco de inadimplência e da probabilidade de pagamento pelo emitente. Emissões classificadas nas quatro categorias mais altas (AAA, AA, A e BBB de S&P e Aaa, Aa, A e Baa da Moody's) são geralmente consideradas como sendo de classe de investimento. Obrigações classificadas como BB, B, CCC, CC e C (Ba, B, Caa e Ca da Moody's), são consideradas como tendo características especulativas significantes.

Grau de Investimento			Grau Especulativo		
S&P	Moody's	Interpretação	S&P	Moody's	Interpretação
AAA	Aaa	Altíssima qualidade de crédito, com risco de crédito mínimo. Capacidade de honrar compromissos financeiros é extremamente forte.	BB+	Ba1	Elementos especulativos e subjetivos substanciais no risco de crédito. Sobre maiores incertezas ou exposições a condições econômicas adversas podem levar a inadequada capacidade de pagamento.
			BB	Ba2	
			BB-	Ba3	
AA+	Aa1	Alta qualidade de crédito, com baixo risco de crédito. Capacidade de honrar compromissos financeiros é muito forte.	B+	B1	Elementos especulativos e subjetivos fortes no risco de crédito. Sobre maiores incertezas ou exposições a condições econômicas adversas provavelmente levam a inadequada capacidade de pagamento.
AA	Aa2		B	B2	
AA-	Aa3		B-	B3	
A+	A1	Forte capacidade de pagamento. Mas suscetível a efeitos adversos de mudanças em circunstâncias e condições econômicas.	CCC+	Caa1	Alto risco de crédito. Vulnerabilidade corrente para estar ou entrar em default. Dependente de condições financeiras e econômicas favoráveis. Sob condições adversas não possui capacidade de pagamento.
A	A2		CCC	Caa2	
A-	A3		CCC-	Caa3	
BBB+	Baa1	Capacidade de pagamento adequada. Mas efeitos adversos de mudanças em circunstâncias e condições econômicas enfraquecem a capacidade de pagamento.	C	Ca	Tipicamente em default com baixa propensão de recuperação. Requerimento de falência é comum.
BBB	Baa2				
BBB-	Baa3				
			D		Falência ou Default

Figura 1 – Resumo das definições de Rating.

Fonte: Compilação de Minardi, Sanvicente e Artes baseado na S&P e Moody's; tradução do autor.

Modelos de *credit rating* são modelos aplicados à classificação de empresas em categorias de risco de crédito, normalmente utilizando-se de critérios quantitativos, qualitativos e julgamentais. As metodologias de análise das agências Moody's e S&P dão um peso maior na análise julgamental (Crouhy *et al.*). A Figura 2 abaixo ilustra os fatores que a agência Moody's considera no *rating* de uma empresa industrial, esses fatores abrangem desde o risco país, até aspectos macroeconômicos, e a estrutura específica da emissão dos títulos.



Figura 2 – Análise de Classificação da Moody's de uma Empresa Industrial.
Fonte: Gerenciamento de Risco (pag. 237).

Mas “Quão precisas são as classificações?”, pergunta Moody's em seu *Credit Rating and Research* (1995, p.5). A resposta está no Gráfico 1, que mostra as taxas médias cumulativas de inadimplência para emitentes de títulos de dívida corporativos para cada nível de *rating*.

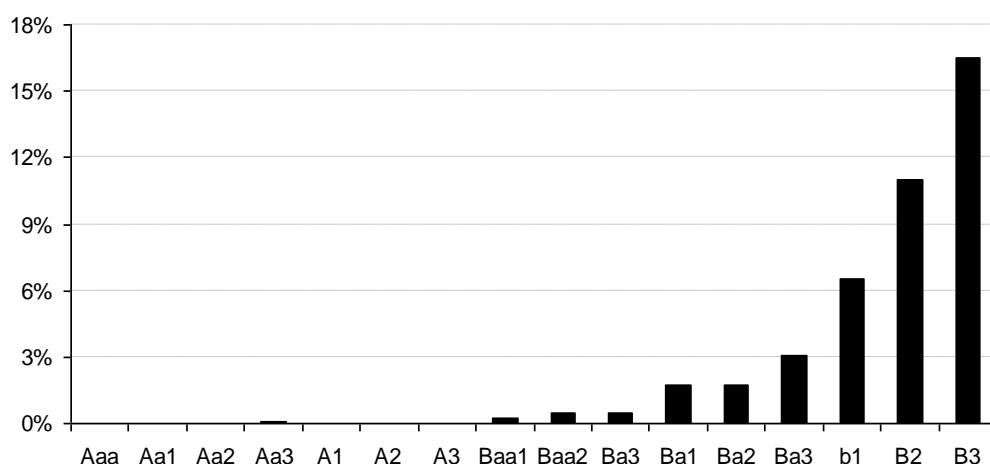


Gráfico 1 – Taxas de Inadimplência em Um Ano por Classificação, 1983-1993.
 Fonte: Credit Rating and Research, Moody's, 1995.

4.2. Dados

Assim como Minardi, Sanvicente e Artes (2007), o público a ser modelado será empresas norte americanas do setor industrial. Os autores testaram diversos índices referentes a: tamanho da empresa, alavancagem financeira, índices de solvência, resultados operacionais e volatilidades das ações. O método de seleção de variáveis utilizado escolheu as variáveis abaixo e para comparabilidade dos resultados essas variáveis também serão utilizadas nesse trabalho:

LN (asset): variável referente ao tamanho da empresa. Em geral, quanto maior o tamanho do ativo da empresa, melhor o *rating* da empresa. A variável é o *logaritmo neperiano do ativo total da empresa*. A transformação logarítmica é utilizada para diminuir a assimetria da variável e reduzir o impacto de possíveis *outliers*. Outra variável relacionada ao tamanho da firma e também muito utilizada na análise de crédito é o *valor contábil do patrimônio líquido (book value of equity)*.

Debt/Asset: variável que mostra o nível de alavancagem financeira. A variável construída é *dívida total sobre o ativo total da empresa*. Além dessa, variáveis comuns na análise de crédito são os índices criados como combinações das contas: *passivo, ativo total, patrimônio líquido e dívida*. Em geral, quanto maior a alavancagem pior o *rating* da empresa.

Ebit/Net Debt: variável referente ao nível de solvência. Uma variante da conta *Ebit* é o *Ebitda*⁷. Além dessas, também se utiliza muito as variáveis *ativo corrente menos passivo corrente dividido pelo ativo total* e *Ebit dividido pelo ativo total*. Claramente, quanto maior o *Ebit* da empresa em relação a sua dívida líquida melhor sua capacidade financeira de pagamento.

Roa: variável de performance operacional. O retorno sobre o ativo é definido como *lucro líquido dividido pelo ativo total*.

Ebit/Net Income: variável também de resultado operacional. *Lucro antes dos efeitos financeiros dividido pelo faturamento líquido*.

Vol: a única variável de mercado do modelo, que reflete a *volatilidade do retorno das ações*. A volatilidade é definida como o desvio padrão dos últimos doze retornos mensais.

Os demonstrativos contábeis das empresas e a volatilidade de suas respectivas ações foram capturados na *Economática* e os dados de *ratings* da Moody's no terminal da *Bloomberg*. O resultado dessa pesquisa gerou uma extensiva base de dados em forma de painel desbalanceado de trezentas e vinte e seis empresas com *ratings* da agência Moody's de dados semestrais de dez/2004 a jun/2009 totalizando dois mil setecentos e trinta e seis registros.

O Gráfico 2 mostra a distribuição das empresas por *rating* no primeiro e último períodos analisados. O formato da distribuição é similar ao formato da maioria dos porta-fólios das instituições financeiras, onde tem-se uma pequena quantidade de empresas extremamente muito boas, e uma pequena quantidade de empresas com a capacidade creditícia ruim. O grande volume da carteira é de empresas ao redor do risco mediano.

⁷ Ebitda = Earnings before interest, tax, depreciation and amortization (Lucro antes dos efeitos financeiros, impostos sobre a renda, depreciação e amortização)

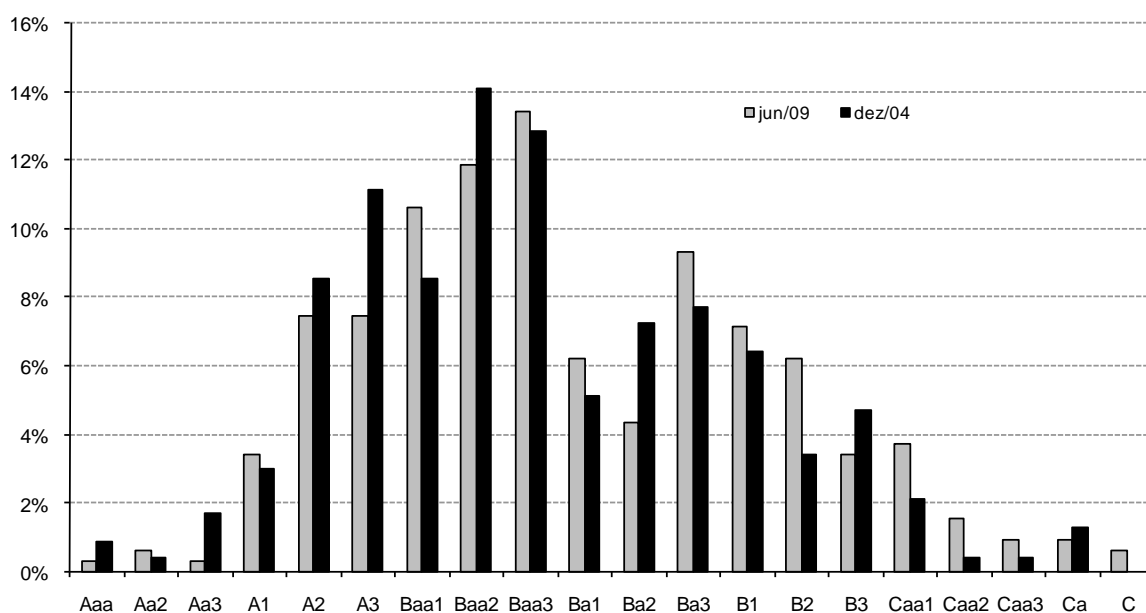


Gráfico 2 – Distribuição das empresas por rating.

Uma análise descritiva dos dados mostra uma média de aproximadamente duzentas e oitenta empresas por período analisado conforme se observa no Gráfico 3. Além disso, setenta e um por cento das empresas possuem dados nos dez períodos que serão utilizados na análise, conforme Gráfico 4. Empresas que possuem informação de demonstrativos contábeis nos dez períodos implicam em empresas que divulgam seus balanços, além do fechamento anual, também no final de cada semestre. A transparência nas informações contábeis é esperada nas empresas de capital aberto. Além disso, onze por cento da base de dados são de empresas que possuem dados em apenas cinco períodos, essas são empresas que divulgam apenas balanços anuais.

Uma segunda análise inicial e importante a ser feita quando o método de estimação é com dados em painel é avaliar a variabilidade dos dados ao longo do período. O resultado obtido está apresentado no Gráfico 5.

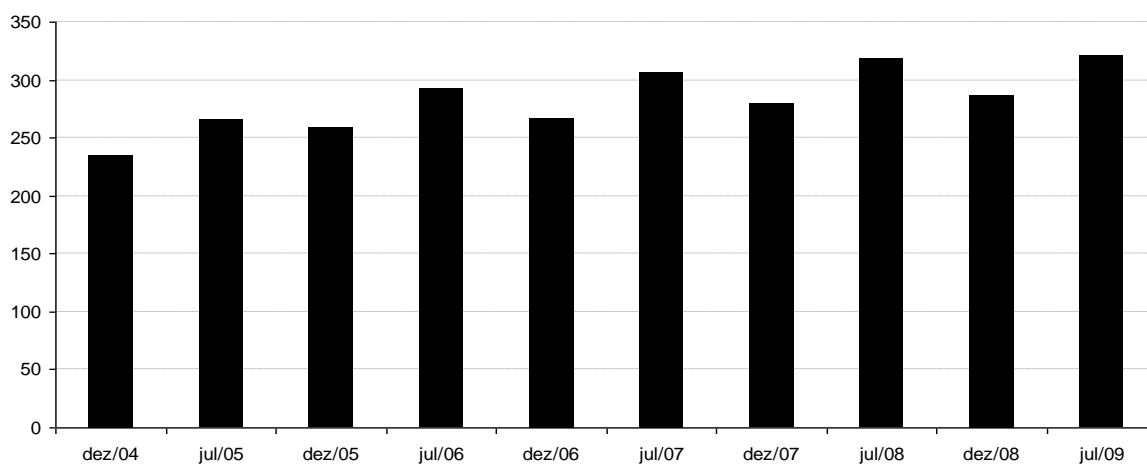


Gráfico 3 – Quantidade de Empresas por Período

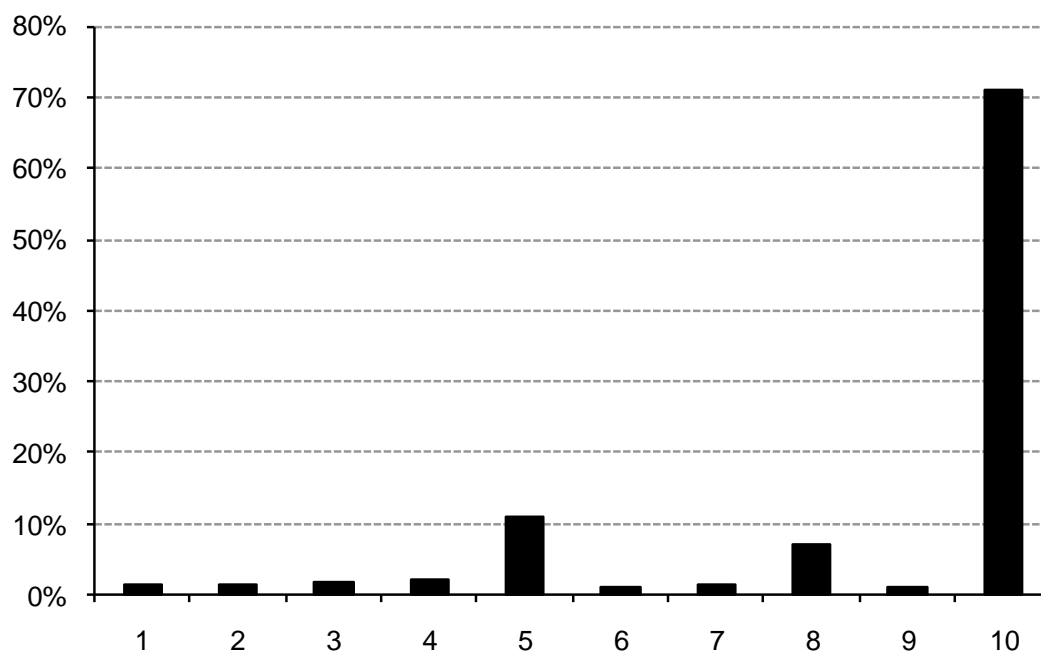


Gráfico 4 – Distribuição da Quantidade de Empresas que possuem dados no maior número de períodos possíveis.

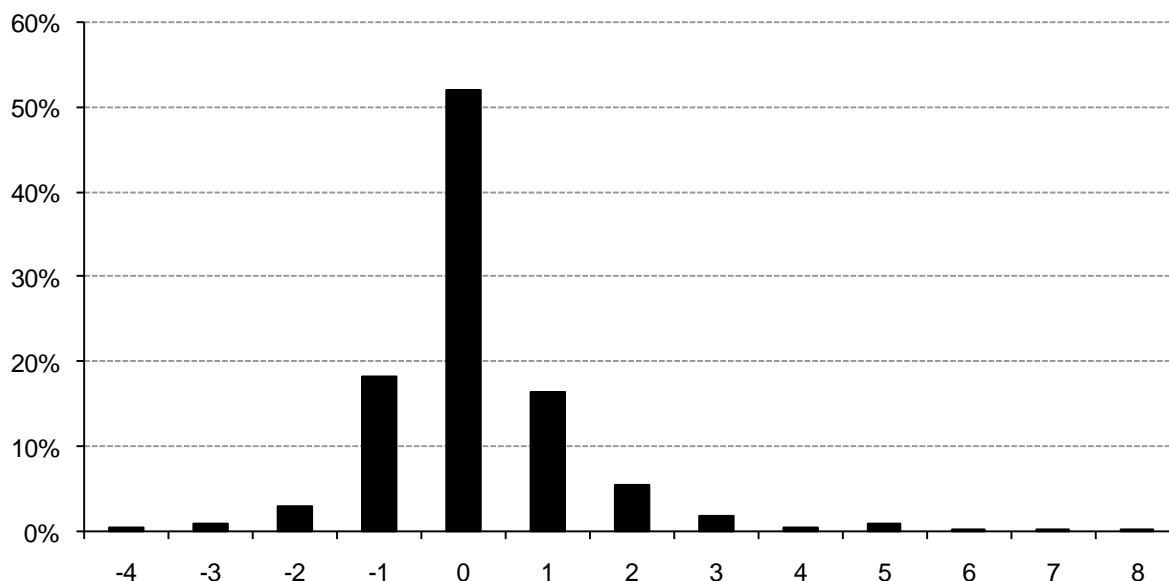


Gráfico 5 – Distribuição das Mudanças de *Ratings* da Moody's.

A migração de *ratings* é moderada, cinquenta e dois por cento das vezes não houve nenhuma alteração de *rating* para nenhuma das trezentas e vinte e seis empresas ao longo dos dez semestres. Aproximadamente dezoito por cento de *upgrades* e dezoito por cento de *downgrades* de apenas um nível em pelo menos um dos semestres. Ao longo do período analisado observou-se uma maior taxa de *downgrades* do que *upgrades*, resultado esperado, pois a distribuição de *ratings* possui uma assimetria à esquerda.

4.3. Correlação: variáveis independentes versus rating

Uma análise descritiva das variáveis independentes com a variável dependente é apresentada na Tabela 2 e na Figura 3, trazendo evidências que o modelo obterá resultados semelhantes aos encontrados por Minardi, Sanvicente e Artes (2007). Para essa análise agrupam-se os *ratings* da Moody's na letra, ou seja, as classificações Aa1, Aa2 e Aa3 foram consolidadas em Aa, e as classificações A1, A2 e A3 em A, e assim por diante.

No estudo de Minardi, Sanvicente e Artes (2007) as variáveis relacionadas ao *Ebit* obtiveram coeficientes estatisticamente iguais a zero. No resultado desse

estudo uma dessas variáveis, *Ebit/Debt*, possui correlação bastante baixa com a variável dependente dando evidência que seu coeficiente estimado poderá ter baixa significância estatística.

Tabela 2 - Análise descritiva das variáveis independentes por classes de rating

Rating	LN Asset	Debt / Asset	EBIT / Debt	ROA	EBIT / Net Income	Vol
	média desv. pad.	média desv. pad.	média desv. pad.	média desv. pad.	média desv. pad.	média desv. pad.
Aaa	19.8 (0.62)	0.05 (0.03)	-166 (191)	9.5 (7.29)	0.2 (0.06)	24.2 (13.69)
Aa	17.0 (0.76)	1.21 (1.36)	-55 (511)	9.6 (5.92)	0.2 (0.05)	19.2 (11.65)
A	16.6 (1.45)	2.84 (3.13)	55 (832)	7.5 (5.39)	0.1 (0.10)	29.3 (15.51)
Baa	16.0 (1.14)	4.57 (4.55)	-46 (4472)	5.6 (6.89)	0.1 (0.13)	33.0 (18.03)
Ba	15.6 (1.21)	7.20 (8.66)	81 (607)	2.9 (9.45)	0.1 (0.26)	44.0 (24.79)
B	15.1 (0.86)	12.56 (13.45)	-37 (1186)	-0.4 (13.44)	0.1 (0.39)	51.4 (29.74)
Caa	15.3 (1.29)	24.10 (41.27)	-19 (203)	-14.1 (24.09)	0.0 (0.55)	75.3 (39.76)
Ca	15.9 (1.42)	11.76 (9.76)	33 (94)	-10.6 (13.80)	0.0 (0.14)	72.8 (43.87)
C	14.9 (0.62)	17.67 (23.68)	10 (117)	-29.7 (22.34)	-0.2 (0.31)	228.6 (-)
Correlação	-0.4035	0.3712	-0.0410	-0.3962	-0.1918	0.4269

Os valores fora do parêntesis são as médias das variáveis por *rating* e os valores entre parêntesis seus respectivos desvios-padrão. A correlação foi estimada através do coeficiente de correlação de spearman.

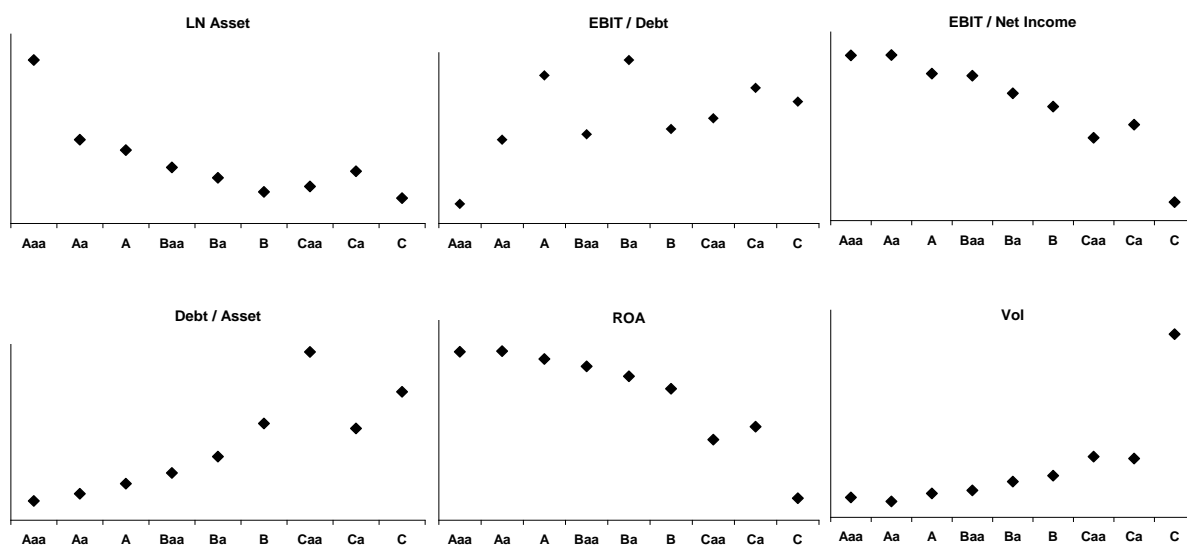


Figura 3 – Análise gráfica de correlação das variáveis independentes vs classes de *rating*.

5. RESULTADOS

Para estimar os modelos de avaliação de risco de empresas serão testadas algumas das metodologias econométricas apresentadas no capítulo 3. As duas primeiras são metodologias onde os dados em painel serão agrupados, formando uma grande *cross-section*. A primeira sendo estimada através de modelos lineares utilizando o estimador de Mínimos Quadrados Ordinários onde a variável resposta é considerada como sendo cardinal, colunas 1 e 2 da Tabela 3. A segunda metodologia será através de um modelo não-linear utilizando a função de ligação probito e a variável resposta tratada como ordenada e o estimador o de máxima verossimilhança, coluna 3. Os resultados desses modelos estão na Tabela 3.

Tabela 3 - Modelos em <i>Cross-Section</i>			
Variável	MQO	MQO	PROBIT
Ln (asset)	-1,017*** (0,043)	-0,840*** (0,041)	-0,340*** (0,017)
Debt/Asset	-0,035 (0,019)	-0,024 (0,018)	-0,006 (0,007)
Ebit/Net Debt	-7,66 (7,36)	-7,57 (6,94)	-4,01 (2,78)
Roa	-0,095*** (0,006)	-0,083*** (0,006)	-0,035*** (0,002)
Ebit/Net Income	1,262*** (0,251)	1,103*** (0,237)	0,480*** (0,095)
Vol	0,035*** (0,002)	0,077*** (0,003)	0,030*** (0,001)
Constante	24,417*** (0,708)	17,131*** (0,784)	-
N	2736	2736	2736
Dummies de Período	Não	Sim	Sim
Log verossimilhança	-	-	-6040,93

Nota: os valores entre parêntesis fornecem o desvio-padrão das estimativas. As significâncias estatísticas das estimativas estão representadas pelos asteriscos com a seguinte legenda: p-valor < 0.001 (***), 0.001 < p-valor < 0.01 (**) e 0.01 < p-valor < 0.05 (*). A ausência de asterisco significa que a estimativa não foi significativa a pelo menos cinco por cento. Para efeito de apresentação os coeficientes estimados da variável Ebit/Net Debt foram multiplicados por 10⁶.

A coluna 1 é o modelo linear sem *dummies* de período, a coluna 2 é o modelo linear inserindo variáveis indicadoras controlando a regressão. E na coluna 3 o modelo probito. Os resultados nos três modelos foram semelhantes. As variáveis *Ln(asset)*, *Roa*, *Ebit/Net Income* e *Vol* foram significativas e tiveram seus sinais de acordo com o esperado. A variável *Ebit/Debt* não foi estatisticamente significativa. Além disso, a variável *Debt/Asset* foi significativa a sete por cento no primeiro modelo e não significativa a dez por cento nos outros dois modelos, porém o sinal do coeficiente da variável não foi conforme o esperado. No modelo probito ordenado é estimada uma constante para cada *rating* que foram omitidas da tabela⁸. A inserção de *dummies* de período na coluna 2 não afetou qualitativamente os resultados das estimações.

Os resultados obtidos com o modelo probito ordenado são semelhantes aos obtidos por Minardi, Sanvicente e Artes (2007). Porém, nesse estudo a variável *Debt/Asset* deixou de ser significativa.

Em seguida, será apresentada a metodologia onde o tratamento para dados em painel é mais adequado. A Tabela 4 apresenta os modelos lineares de efeitos fixos e os modelos lineares de efeitos aleatórios.

⁸ A saída completa da regressão encontra-se no apêndice onde as constantes podem ser analisadas.

Tabela 4 - Modelos em Painel.		
Variável	Efeitos Fixos	Efeitos Aleatórios
Ln (asset)	-0,304*** (0,077)	-0,578*** (0,062)
Debt/Asset	0,011 (0,008)	0,012 (0,009)
Ebit/Net Debt	1,66 (2,6)	0,935 (2,72)
Roa	-0,027*** (0,003)	-0,027*** (0,003)
Ebit/Net Income	0,301** (0,089)	0,303** (0,092)
Vol	0,020*** (0,0017)	0,023*** (0,002)
Constante	12,801*** (1,244)	17,019*** (1,020)
N	2736	2736
Dummies de Período	Sim	Sim
Teste F	34,86*** F(15,24)	-
Teste Wald		685,00*** Chi2(15)
Teste de Hausman: $\text{Chi}^2(14) = 229,80^{***}$		

Nota: os valores entre parêntesis fornecem o desvio-padrão das estimativas. As significâncias estatísticas das estimativas estão representadas pelos asteriscos com a seguinte legenda: p-valor < 0.001 (***), 0.001 < p-valor < 0.01 (**) e 0.01 < p-valor < 0.05 (*). A ausência de asterisco significa que a estimativa não foi significativa a pelo menos cinco por cento. Para efeito de apresentação os coeficientes estimados da variável Ebit/Net Debt foram multiplicados por 10^6 .

Os resultados obtidos com os modelos de efeitos fixos e efeitos aleatórios foram muito semelhantes. Apesar das diferenças entre os modelos de efeitos fixos e efeitos aleatórios serem pequenas, o teste de Hausman trouxe evidência contra a suposição dos efeitos serem aleatórios, corroborando a hipótese feita anteriormente.

O último modelo testado é o modelo não-linear de efeitos fixos utilizando a função de ligação logito. Como visto no capítulo três, o método para estimar o

modelo logito ordinal para p categorias consiste em fazer combinações das categorias adjacentes tal que a variável dependente seja resumida como uma variável binária. Para usar essa metodologia agrupou-se os ratings em 8 categorias (Aaa, Aa, A, Baa, Ba, B1, Caa, Ca). Dessa forma tem-se 7 regressões com estimativas distintas para cada um dos parâmetros. A tabela 5 representa as recodificações da variável resposta utilizadas nas regressões.

Tabela 5: Recodificação das Variáveis Respostas			
Regressão	y = 0	y = 1	N (y=0 ; y=1)
(1)	Aaa	caso contrário	(18 ; 2718)
(2)	Aaa, Aa	caso contrário	(53 ; 2683)
(3)	Aaa, Aa, A	caso contrário	(606 ; 2130)
(4)	Aaa, Aa, A, Baa	caso contrário	(1629 ; 1107)
(5)	Aaa, Aa, A, Baa, Ba	caso contrário	(2164 ; 572)
(6)	Aaa, Aa, A, Baa, Ba, B1	caso contrário	(2633 ; 103)
(7)	Aaa, Aa, A, Baa, Ba, B1, Caa	caso contrário	(2725 ; 11)

Nota: A última coluna contém o número de registros para cada uma das sete regressões sempre totalizando 2.736 registros.

A Tabela 6 fornece o resultado do modelo logito de efeitos fixos para cada uma das sete regressões.

Tabela 6: Modelos Não-Lineares de Efeitos Fixos.							
Variável	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln (asset)	-0,009*** (0,01)	- 0,006*** (0,0003)	-0,644 (0,904)	-7,66*** (1,321)	-0,994* (0,4003)	-0,959 0,640)	-10,57 (5,88)
Debt/Asset	-0,002 (0,0017)	- 0,011*** (0,001)	0,082 (0,123)	-0,040 (0,099)	-0,025 (0,055)	0,095 (0,078)	0,266 (0,334)
Ebit/Net Debt	2,093 (1,334)	-20,0*** (4,4)	-60,0 (370)	377,0 (271,0)	-40,0 (257,0)	-0,473 (50,0)	70 (44200)
Roa	-0,010*** (0,0008)	- 0,003*** (0,0005)	-0,190* (0,079)	- 0,138*** (0,034)	-0,079*** (0,020)	-0,029 (0,018)	-0,315 (0,188)
Ebit/Net Income	-0,105** (0,034)	-0,032 (0,021)	6,268 (4,701)	-2,161 (2,261)	0,669 (0,589)	0,184 (0,478)	-10,027 (13,07)
Vol	0,004*** (0,0002)	0,005*** (0,0001)	0,039*** (0,014)	0,0069 (0,0093)	-0,0001 (0,0058)	0,019*** (0,0067)	-0,007 (0,018)
Constante	5,04*** (0,01)	3,94*** (0,001)	30,33 (746,8)	151,20 (1055,4)	15,97* (6,97)	11,10 (11,72)	173,50 (788,80)
Dummies de Período	Não	Não	Não	Não	Não	Não	Não
Log verossimilhança	199,21	498,20	234,17	285,73	450,44	282,10	33,11

Nota: Os valores entre parêntesis fornecem o desvio-padrão das estimativas. As significâncias estatísticas das estimativas estão representadas pelos asteriscos com a seguinte legenda: p-valor < 0,001 (***), 0,001 < p-valor < 0,01 (**) e 0,01 < p-valor < 0,05 (*). A ausência de asterisco significa que a estimativa não foi significativa a pelo menos cinco por cento. Para efeito de apresentação os coeficientes estimados da variável Ebit/Net Debt foram multiplicados por 10⁶.

Tabela 7: Estatísticas dos coeficientes dos modelos não-lineares de efeitos fixos.							
Variável	Ln (asset)	Debt /Asset	Ebit/Net Debt	Roa	Ebit/Net Income	Vol	Constante
Média	-2,98	0,05	46,95	-0,11	-2,53	0,01	55,87
Desvio Padrão	4,29	0,11	151,22	0,11	4,08	0,02	73,55
Coeficiente de Variação	-1,44	2,04	3,22	-1,05	-1,61	1,59	1,32

As estimativas para cada um dos modelos foram bastante distintas, com coeficientes variando de negativo para positivo para algumas variáveis. As duas primeiras regressões, que consideram os melhores *ratings*, forneceram estimativas significantes para a maior parte das variáveis. A quarta regressão, que divide as empresas em grau de investimento e grau especulativo, obteve resultados divergentes dos demais modelos e com poucas estimativas significantes. E a última regressão obteve coeficientes não significantes para todas as variáveis.

A variável *Roa* foi a que obteve estimativas mais estáveis com um coeficiente de variação de apenas 1,05. Em seguida *Ln(asset)*, *Vol* e *Ebit/Net Income* com uma variação um pouco mais elevada. E finalmente, *Debt/Asset* e *Ebit/Net Debt* foram as variáveis que obtiveram estimativas com as maiores variações. A tabela 7 mostra as estatísticas das estimativas para cada um dos modelos binários.

6. CONCLUSÃO

Nesse trabalho foram estudadas diferentes abordagens quantitativas na avaliação do risco de crédito de empresas, de modelos estatísticos a modelos financeiros. Em um contexto de poucos defaults e sob a ótica de Basiléia II, que tem como requisito para as instituições financeiras o desenvolvimento de modelos internos e que forneçam uma probabilidade de inadimplência, a abordagem de replicar os *ratings* de agências externas de classificação é a mais adequada.

Para estimação dos *credit ratings* utilizou-se as abordagens considerando os dados como uma *cross-section* e também em painel. Nos dados agrupados em *cross-section* os três modelos testados apresentaram resultados semelhantes. A inserção de variáveis indicadoras de semestre controlando a regressão gerou mudanças razoáveis nos parâmetros estimados indicando que provavelmente havia um viés nos estimadores quando comparado ao primeiro modelo dando evidência da necessidade de utilizar metodologias adequadas para dados em painel.

Apesar da hipótese de exogeneidade estrita ser muito forte para empresas de grande porte, os resultados dos modelos de efeitos fixos e efeitos aleatórios, considerando a variável resposta como cardinal, foram muito semelhantes. Em contrapartida, o teste de Hausman trouxe evidência contra a suposição dos efeitos serem aleatórios, corroborando a hipótese de que há correlação entre o efeito individual da firma e as covariáveis do modelo.

As conclusões em todos os modelos são semelhantes ao obtido por Minardi, Sanvicente e Artes quando analisada as significâncias estatísticas e o sinal obtido nos coeficientes. Nesse estudo, os principais fatores determinantes na avaliação do risco das empresas foram tamanho da empresa, retorno sobre o ativo, volatilidade das ações e *Ebit/faturamento líquido*. Em todos os modelos testados essas variáveis foram significantes com exceção da última metodologia no qual a variável *Ebit/faturamento líquido* mostrou-se pouco robusta. A variável retorno sobre o ativo foi a que obteve estimativas mais robustas em todas as metodologias aplicadas. E finalmente, no modelo logito ordenado de efeitos fixos, os fatores considerados na estimação do *credit rating* foram semelhantes nos modelos binários para estimar as empresas com melhores *ratings*.

REFERÊNCIAS

AFONSO, António; GOMES, Pedro; ROTHER, Philipp. What “hides” behind sovereign debt ratings? **European Central Bank**, Working Paper Series no. 711, january 2007.

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

BASEL COMMITTEE ON BANKING SUPERVISION (BCBS). **International convergence of capital measurement and capital standards: a revised framework, 2006**. Disponível em: <<http://www.bis.org/publ/bcbs128.pdf>>. Acesso em: 12 out. 2009.

BERNSTEIN, Peter L. **Desafio aos deuses: a fascinante história do risco** / Peter L. Bernstein; tradução Ivo Korytowski. – Rio de Janeiro: Campus, 1997, 389 p.

BISSOONDOYAL-BHEENICK, Emawtee. An analysis of the determinants of sovereign ratings. **Global Finance Journal**, v. 15 (3), p. 251-280, 2005.

BISSOONDOYAL-BHEENICK, Emawtee; BROOKS, Robert; YIP, Angela. Determinants of sovereign ratings: A comparison of case-based reasoning and ordered probit approaches. **Monash Econometrics and Business Statistics Working Papers** 9/05, Monash University, Department of Econometrics and Business Statistics, may, 2005.

BLOOMBERG. <<http://www.bloomberg.com>>.

BUTLER, J.S.; MOFFITT, Robert. A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model, **Econometrica**, v. 50, n. 3, p. 761-764, 1982.

CHAMBERLAIN, Gary. Analysis of covariance with qualitative data. **Review of Economic Studies**, v. 47, p. 225-238, 1980.

CHAMBERLAIN, Gary. Panel Data. In: Griliches, Z., Intiligator, M.D. (Eds.), **Handbook of Econometrics**, v. II, North-Holland, Amsterdam, p. 1247-1318, 1984.

CROUHY, Michel. **Gerenciamento de risco: abordagem conceitual e prática: uma visão integrada dos riscos de crédito operacional e de mercado** / Michel Crouhy, Dan Galai, Robert Mark; [tradução Carlos Henrique Trieschmann e Luiz Frazão Filho ; supervisão técnica João Carlos Douat]. – Rio de Janeiro: Qualitymark : São Paulo : SERASA, 2004, 635 p.

DAS, M.; VAN SOEST, Arthur. A Panel Data Model for Subjective Information in Household Income Growth, **Journal of Economic Behavior and Organization**, v. 40, p. 409-426, 1999.

De ANDRADE, Fabio W. M. **Desenvolvimento de Modelo de Risco de Portfólio para Carteiras de Crédito a Pessoas Físicas**. São Paulo, 2004. 196 f. Dissertação

(Doutorado em Administração de Empresas) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas - FGV/EAESP, São Paulo, 2004.

De SERVIGNY, Arnaud; RENAULT, Oliver. **Measuring and managing credit risk**. The McGraw-Hill, New York, 2004, 466 p.

FERRER-I-CARBONELL, Ada; FRIJTERS, Paul. How Important is Methodology for the Estimates of the Determinants of Happiness? **Working paper, University of Amsterdam**, 2004.

FISHER, Ronald A. The use of multiple measurements in taxonomic problem. **Annals of Eugenics**, v. 7, p. 179-188, 1936.

GOURIEROUX, C., A. MONFORT. **Simulation-Based Methods Econometric Methods**. Oxford: Oxford University Press, 1996.

GREENE, William H. **Econometric Analysis**, 5th ed., Prentice Hall, New York. 2002.

GREENE, William H. **Econometric Analysis**, 6th ed., Prentice Hall, New York. 2008, 1178 p.

GUPTON, Greg M. et al. **CreditMetrics: technical report**. New York: J.P. Morgan & Co. Incorporated, 1997.

HECKMAN, James; MACURDY, Thomas. A life Cycle Model of Female Labor Supply, **Review of Economic Studies**, v. 47, p. 247-283, 1980.

HORRIGAN, John O. The determination of long term credit standards with financial ratios. Empirical Research in Accounting 1966, **Journal of Accounting Research**, v. 4, supplement, p. 44-62, 1966.

HU, Yen-Ting; KIESEL, R. e PERRAUDIN, William. The Estimation of transition matrices for sovereign credit ratings. **Journal of Banking & Finance**, v. 26 (7), p. 1383-1406, 2002.

KAPLAN, Robert S.; URWITZ, G. Statistical Models of Bond Ratings: A Methodological Inquiry, **The Journal of Business**, v. 52, n.2, p. 231-261, 1979.

LEWIS, Edward M. **An Introduction to Credit Scoring**. San Rafael: Fair, Isaac and Co., Inc. 1992.

MERTON, R. C.; On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. **Journal of Finance**, v. 29, p. 449-470, 1974.

MINARDI, Andrea Maria A. F.; SANVICENTE, Antonio Z.; ARTES, Rinaldo. **Methodology for Estimating Credit Ratings and the Cost of Debt for Business units and Privately-held Companies**. In: Credit Scoring and Credit Control X, 2007, Edinburgh. Papers Accepted for Presentation, 2007

NEYMAN, Jerzy; SCOTT, Elizabeth. Consistent Estimates Based on Partially Consistent Observations. **Econometrica**, v. 16, p. 1-32, 1948.

PINCHES, G. and MINGO, K. A multivariate analysis of industrial bond ratings. **Journal of Finance**, v. 28, p. 201-206, 1973.

PLUTO, Katja; TASCHE, Dirk. Estimating Probabilities of Default for Low Default Portfolios, July 28, 2005. **The Basel II Risk Parameters: Estimation, Validation, and Stress Testing**. Springer Berlin - Heidelberg, 2006

SAUNDERS, Anthony. **Medindo Risco Credito**. Qualitymark, 2000.

TRAIN, K. **Discrete choice methods with simulation**. Cambridge: Cambridge University Press, 2003.

VASICEK, O. The loan loss distribution. Working paper, **KMV Corporation**, 1997.

WEST, R.R. An alternative approach to predicting corporate bond ratios, **Journal of Accounting Research**, v. 7, spring, p. 118-127, 1970.

WOOLDRIDGE, Jeffrey M. **Econometric analysis of cross section and panel data**, The MIT Press, Cambridge, Massachusetts, London, 2002.

APÊNDICES

Modelos Lineares em Cross-Section

Modelo 1 – apenas variáveis explicativas

Sintaxe SATA v11.0:

```
regress rating_moddds1 ln_asset debt_asset ebit_debt roa ebit_net_income vol
```

Output SATA v11.0:

Source	SS	df	MS			
Model	11477.6071	6	1912.93451	Number of obs = 2736		
Residual	19648.5903	2729	7.19992316	F(6, 2729) = 265.69		
Total	31126.1974	2735	11.3806937	Prob > F = 0.0000		
				R-squared = 0.3687		
				Adj R-squared = 0.3674		
				Root MSE = 2.6833		

rating_mod~1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ln_asset	-1.017201	.0431427	-23.58	0.000	-1.101797	-.9326055
debt_asset	-.0353286	.0197045	-1.79	0.073	-.0739657	.0033086
ebit_debt	-7.66e-06	7.36e-06	-1.04	0.298	-.0000221	6.78e-06
roa	-.0953624	.0063295	-15.07	0.000	-.1077735	-.0829513
ebit_net_i~e	1.262556	.2513372	5.02	0.000	.7697257	1.755387
vol	.0353584	.0023	15.37	0.000	.0308485	.0398684
_cons	24.41705	.7088868	34.44	0.000	23.02704	25.80705

Modelo 2 – variáveis explicativas mais dummies de período

Sintaxe SATA v11.0:

```
regress rating_moddds1 ln_asset debt_asset ebit_debt roa ebit_net_income vol d_per_1-d_per_9
```

Output SATA v11.0:

Source	SS	df	MS			
Model	13798.3672	15	919.891144	Number of obs = 2736		
Residual	17327.8302	2720	6.37052581	F(15, 2720) = 144.40		
Total	31126.1974	2735	11.3806937	Prob > F = 0.0000		
				R-squared = 0.4433		
				Adj R-squared = 0.4402		
				Root MSE = 2.524		

rating_mod~1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ln_asset	-.8406285	.0419213	-20.05	0.000	-.9228293	-.7584277
debt_asset	-.0244524	.0185534	-1.32	0.188	-.0608326	.0119279
ebit_debt	-7.57e-06	6.94e-06	-1.09	0.275	-.0000212	6.04e-06
roa	-.0836272	.006021	-13.89	0.000	-.0954334	-.0718211
ebit_net_i~e	1.103088	.2374076	4.65	0.000	.6375707	1.568606
vol	.0775651	.0031523	24.61	0.000	.0713839	.0837463
d_per_1	3.292154	.2657473	12.39	0.000	2.771067	3.813241
d_per_2	3.366895	.2612728	12.89	0.000	2.854582	3.879209
d_per_3	3.485449	.2591527	13.45	0.000	2.977292	3.993605
d_per_4	3.4579	.2509569	13.78	0.000	2.965815	3.949986
d_per_5	3.704544	.2552046	14.52	0.000	3.204129	4.204958
d_per_6	3.877311	.2553859	15.18	0.000	3.376541	4.378081
d_per_7	3.6676	.2494981	14.70	0.000	3.178375	4.156825
d_per_8	2.917902	.2312398	12.62	0.000	2.464479	3.371326
d_per_9	.4214492	.2096259	2.01	0.044	.0104071	.8324912
_cons	17.13164	.7849772	21.82	0.000	15.59243	18.67085

Modelos Não- Lineares em Cross-Section

Modelo 4 – Probit Ordenado

Sintaxe SATA v11.0:

oprobit rating_moddys1 ln_asset debt_asset ebit_debt roa ebit_net_income vol d_per_1-d_per_9

Output SATA v11.0:

```
Iteration 0: log likelihood = -7031.7642
Iteration 1: log likelihood = -6246.3923
Iteration 2: log likelihood = -6238.6381
Iteration 3: log likelihood = -6238.428
Iteration 4: log likelihood = -6238.4248
Iteration 5: log likelihood = -6238.4248
```

```
ordered probit regression          Number of obs   =      2736
LR chi2(15)                       =     1586.68
Prob > chi2                         =      0.0000
Pseudo R2                           =      0.1128

Log likelihood = -6238.4248
```

rating_mod~1	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ln_asset	-.3408008	.0175946	-19.37	0.000	-.3752856	-.3063159
debt_asset	-.006825	.0074267	-0.92	0.358	-.0213811	.007731
ebit_debt	-4.01e-06	2.78e-06	-1.44	0.149	-9.45e-06	1.43e-06
roa	-.0356595	.0024787	-14.39	0.000	-.0405177	-.0308012
ebit_net_i~e	.4803667	.0953699	5.04	0.000	.2934452	.6672882
vol	.0305238	.0013611	22.43	0.000	.027856	.0331916
d_per_1	1.281029	.1085	11.81	0.000	1.068373	1.493685
d_per_2	1.31415	.1067612	12.31	0.000	1.104902	1.523398
d_per_3	1.358614	.1059418	12.82	0.000	1.150972	1.566256
d_per_4	1.353412	.1026286	13.19	0.000	1.152263	1.55456
d_per_5	1.454485	.1046615	13.90	0.000	1.249353	1.659618
d_per_6	1.517116	.1049518	14.46	0.000	1.311414	1.722818
d_per_7	1.439443	.1022705	14.07	0.000	1.238996	1.639889
d_per_8	1.13798	.0941376	12.09	0.000	.9534735	1.322486
d_per_9	.1485085	.0839001	1.77	0.077	-.0159326	.3129497
/cut1	-6.485879	.3538881			-7.179486	-5.792271
/cut2	-6.307348	.3465831			-6.986639	-5.628058
/cut3	-5.89808	.3364526			-6.557515	-5.238646
/cut4	-5.374102	.3303017			-6.021482	-4.726723
/cut5	-4.749902	.3265051			-5.38984	-4.109964
/cut6	-4.28385	.3243548			-4.919573	-3.648126
/cut7	-3.889923	.3229379			-4.52287	-3.256976
/cut8	-3.425928	.3216257			-4.056303	-2.795553
/cut9	-2.97218	.3208672			-3.601068	-2.343292
/cut10	-2.763186	.3207515			-3.391847	-2.134525
/cut11	-2.526165	.3205664			-3.154464	-1.897867
/cut12	-2.173252	.3202325			-2.800896	-1.545607
/cut13	-1.668402	.3206558			-2.296876	-1.039928
/cut14	-1.277883	.322277			-1.909534	-.6462313
/cut15	-.7786391	.3262635			-1.418104	-.1391744
/cut16	-.1576248	.3346927			-.8136104	.4983609
/cut17	.508178	.3533985			-.1844703	1.200826
/cut18	.6046299	.3565714			-.0942372	1.303497
/cut19	2.732463	1.078996			.6176696	4.847256

Modelos Lineares em Painel

Modelo 6 – Efeitos Fixos

Sintaxe SATA v11.0:

xtreg rating_mod~1 ln_asset debt_asset ebit_debt roa ebit_net_income vol d_per_1-d_per_9, fe

Output SATA v11.0:

```

Fixed-effects (within) regression          Number of obs   =   2736
Group variable: id                       Number of groups =    326

R-sq:  within = 0.1792                   Obs per group:  min =    1
        between = 0.5105                  avg   =    8.4
        overall = 0.4298                  max   =   10

corr(u_i, Xb) = 0.5298                   F(15,2395)     =   34.86
                                                Prob > F       =   0.0000
    
```

rating_mod~1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ln_asset	-.3049379	.0772845	-3.95	0.000	-.4564893	-.1533864
debt_asset	.0114133	.0082762	1.38	0.168	-.0048161	.0276426
ebit_debt	1.66e-06	2.62e-06	0.63	0.527	-3.48e-06	6.80e-06
roa	-.0274778	.002616	-10.50	0.000	-.0326078	-.0223479
ebit_net_income	.3017519	.089005	3.39	0.001	.1272171	.4762867
vol	.020083	.0017329	11.59	0.000	.0166849	.0234812
d_per_1	.7800383	.1094858	7.12	0.000	.5653416	.9947351
d_per_2	.7731138	.1082304	7.14	0.000	.5608788	.9853488
d_per_3	.7248377	.1068551	6.78	0.000	.5152997	.9343757
d_per_4	.6934615	.1023778	6.77	0.000	.4927032	.8942198
d_per_5	.7504753	.103962	7.22	0.000	.5466104	.9543401
d_per_6	.7986039	.1056125	7.56	0.000	.5915025	1.005705
d_per_7	.7238169	.1006018	7.19	0.000	.5265413	.9210924
d_per_8	.57948	.088465	6.55	0.000	.4060041	.7529559
d_per_9	-.0365845	.0692717	-0.53	0.597	-.1724233	.0992542
_cons	12.80102	1.244813	10.28	0.000	10.36	15.24204
sigma_u	2.9195156					
sigma_e	.81297617					
rho	.92803853	(fraction of variance due to u_i)				
F test that all u_i=0:			F(325, 2395) =	73.30	Prob > F = 0.0000	

Modelo 8 – Efeitos Aleatórios

Sintaxe SATA v11.0:

xtreg rating_moddds1 ln_asset debt_asset ebit_debt roa ebit_net_income vol d_per_1-d_per_9, re

Output SATA v11.0:

```

Random-effects GLS regression           Number of obs   =    2736
Group variable: id                     Number of groups =    326

R-sq:  within = 0.1762                  Obs per group:  min =     1
        between = 0.4413                  avg           =    8.4
        overall = 0.4011                  max           =    10

Random effects u_i ~ Gaussian          wald chi2(15)   =   685.00
corr(u_i, X) = 0 (assumed)             Prob > chi2     =   0.0000

```

rating_mod~1	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
ln_asset	-.5777436	.062292	-9.27	0.000	-.6998336 - .4556535
debt_asset	.012302	.0085482	1.44	0.150	-.0044521 .0290562
ebit_debt	9.35e-07	2.72e-06	0.34	0.731	-4.39e-06 6.26e-06
roa	-.026907	.0026815	-10.03	0.000	-.0321626 -.0216514
ebit_net_i~e	.3030119	.092245	3.28	0.001	.122215 .4838088
vol	.0235241	.0017694	13.30	0.000	.0200562 .0269919
d_per_1	.8488228	.1120047	7.58	0.000	.6292976 1.068348
d_per_2	.8540824	.1108359	7.71	0.000	.6368481 1.071317
d_per_3	.821284	.1096533	7.49	0.000	.6063675 1.0362
d_per_4	.7970423	.1052053	7.58	0.000	.5908437 1.003241
d_per_5	.8746991	.1069939	8.18	0.000	.6649949 1.084403
d_per_6	.9495666	.1087358	8.73	0.000	.7364483 1.162685
d_per_7	.8774856	.1036907	8.46	0.000	.6742555 1.080716
d_per_8	.7148283	.0913168	7.83	0.000	.5358507 .8938058
d_per_9	-.0094634	.072065	-0.13	0.896	-.1507082 .1317813
_cons	17.01852	1.020378	16.68	0.000	15.01862 19.01842
sigma_u	2.1439212				
sigma_e	.81297617				
rho	.87428401	(fraction of variance due to u_i)			

Teste de Hasuman

	Coefficients		(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
	(b) fixed	(B) .		
ln_asset	-.3049379	-.5777436	.2728057	.051069
debt_asset	.0114133	.012302	-.0008887	.0011557
ebit_debt	1.66e-06	9.35e-07	7.22e-07	2.76e-07
roa	-.0274778	-.026907	-.0005708	.0004939
ebit_net_i~e	.3017519	.3030119	-.00126	.0098169
vol	.020083	.0235241	-.003441	.0003625
d_per_1	.7800383	.8488228	-.0687845	.0218294
d_per_2	.7731138	.8540824	-.0809686	.0209782
d_per_3	.7248377	.821284	-.0964463	.0194804
d_per_4	.6934615	.7970423	-.1035808	.0178197
d_per_5	.7504753	.8746991	-.1242239	.0171201
d_per_6	.7986039	.9495666	-.1509627	.0171193
d_per_7	.7238169	.8774856	-.1536687	.0155668
d_per_8	.57948	.7148283	-.1353483	.0127538
d_per_9	-.0365845	-.0094634	-.0271211	.0043943

b = consistent under Ho and Ha; obtained from xtreg
 B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

$$\begin{aligned} \text{chi2}(14) &= (b-B)'[(V_b-V_B)^{-1}](b-B) \\ &= 229.80 \\ \text{Prob}>\text{chi2} &= 0.0000 \end{aligned}$$

Modelo 10 – Binários de Efeito Fixo

Modelo 10.1 - Regressão rating 1

Sintaxe SAS 9.2:

```
proc logistic data=IMPW desc;
model d_rm2_1 = ln_asset debt_asset ebit_debt roa ebit_net_income vol
d_firma_1-d_firma_325;
run;
```

Output SAS 9.2:

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	5.0442	0.00816	382387.519	<.0001
ln_asset	1	-0.00913	0.000513	316.7064	<.0001
debt_asset	1	-0.00246	0.00167	2.1684	0.1409
ebit_debt	1	2.093E-6	1.334E-6	2.4620	0.1166
roa	1	-0.0103	0.000818	157.9089	<.0001
ebit_net_income	1	-0.1052	0.0337	9.7610	0.0018
vol	1	0.00390	0.000190	421.0094	<.0001

Modelo 10.2 - Regressão rating 2

Sintaxe SAS 9.2:

```
proc logistic data=IMPW desc;
model d_rm2_2 = ln_asset debt_asset ebit_debt roa ebit_net_income vol
d_firma_1-d_firma_325;
run;
```

Output SAS 9.2:

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	3.9353	0.00535	540483.007	<.0001
ln_asset	1	-0.00632	0.000337	350.9477	<.0001
debt_asset	1	-0.0106	0.00109	93.6430	<.0001
ebit_debt	1	-0.00002	4.407E-6	15.3448	<.0001
roa	1	-0.00276	0.000511	29.2418	<.0001
ebit_net_income	1	-0.0323	0.0208	2.4026	0.1211
vol	1	0.00451	0.000122	1369.9156	<.0001

Modelo 10.3 - Regressão rating 3

Sintaxe SAS 9.2:

```
proc logistic data=IMPW desc;
model d_rm2_3 = ln_asset debt_asset ebit_debt roa ebit_net_income vol
d_firma_1-d_firma_325;
run;
```

Output SAS 9.2:

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	30.3312	746.8	0.0016	0.9676
ln_asset	1	-0.6436	0.9037	0.5071	0.4764
debt_asset	1	0.0822	0.1227	0.4488	0.5029
ebit_debt	1	-0.00006	0.000370	0.0291	0.8646
roa	1	-0.1900	0.0793	5.7391	0.0166
ebit_net_income	1	-6.2677	4.7009	1.7777	0.1824
vol	1	0.0388	0.0141	7.5873	0.0059

Modelo 10.4 - Regressão rating 4

Sintaxe SAS 9.2:

```
proc logistic data=IMPW desc;
model d_rm2_4 = ln_asset debt_asset ebit_debt roa ebit_net_income vol
d_firma_1-d_firma_325;
run;
```

Output SAS 9.2:

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	151.2	1055.4	0.0205	0.8861
ln_asset	1	-7.6605	1.3205	33.6560	<.0001
debt_asset	1	-0.0400	0.0992	0.1625	0.6869
ebit_debt	1	0.000377	0.000271	1.9285	0.1649
roa	1	-0.1383	0.0339	16.6195	<.0001
ebit_net_income	1	-2.1609	2.2614	0.9130	0.3393
vol	1	0.00690	0.00925	0.5570	0.4555

Modelo 10.5 - Regressão rating 5

Sintaxe SAS 9.2:

```
proc logistic data=IMPW desc;
model d_rm2_5 = ln_asset debt_asset ebit_debt roa ebit_net_income vol
d_firma_1-d_firma_325;
run;
```

Output SAS 9.2:

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	15.9670	6.9698	5.2481	0.0220
ln_asset	1	-0.9937	0.4003	6.1625	0.0130
debt_asset	1	-0.0252	0.0548	0.2123	0.6450
ebit_debt	1	-0.00004	0.000257	0.0213	0.8841
roa	1	-0.0789	0.0197	16.1232	<.0001
ebit_net_income	1	0.6692	0.5889	1.2913	0.2558
vol	1	-0.00008	0.00577	0.0002	0.9896

Modelo 10.6 - Regressão rating 6

Sintaxe SAS 9.2:

```
proc logistic data=IMPW desc;
model d_rm2_6 = ln_asset debt_asset ebit_debt roa ebit_net_income vol
d_firma_1-d_firma_325;
run;
```

Output SAS 9.2:

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	11.0986	11.7229	0.8963	0.3438
ln_asset	1	-0.9593	0.6397	2.2490	0.1337
debt_asset	1	0.0949	0.0779	1.4834	0.2232
ebit_debt	1	-4.73E-7	0.000050	0.0001	0.9925
roa	1	-0.0289	0.0175	2.7379	0.0980
ebit_net_income	1	0.1837	0.4775	0.1480	0.7005
vol	1	0.0189	0.00674	7.8388	0.0051

Modelo 10.7 - Regressão rating 7

Sintaxe SAS 9.2:

```
proc logistic data=IMPW desc;  
model d_rm2_7 = ln_asset debt_asset ebit_debt roa ebit_net_income vol  
d_firma_1-d_firma_325;  
run;
```

Output SAS 9.2:

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	173.5	788.8	0.0484	0.8259
ln_asset	1	-10.5658	5.8755	3.2338	0.0721
debt_asset	1	0.2663	0.3343	0.6348	0.4256
ebit_debt	1	0.000070	0.0442	0.0000	0.9987
roa	1	-0.3145	0.1883	2.7890	0.0949
ebit_net_income	1	-10.0270	13.0697	0.5886	0.4430
vol	1	-0.00654	0.0179	0.1342	0.7142