



UNIVERSIDADE DE BRASÍLIA
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO, CONTABILIDADE E GESTÃO
DE POLÍTICAS PÚBLICAS
DEPARTAMENTO DE CIÊNCIAS CONTÁBEIS E ATUARIAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS CONTÁBEIS
LINHA DE PESQUISA 2: IMPACTOS DA CONTABILIDADE NO SETOR
PÚBLICO, NAS ORGANIZAÇÕES E NA SOCIEDADE

SÉRGIO CARLOS DOS SANTOS

DESEMPENHOS CONTÁBIL E DE MERCADO DE CURTO E
LONGO PRAZOS: UMA ABORDAGEM MULTINOMIAL

BRASÍLIA (DF)

2020

Professora Doutora Márcia Abrahão Moura

Reitora da Universidade de Brasília

Professor Doutor Enrique Huelva Unternbäumen

Vice-Reitor da Universidade de Brasília

Professora Doutora Adalene Moreira Silva

Decana de Pós-Graduação

Professor Doutor Eduardo Tadeu Vieira

**Diretor da Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas
Públicas**

Professor Doutor Paulo César de Melo Mendes

Chefe do Departamento de Ciências Contábeis e Atuariais

Professor Doutor César Augusto Tibúrcio Silva

Coordenador do Programa de Pós-Graduação em Ciências Contábeis

SÉRGIO CARLOS DOS SANTOS

**DESEMPENHOS CONTÁBIL E DE MERCADO DE CURTO E
LONGO PRAZOS: UMA ABORDAGEM MULTINOMIAL**

Tese submetida à apreciação da banca examinadora do Programa de Pós-Graduação em Ciências Contábeis – PPGCont da Universidade de Brasília (UnB), como requisito parcial à obtenção do título de Doutor em Ciências Contábeis.

Orientador: Prof. Dr. Paulo Roberto Barbosa Lustosa.

BRASÍLIA (DF)

2020

SÉRGIO CARLOS DOS SANTOS

**DESEMPENHOS CONTÁBIL E DE MERCADO DE CURTO E
LONGO PRAZOS: UMA ABORDAGEM MULTINOMIAL**

Tese submetida à apreciação da banca examinadora do Programa de Pós-Graduação em Ciências Contábeis – PPGCont da Universidade de Brasília (UnB), como requisito parcial à obtenção do título de Doutor em Ciências Contábeis.

Professor Doutor Paulo Roberto Barbosa Lustosa

Universidade de Brasília – UnB

Orientador

Professor Doutor José Elias Feres de Almeida

UFES – Universidade Federal do Espírito Santo

Examinador (Membro externo)

Professor Doutor Rodrigo dos Santos Verdi

MIT – *Massachusetts Institute of Technology*

Examinador (Membro externo)

Professor Doutor César Augusto Tibúrcio Silva

Universidade de Brasília – UnB

Examinador (Membro interno)

Professor Paulo Augusto Pettenuzzo de Brito, Ph.D.

Universidade de Brasília – UnB

Examinador (Suplente)

DEDICATÓRIA

À minha esposa Nádia, meus filhos Carolina e Gabriel, e todos os meus familiares, em especial minha mãe e meus irmãos Sara e Sandro.

A confiança de vocês é que me anima a continuar buscando transpor todos os obstáculos.

AGRADECIMENTOS

Ao Professor Doutor Paulo Roberto Barbosa Lustosa que aceitou o desafio de orientar-me neste trabalho e mais do que um orientador, foi um grande amigo durante essa jornada. Minha eterna admiração pelo seu caráter, inteligência, disposição e companheirismo, tornando possível a realização deste trabalho.

Ao colega de Banco do Brasil (BB) Maurício Marchiori Ono, Superintendente Comercial do BB à época de meu ingresso no programa de doutorado, que acreditou e me apoiou nessa empreitada de conciliar estudo e trabalho. Em seu nome agradeço à toda família BB, presente na minha formação pessoal e profissional desde os meus 15 anos de idade quando ingressei naquela instituição. Minha especial gratidão à minha colega de BB Alice Pereira Oliveira Barreto que com muita excelência e profissionalismo conduziu nosso Escritório de Negócios durante minhas diversas ausências em razão da realização do curso de doutorado. Além destas, muitas foram as pessoas do BB que contribuíram para que eu realizasse este curso, mesmo não citadas, sintam-se lembradas.

À Professora Soraya Pedrosa, Coordenadora do Curso de Ciências Contábeis da Faculdade Araguaia de Goiânia-GO (FARA). Em seu nome agradeço aos Dirigentes, Coordenadores, Professores e Alunos da FARA que me apoiaram durante meu curso de doutorado, permitindo que eu conciliasse meus estudos e o trabalho de Professor nos cursos de Ciências Contábeis e Gestão Comercial daquela instituição.

Aos professores Otávio Medeiros, João Bilhim, José Matias-Pereira, Andréa Gonçalves e Suresh Radhakrishnan com os quais muito aprendi durante o curso. Aos companheiros de turma, Eliedna de Sousa Barbosa, Nara Cristina Ferreira Mendes, Maria Lizete da Silveira, Edmilson Soares Campos, Carlos Vicente Berner, Olavo Venturim Caldas, José Bonifácio de Araújo Junior, Sérgio Ricardo Miranda Nazaré, Karla Roberta Castro Pinheiro Alves e, em especial, Luis Gustavo do Lago Quinteiro que me convidou e insistiu que eu participasse do processo seletivo para ingresso neste Programa de Pós-Graduação. Ao colega Alex Laquis Resende, em seu nome agradeço aos colegas da turma anterior com os quais convivi durante trabalhos e disciplinas, e ao colega Joaquim Ramalho de Albuquerque que apesar de pertencer à turma seguinte, como aluno especial em nossa turma compartilhou trabalhos e disciplinas, auxiliando-me na etapa preliminar do desenvolvimento da metodologia aplicada neste trabalho.

Aos professores César Augusto Tibúrcio Silva (PPGCont/UnB), Jorge Katsumi Niyama (PPGCont/UnB), José Elias Feres de Almeida (UFES – Universidade Federal do Espírito Santo) e Rodrigo dos Santos Verdi (MIT – *Massachusetts Institute of Technology*) pelas sugestões e

comentários apresentados durante a qualificação do projeto desta tese e, posteriormente, na defesa da tese, contribuindo para desenvolver o relatório final e estimular-me a novas pesquisas por meio de seus exemplos em produção científica.

Aos colaboradores do Departamento de Ciências Contábeis da UnB, principalmente aqueles vinculados ao Programa de Pós-Graduação em Ciências Contábeis (PPGCont).

À Nádia, minha esposa e amiga, e meus filhos Carolina e Gabriel pela paciência e compreensão durante esse período de trabalho em que muitas vezes eu não podia oferecer-lhes toda a atenção que vocês merecem. A alegria de vocês é que me motiva a buscar sempre mais!

À minha mãe, sempre presente, e ao meu pai, que já não está entre nós. Seus incentivos e confiança, desde a minha infância, permanecerão eternamente em meu coração.

A Deus, Nosso Senhor, que foi e será refúgio seguro nos momentos de dificuldade.

Muito obrigado!

EPÍGRAFE

“Quanto maiores são as dificuldades a vencer,
maior será a satisfação.”

Cícero (106–43 a.C.).

RESUMO

DESEMPENHOS CONTÁBIL E DE MERCADO DE CURTO E LONGO PRAZOS: UMA ABORDAGEM MULTINOMIAL

O propósito da pesquisa é verificar se as informações de longo prazo seriam mais úteis para explicar a relação entre informações contábeis e retornos de ações do que as de curto prazo sob o pressuposto de que no longo prazo os indicadores contábeis estão associados aos retornos das ações, permitindo a realização de previsões. O estudo parte de um constructo de desempenho de mercado, estabelecendo três categorias que permitem utilizar modelos de regressão logística multinomial. São comparados os percentuais de acertos das previsões realizadas em dois modelos, um de curto e outro de longo prazo. A amostra com 500.479 observações no período de 23 anos (1995 a 2017), de 50.998 empresas, de 18 países, distribuídas pelos cinco continentes e englobando os cinco países do BRICS, confere caráter global à pesquisa. Na comparação entre os modelos de dados anuais e quinquenais, foram encontrados indícios de que no curto prazo os desempenhos contábil e de mercado das empresas não tem relação, ou se tiverem é bastante limitado, por outro lado essa relação existe no longo prazo e os indicadores contábeis mais significativos são os de giro do ativo e lucratividade. O percentual total de acertos no curto prazo foi de 36,72%, enquanto no longo prazo foi 44,14%. A análise e comparação, em recorte da amostra com empresas dos EUA, contra o resultado da amostra de 18 países, foi um teste de robustez para os resultados encontrados. A verificação de que no longo prazo a informação contábil é mais útil foi uma inovação do estudo, o qual apresenta limitações em função de que seu foco não abrangia analisar ou explicar as inúmeras informações e variáveis que afetam o retorno de ações nos diversos países que compõem a amostra, pois, para tal verificação não era necessário prever o valor do retorno futuro e sim em qual categoria de nível de retorno a ação da empresa estaria no futuro dentre as categorias baixo, médio e alto retorno, sendo confirmada a hipótese pela comparação entre os resultados de previsão dos dois modelos, um de curto e outro de longo prazo. Os modelos de previsão desenvolvidos têm por base informações públicas e sinalizam ter utilidade para investidores em ações com objetivos de longo prazo. O caráter global do estudo abre as portas para futuras pesquisas que busquem explicar os diferentes comportamentos de mercados nacionais identificados neste trabalho.

Palavras-chave: hipótese de eficiência de mercado; relevância da informação contábil; retorno de ações; análise de desempenho; modelo logit multinomial.

ABSTRACT

SHORT AND LONG TERM ACCOUNTING AND MARKET PERFORMANCES: A MULTINOMIAL APPROACH

The purpose of the research is to verify whether long-term information would be more useful to explain the relationship between accounting information and stock returns than short-term information under the assumption that in the long term, financial ratios are associated with stock returns, allowing to make predictions. The study is based on a market performance construct, establishing three categories that allow using of multinomial logistic regression models. The percentages of correct predictions made in two models (short and long term) are compared. The sample with 500,479 observations over the 23-year period (1995 to 2017), from 50,998 companies, from 18 countries, spread over the five continents and encompassing the five BRICS countries, gives the research a global character. In the comparison between the annual and five-year models, evidence was found that in the short term the accounting and market performances of the companies are unrelated, or if they are very limited, on the other hand this relationship exists in the long term and the Asset Utilization and Profitability Ratios are the most significant financial ratios. The total percentage of correct predictions in the short term was 36.72%, while in the long term it was 44.14%. The analysis and comparison, in a section of the sample with US companies, against the result of the sample with 18 countries, was a robustness testing for the results found. The verification that in the long term the accounting information is more useful was an innovation of the study, which has limitations due to the fact that its focus did not include analyzing or explaining the countless information and variables that affect stock returns in the different countries in the sample, because for that verification it was not necessary to predict the value of the future stock return, but in which category of stock return level the company's stock would be in the future among the low, medium and high stock return categories and the hypothesis would be confirmed by the comparison between the prediction results of the two models, one short term and the other long term. The prediction models developed are based on public information and indicate that they are useful for investors in stocks with long term objectives. The global character of the study opens the door for future research that seeks to explain the different behaviors of national markets identified in this work.

Keywords: Efficient markets hypothesis; value relevance; stock returns; performance analysis; multinomial logit model.

LISTA DE FIGURAS

Figura 1 – Procedimentos esquematizados do MLM de curto prazo	48
Figura 2 – Procedimentos esquematizados do MLM de longo prazo	49

LISTA DE TABELAS

Tabela 1 – Distribuição amostral	49
Tabela 2 – Distribuição em tercís dos retornos anuais e quinquenais	50
Tabela 3 – Curto prazo – toda amostra utilizada (1996 a 2017) – estatística descritiva	51
Tabela 4 – Longo prazo – toda amostra utilizada (2000 a 2017) – estatística descritiva	51
Tabela 5 – Curto prazo – toda amostra utilizada (1996 a 2017) – distribuição em percentis	52
Tabela 6 – Longo prazo – toda amostra utilizada (2000 a 2017) – distribuição em percentis	52
Tabela 7 – Curto prazo – amostra utilizada na estimação (1996 a 2012) – estatística descritiva	94
Tabela 8 – Curto prazo – amostra utilizada na estimação (1996 a 2012) – distribuição em percentis	94
Tabela 9 – Curto prazo – amostra utilizada na previsão (2013 a 2017) – estatística descritiva	95
Tabela 10 – Curto prazo – amostra utilizada na estimação (2013 a 2017) – distribuição em percentis	95
Tabela 11 – Longo prazo – amostra utilizada na estimação (2000 a 2012) – estatística descritiva	96
Tabela 12 – Longo prazo – amostra utilizada na estimação (2000 a 2012) – Distribuição em percentis	96
Tabela 13 – Longo prazo – amostra utilizada na previsão (2013 a 2017) – estatística descritiva	97
Tabela 14 – Longo prazo – amostra utilizada na estimação (2013 a 2017) – Distribuição em percentis	97
Tabela 15 – Correlação das variáveis do MLM de curto prazo	53
Tabela 16 – Correlação das variáveis do MLM de longo prazo	53
Tabela 17 – Estimativas dos Parâmetros do MLM de curto prazo	55
Tabela 18 – Estimativas dos Parâmetros do MLM de longo prazo	56
Tabela 19 – Razão de Risco Relativo (RRR) do MLM de curto prazo	57
Tabela 20 – Razão de Risco Relativo (RRR) do MLM de longo prazo	58
Tabela 21 – Indicadores de desempenho dos MLM	59
Tabela 22 – MLM de curto prazo: acertos de predição	59

Tabela 23 – MLM de longo prazo: acertos de predição	60
Tabela 24 – Efeito marginal dos valores preditos – MLM curto prazo	61
Tabela 25 – Efeito marginal dos valores preditos – MLM longo prazo	61
Tabela 26 – Comparação entre acertos de predição dos MLM de curto e longo prazos – Total geral	63
Tabela 27 – Comparação entre acertos de predição dos MLM de curto e longo prazos – ano 2013	98
Tabela 28 – Comparação entre acertos de predição dos MLM de curto e longo prazos – ano 2014	99
Tabela 29 – Comparação entre acertos de predição dos MLM de curto e longo prazos – ano 2015	100
Tabela 30 – Comparação entre acertos de predição dos MLM de curto e longo prazos – ano 2016	101
Tabela 31 – Comparação entre acertos de predição dos MLM de curto e longo prazos – ano 2017	102
Tabela 32 – Indicadores de desempenho dos MLM – amostra EUA	65
Tabela 33 – MLM de curto prazo: acertos de predição EUA	65
Tabela 34 – MLM de longo prazo: acertos de predição EUA	65
Tabela 35 – Comparação de acertos totais entre as amostras EUA e 18 países	66

LISTA DE QUADROS

Quadro 1 – Índices de mercado por país	38
Quadro 2 – Lista das variáveis utilizadas no trabalho	75
Quadro 3 – MLM de curto e longo prazos: variáveis que representam o constructo de desempenho contábil	40
Quadro 4 – MLM de curto prazo: variáveis que representam os dados macroeconômicos e índices de mercado por país	41
Quadro 5 – MLM de longo prazo: variáveis que representam os dados macroeconômicos e índices de mercado por país	41
Quadro 6 – Lista das variáveis criadas para execução dos modelos anual e quinquenal	91
Quadro 7 – Resumo das variáveis dos MLM de curto e de longo prazos	42

LISTA DE GRÁFICOS

Gráfico 1 – Efeitos marginais do desempenho contábil – MLM de curto prazo	62
Gráfico 2 – Efeitos marginais do desempenho contábil – MLM de longo prazo	62

LISTA DE SIGLAS E ABREVIATURAS

AIC	Critério de Informação de Akaike
BIC	Critério de Informação Bayesiano
BM&FBOVESPA	Bolsa de Valores e de Mercadorias e Futuros de São Paulo
BRICS	Termo que agrupa os principais países de mercado emergente no mundo: Brasil, Rússia, Índia, China e África do Sul
CPC	Comitê de Pronunciamentos Contábeis
EBITDA	<i>Earnings before interest, taxes, depreciation and amortization</i> ou lucro antes de juros, impostos, depreciação e amortização
EUA	Estados Unidos da América
EVA	<i>Economic Value Added</i> ou Valor Econômico Agregado
FASB	<i>Financial Accounting Standards Board</i>
IASB	<i>International Accounting Standards Board</i>
LR	<i>Likelihood ratio</i> ou teste da razão de verossimilhança
MDA	Análise Discriminante Múltipla
MLM	Modelos logit multinomial
RRR	<i>Relative-risk ratios</i> ou razão de risco relativo
STATA	Software estatístico <i>Statistics/Data Analysis</i> versão 15.1
WFA	<i>Worldscope - Fundamentals Annual</i>
WRDS	<i>Wharton Research Data Services</i>
WSD	<i>Worldscope - Stock Data</i>

SUMÁRIO

1. INTRODUÇÃO.....	18
1.1. Contextualização e motivação do estudo.....	18
1.2. Questão da pesquisa.....	21
1.3. Objetivos.....	21
1.3.1 Objetivo geral.....	21
1.3.2. Objetivos específicos.....	22
1.4. Relevância e caráter inovador do estudo	22
1.5. Delimitação do estudo	24
1.6. Estrutura do trabalho.....	25
2. REFERENCIAL TEÓRICO.....	25
2.1. Hipótese de eficiência do mercado	25
2.2. <i>Value relevance</i> – relevância da informação contábil	27
2.2.1. <i>Value relevance</i> e diferenciação pelo mercado do desempenho de longo prazo.....	28
2.2.2. <i>Value relevance</i> , demonstrações financeiras e indicadores contábeis.....	29
2.3. Modelos logit multinomial (MLM)	32
2.3.1. Detalhando o MLM.....	33
3. DESENVOLVIMENTO DA HIPÓTESE DA PESQUISA	35
4. METODOLOGIA DA PESQUISA.....	36
4.1. Justificativa da aplicação do MLM.....	36
4.2. Amostra e fonte de dados.....	37
4.2.1. Extração de dados.....	37
4.2.2. Variáveis utilizadas	38
4.2.2.1. Informações contábeis.....	39
4.2.2.2. Dados macroeconômicos e índices de mercado	40
4.3. Estruturação dos modelos de análises.....	41
4.3.1. Modelos para relação entre desempenhos contábil e de mercado de curto e longo prazos	42

4.3.2. Variáveis dependentes.....	43
4.3.3. Variáveis independentes.....	43
4.3.4. Modelo Econométrico	44
4.4. Procedimentos.....	45
5. RESULTADOS	48
5.1. Estatística descritiva	49
5.2. Análises dos MLM.....	54
5.3. Análises das predições realizadas	59
5.4. Teste e procedimentos de robustez	64
5.5. Síntese dos resultados em relação às hipóteses da pesquisa	66
6. CONSIDERAÇÕES FINAIS	68
REFERÊNCIAS	71
APÊNDICE A – Lista das variáveis utilizadas no trabalho	75
APÊNDICE B – Do-file – Amostra 18 países.....	79
APÊNDICE C – Lista das variáveis criadas para execução dos modelos anual e quinquenal	91
APÊNDICE D – Estatísticas descritivas suplementares – MLM curto prazo.....	94
APÊNDICE E – Estatísticas descritivas suplementares – MLM longo prazo	96
APÊNDICE F – Comparação entre acertos de predição dos MLM de curto e longo prazos, por países e ano a ano	98
APÊNDICE G – Do-file – amostra EUA	103

1. INTRODUÇÃO

1.1. Contextualização e motivação do estudo

Existe controvérsia na literatura acerca da capacidade preditiva das informações contábeis em relação aos retornos futuros. Os trabalhos de Ball & Brown (1968) e Beaver (1968) são marcos nos estudos da relação entre a contabilidade e o mercado de capitais, sendo que Kothari (2001, p.207) incluiu também o estudo de Fama et al. (1969). Ball & Brown concluíram que mais da metade de toda a informação de uma empresa disponibilizada durante o ano é capturada naquele ano pelos números de seus resultados, enquanto Beaver observou que as expectativas dos investidores e do mercado como um todo são alteradas pelos anúncios dos resultados e se refletem em mudanças no equilíbrio dos preços das ações. Por outro lado, Lev (1989) concluiu que a correlação entre resultado contábil e retorno de ações é muito baixa e, além do mais, a relação entre essas variáveis é instável ao longo do tempo, sugerindo que informações trimestrais e anuais teriam utilidade limitada para tomada de decisão de investidores.

Pouco antes dos estudos de Ball & Brown (1968) e Beaver (1968), Horrigan (1966) afirmou que indicadores financeiros e dados contábeis podem ser úteis na gestão de crédito de longo prazo enquanto de maneira mais conservadora Benston (1967) concluiu que, apesar de significativa, somente uma pequena relação foi encontrada entre publicação de dados contábeis e preços de ações, de forma que seu estudo apresenta a informação contida em relatórios contábeis publicados como sendo uma porção relativamente pequena da informação utilizada pelos investidores na tomada de decisão.

Inspirado por essa controvérsia acerca da relação entre informações contábeis e retornos de ações e pelo estudo de Rosa & Lustosa (2014) cujos resultados sinalizaram fortes evidências de que o mercado efetivamente diferencia as empresas de alto desempenho operacional de longo prazo, nasceu a ideia de se verificar se as informações de longo prazo seriam mais úteis para explicar a relação entre informações contábeis e retornos de ações.

Verificar a relação entre as informações contábeis e o retornos das ações se alinha com o objetivo do relatório contábil que segundo o FASB – *Financial Accounting Standards Board* (FASB, 2016), o IASB – *International Accounting Standards Board* (IASB, 2018) e, no Brasil, o Pronunciamento Conceitual Básico (R1) do Comitê de Pronunciamentos Contábeis (CPC), convergente com as Normas Internacionais de Contabilidade, seria fornecer informações contábil-financeiras acerca da entidade que reporta essa informação que sejam úteis a

investidores existentes e em potencial quando da tomada de decisão ligada ao fornecimento de recursos para a entidade (CPC, 2011). Depreende-se deste objetivo que as informações contábeis apresentariam poder preditivo do desempenho econômico futuro da entidade que reporta e dessa forma servir de base para a predição de retornos das ações de empresas comercializadas em bolsas de valores.

Quanto à metodologia do modelo de análise, este estudo partiu de um constructo de desempenho de mercado que ordena o retorno das ações das empresas que compõem a amostra, dividindo-os em tercís de forma a estabelecer três categorias (baixo, médio e alto retornos), semelhante a segregação utilizada no estudo de Rosa & Lustosa (2014), embora aquele trabalho utilizasse o desempenho operacional para segregar as empresas em alto, médio e baixo desempenho.

O estabelecimento das três categorias de retornos das ações permitiu utilizar um modelo de regressão logística multinomial que em estudos contábeis normalmente tem se destinado a modelos de predição de insolvência. Nesse tipo de estudo, Ohlson (1980) apresentou vantagens do logit sobre os modelos clássicos e a escolha do modelo logit neste estudo se justifica por se equiparar aos métodos de análise mais modernos, como demonstraram Mselmi, Lahiani & Hamza (2017) que compararam em seu trabalho um modelo logit com outros quatro modelos de metodologias mais modernas, como rede neural e inteligência artificial, concluindo que todos os cinco métodos utilizados atingiam as expectativas de poder de predição, superando em eficiência estudos anteriores com modelos lineares clássicos e sem conseguir provar a superioridade dos quatro métodos mais modernos sobre o modelo logit, ou seja, nenhum dos métodos testados por Mselmi, Lahiani & Hamza (2017, p.78), segundo os autores, apresentou absoluta superioridade de performance sobre o modelo logit.

A verificação de se as informações de longo prazo seriam mais úteis para explicar a relação entre informações contábeis e retornos de ações foi realizada neste estudo por meio da comparação entre os resultados de predição entre dois modelos, um de curto prazo que utilizou informações anuais e outro de longo prazo com informações quinquenais.

O pressuposto desta pesquisa é que ao alargar a janela temporal, as tendências do desempenho contábil melhor se alinhariam às do desempenho de mercado, pois os dados contábeis anuais divulgados pelas empresas estariam influenciados por diversos fatores políticos, econômicos e sociais presentes naquele exercício que ao mesmo tempo que impactam as decisões, as estratégias e as atividades das empresas na gestão de seus negócios, atuam nas expectativas dos investidores e suas decisões de compra e venda, muitas vezes tornando menos

relevante o desempenho contábil da empresa em face de acontecimentos que afetam o mercado de capitais como um todo, como acontece em crises ou bolhas financeiras, por exemplo.

Assim, os resultados deste estudo sinalizaram que os desempenhos contábil e de mercado num período de 5 anos apresentam mais informações relevantes acerca da tendência da gestão da empresa e sua interação com o mercado, suavizando tanto na empresa quanto no mercado os impactos de fatores externos a essas empresas, tais como: crises ou bolhas econômicas, movimentos pontuais de crescimento ou de recessão da economia nacional ou mundial etc. Tal confirmação veio dos resultados da aplicação dos dois modelos, com informações anuais e quinquenais, cujos resultados de previsões comparados demonstraram que o modelo com informações quinquenais ofereceu maior capacidade preditiva do que o de informações anuais, graças ao melhor ajuste entre as tendências de desempenhos contábil e de mercado no longo prazo quando comparados com o modelo de curto prazo.

Na construção dos modelos, ao invés de adotar variáveis independentes consagradas pela literatura de *value relevance* (relevância da informação contábil), como lucro e fluxo de caixa, foram adotados neste estudo indicadores contábeis disponíveis e, normalmente, utilizados no mercado que analisados na amostra deste trabalho apresentaram maior correlação com o retorno das ações das empresas analisadas. A utilização de indicadores contábeis como base de predição é prática comum em estudos contábeis, como os trabalhos seminais de Beaver (1966) e Altman (1968), por exemplo.

Além dos indicadores contábeis, foram acrescentados aos modelos informações macroeconômicas e de mercado sob a premissa de que tal combinação melhoraria a acurácia dos modelos, conforme se percebeu no estudo de Shumway (2001) e He et al. (2005).

A utilização de regressão logística se amparou em estudos contábeis como os de Ohlson (1980), Shumway (2001), Johnsen & Melicher (1994), Ward (1999) e Lau (1987) que apresentaram que tal metodologia apresentou resultados preditivos superiores aos realizados com regressões lineares e, ainda, que a adição de uma terceira categoria melhorava o poder preditivo dos modelos, avançando de binomiais para multinomiais.

Por fim, a ideia principal é que o preço das ações contém em sua formação a expectativa dos investidores que, apesar de ser influenciada por inúmeras outras informações, utilizam-se também das informações contábeis para a tomada de decisão. Sob a premissa de eficiência de mercado semiforte em que as informações de desempenho de uma empresa são assimiladas pelo mercado e prontamente impostas ao preço de suas ações (HE ET AL., 2005, p.6) e dada a sazonalidade de divulgação ao público das informações contábeis, este trabalho de maneira inovadora verificou que as informações de longo prazo são mais úteis para explicar a relação

entre informações contábeis e retornos de ações por meio da comparação das predições com dados anuais e quinquenais.

1.2. Questão da pesquisa

Dada a controvérsia existente na literatura acerca da capacidade preditiva das informações contábeis em relação aos retornos futuros, a pesquisa buscou responder a seguinte questão: **as informações de longo prazo são mais úteis para explicar a relação entre informações contábeis e retornos de ações do que as de curto prazo?**

A premissa de desenvolvimento do trabalho foi de que ao se ampliar o prazo de coleta de informações dos desempenhos contábil e de mercado, tais constructos melhor capturariam as tendências de desempenho, permitindo ao modelo identificar suas relações e, conseqüentemente, apresentar maior capacidade preditiva.

Os resultados deste trabalho sinalizaram que tal premissa é verdadeira, pois a utilidade das informações contábeis para a predição de retornos foi ampliada com tais informações em prazo mais longo (quinquenais) em comparação com as de prazo mais curto (anuais), em linha com os achados de Lev (1989) de que é muito baixa a correlação entre resultado contábil e retorno de ações no curto prazo uma vez que as informações trimestrais e anuais, segundo Lev (1989), apresentariam utilidades limitadas para a tomada de decisão de investidores.

1.3. Objetivos

Baseado no problema da pesquisa definido são apresentados os seguintes objetivos geral e específicos:

1.3.1 Objetivo geral

Verificar se as informações de longo prazo seriam mais úteis para explicar a relação entre informações contábeis e retornos de ações do que as de curto prazo no intuito de embasar a hipótese de que as informações contábeis têm poder preditivo para o desempenho de mercado das empresas, representado pelo preço de suas ações, quando observadas no longo prazo. Para efeito deste trabalho, longo prazo foi considerado cinco anos, enquanto o curto prazo foi representado por informações anuais.

1.3.2. Objetivos específicos

Aplicar a metodologia de regressão logística multinomial por meio de constructo de desempenho de mercado para representar categorias da variável dependente a fim de apurar o poder preditivo de dois modelos, um de longo prazo (informações quinquenais) e outro de curto prazo (informações anuais), para por meio da comparação entre os resultados desses modelos verificar o atingimento do objetivo geral deste trabalho, bem como:

- a) analisar e comparar os impactos no poder preditivo do modelo ao se utilizar dados anuais e quinquenais;
- b) analisar e comparar os impactos no poder preditivo do modelo ao utilizar-se de dados de empresas de 18 países diferentes, de todos os continentes, conferindo um caráter global à pesquisa;
- c) analisar e comparar, em recorte especial da amostra, as previsões envolvendo um país específico, contra o resultado global alcançado com os dados de empresas de 18 países, verificando se a metodologia é robusta o suficiente para ser replicada em outras situações.

1.4. Relevância e caráter inovador do estudo

Confirmar que as informações de longo prazo são mais úteis para explicar a relação entre informações contábeis e retornos de ações do que as de curto prazo permitiria a construção de modelos de predição do nível de retorno das ações auxiliando o usuário da informação mitigar o risco de perda e potencializar suas chances de ganho, bem como orientar a atuação da empresa e de seus *stakeholders* a fim de melhorar sua performance, sendo que no atual ambiente competitivo global em que as empresas estão inseridas, são cada vez mais frequentes os momentos de crises econômicas, potencializando-se a necessidade de modelos preditivos confiáveis.

Dado que a construção dos modelos anual e quinquenal baseiam-se em informações públicas e facilmente disponíveis ao mercado, a validação da relação entre informações de longo prazo dos desempenhos contábil e de mercado permitiria a utilização dos modelos para tomada de decisão de investidores em mercado de capitais, segmentando seus investimentos de acordo com suas estratégias de curto ou longo prazo, tornando-se ferramenta de apoio à decisão.

Considerando que a principal função da Contabilidade é gerar informação para a tomada de decisão, em especial aos usuários credores e investidores (CPC, 2011), caberia a

Contabilidade abastecer de dados esses modelos preditivos e nesse sentido esta Ciência vem implementando mudanças que visam melhorar a qualidade das informações aos usuários e, conseqüentemente, oferecer informação ainda mais útil para modelos preditivos, afinal, cabe à Contabilidade a missão de levar informação clara e completa a cada usuário dessa informação, sendo que um de seus desafios seria a maneira pela qual se proveria ao usuário a informação contábil para o processo decisório (SANTOS et al., 2015).

Nesse sentido, trabalhos como o de Lev (1989) ao sinalizar a baixa correlação entre variáveis contábeis e retornos futuros retira importância relativa da Contabilidade em explicar desempenho econômico futuro das empresas e o preço de suas ações no mercado que, de certa forma, é o objetivo e principal função da Contabilidade, estimulando o avanço de pesquisas no sentido de revisitar as práticas contábeis, bem como os modelos de análise e predição a fim de que tais inovações contribuam para a comprovação empírica do referido objetivo da Contabilidade.

O caráter inovador do presente estudo ao focar nas informações de longo prazo para buscar a melhoria da qualidade de predição de retornos de ações a partir de indicadores contábeis utiliza-se da metodologia logit multinomial inspirada em modelos logit binomiais utilizados em estudos como os de Ohlson (1980), Ou & Penman (1989) e Holthausen & Larcker (1992). Para tanto, foi utilizado o retorno de ações como parâmetro de construção das três categorias de forma que as variáveis dependentes estão vinculadas a uma variável de mercado e as variáveis independentes são dados contábeis, macroeconômicos e de mercado sob a premissa de que índices de mercado e contábeis combinados em um mesmo modelo auxilia na melhoria da capacidade preditiva como destacaram em seus estudos Shumway (2001) e He et al. (2005).

A seleção da amostra traz relevância global ao trabalho, pois se compõe de empresas pertencentes a 18 países diferentes, distribuídos por todos os continentes (África, América, Ásia, Europa e Oceania) e englobando os cinco países do BRICS – termo que agrupa os principais países de mercado emergente no mundo: Brasil, Rússia, Índia, China e África do Sul.

As informações contábeis das empresas e os preços de suas ações foram obtidas pelo acesso à base Thomson Reuters por meio do WRDS – *Wharton Research Data Services, University of Pennsylvania*, sendo a amostra composta de 500.479 observações no período de 23 anos (1995 a 2017), envolvendo 18 países e 50.998 empresas diferentes. Essa amostra gerou no modelo de informações anuais 296.625 observações válidas e no modelo de informações quinzenais 153.929.

Dentre os indicadores contábeis disponíveis na base de dados foram testadas suas correlações com os retornos anuais e quinquenais, selecionando o indicador de maior correlação dentro de cada uma das seguintes categorias: liquidez, alavancagem, lucratividade e giro do ativo. Os indicadores contábeis utilizados como constructo de desempenho contábil para os modelos anual e quinquenal estão detalhados na seção 4.2.2.1. (Informações contábeis).

1.5. Delimitação do estudo

A pesquisa foi realizada com base em informações públicas de empresas listadas em bolsas de valores de 18 países, distribuídos por todos os continentes, conferindo um caráter global na análise dos resultados. Foram excluídas as empresas do sistema financeiro dada a peculiaridade da atividade. As informações contábeis das empresas e os preços de suas ações foram as disponíveis na base Thomson Reuters e a amostra selecionada conta com informações de 50.998 empresas diferentes distribuídas ao longo de série histórica que compreende o intervalo entre os anos de 1995 e 2017, nos respectivos anos em que essas empresas tiveram dados disponibilizados na referida base. A série histórica iniciando-se em 1995 faz com que as informações das empresas brasileiras presentes na amostra estejam no âmbito do Plano Real.

Dadas as inúmeras informações e variáveis que afetam o retorno de ações que é a base da definição das variáveis categóricas dos modelos testados, não foi intenção deste trabalho determinar tais retornos e sim verificar se as informações de longo prazo seriam mais úteis para explicar a relação entre informações contábeis e retornos de ações do que as de curto prazo sob o pressuposto de que o alargamento da janela temporal das informações contábeis aumenta sua relação com o desempenho de mercado das empresas, representado pelo preço de suas ações.

Para tal verificação não era necessário prever o valor do retorno futuro e sim em qual categoria de nível de retorno a ação da empresa estaria no futuro dentre as categorias baixo, médio e alto retorno. Assim, foi realizada comparação entre os resultados de predição entre dois modelos, um de curto prazo (informações anuais) e outro de longo prazo (informações quinquenais), cujo maior percentual de acertos no modelo de longo prazo sinalizou na direção de que as informações de longo prazo são mais úteis para explicar a relação entre informações contábeis e retornos de ações do que as de curto prazo.

Também não é intenção deste trabalho explicar os diferentes níveis de acertos por anos e/ou países da amostra ou a concentração de acertos em determinada categoria, dentre os três níveis de retorno dos modelos, uma vez que a comparação entre os percentuais totais de acertos de predições dos modelos de curto e de longo prazos foi suficiente para demonstrar que as

informações de longo prazo são mais úteis para explicar a relação entre informações contábeis e retornos de ações do que as de curto prazo. Entretanto, os resultados disponíveis nas tabelas comparativas ano a ano e segmentadas por países (APÊNDICE F) podem ser exploradas como inspiração para novas pesquisas que busquem explicar tais pontos e outros que o autor não explorou neste trabalho.

1.6. Estrutura do trabalho

Além desta primeira seção que realiza a introdução do estudo, o trabalho contempla:

- a) seção 2: que apresenta o referencial teórico utilizado para definição da questão da pesquisa que se alicerça basicamente na hipótese de eficiência do mercado e na relevância da informação contábil (*value relevance*), bem como no modelo logit multinomial que será a ferramenta de análise empírica;
- b) seção 3: que explicita o desenvolvimento da hipótese desta pesquisa que versa sobre a relação entre desempenhos contábil e de mercado de curto e de longo prazos;
- c) seção 4: que detalha a metodologia da pesquisa a partir da justificativa da aplicação do modelo logit multinomial, apresenta informações sobre a amostra utilizada e a fonte de tais dados, desenvolve a estrutura dos modelos de análises e define os procedimentos que serão cumpridos para efetivação da pesquisa;
- d) seção 5: que reporta os resultados desde a estatística descritiva, passando pelas análises dos modelos e das predições realizadas, pelo teste e procedimentos de robustez, até concluir com a síntese dos resultados em relação às hipóteses da pesquisa;
- e) seção 6: as considerações finais sobre a pesquisa.

2. REFERENCIAL TEÓRICO

2.1. Hipótese de eficiência do mercado

O trabalho seminal de Fama (1970) revisa de maneira abrangente a literatura teórica e empírica do modelo de mercados eficientes, explorando a eficiência em suas formas fraca, semi-forte e forte, passando pela teoria dos mercados eficientes, pelas evidências em pesquisas empíricas e concluindo que o apoio ao modelo de mercados eficientes é extenso, enquanto evidências contrárias são esparsas.

A teoria de mercados eficientes prevê que os preços das ações refletem o conhecimento dos investidores e sua capacidade em processar as informações disponíveis, sendo que o processo de formação dos preços é complexo e não totalmente compreendido (SCOTT, 2014, p.120), por exemplo, não é claro os mecanismos pelos quais as projeções de analistas afetam os preços de ativos e retornos esperados (KOTHARI, SO & VERDI, 2016).

Em um mercado eficiente, os preços das ações refletem todas as informações disponíveis e as variações de preços nesse mercado se comportarão aleatoriamente ao longo do tempo. Eficiência é definida em relação a um estoque de informações e se esse estoque de informações estiver incompleto ou suas informações erradas, os preços das ações também estarão errados. Assim, a eficiência do mercado não garante que os preços das ações reflitam o valor real da empresa e sim que os preços não apresentam viés em relação às informações publicamente disponíveis e reagem rapidamente a novas informações, até porque individualmente os investidores podem ter diferentes crenças e/ou podem interpretar as mesmas informações de maneira diferente e, conseqüentemente, apontar valores diferentes para os preços das ações (SCOTT, 2014, p.126).

Na hipótese de eficiência do mercado, a contabilidade concorre com outras fontes de informação, como mídia, analistas financeiros e até o próprio comportamento do preço de mercado das ações. Como veículo para informar os investidores, a contabilidade sobreviverá apenas se mantiver-se útil, oportuna e com custo compatível com outras fontes, sendo que a teoria de eficiência do mercado também alerta para a principal razão da existência da contabilidade que seria mitigar a assimetria da informação. Assim, podemos pensar na contabilidade como um mecanismo que permite a comunicação de informações úteis de dentro para fora da empresa, propiciando melhor qualidade nas decisões dos investidores. Os teóricos da contabilidade começaram a perceber a importância da eficiência do mercado no final dos anos 1960 e desde então a teoria tem guiado muita pesquisa, com grandes implicações para a prática contábil. Para os profissionais da contabilidade deve-se enfatizar que a hipótese de eficiência do mercado é um modelo de como o mercado de capitais opera e como qualquer modelo, ela não captura toda a complexidade de tal mercado, portanto, a questão relevante é o grau de eficiência – ou seja, quão próximos da plena eficiência estão os mercados, sendo que nos últimos anos muito se tem perguntado sobre se o investidor médio, cujo comportamento está subjacente à eficiência do mercado, é tão racional conforme o modelo assume, e cada vez mais evidências questionam a própria eficiência do mercado (SCOTT, 2014, p.121).

Nesse sentido Scott (2014) concluiu que, embora os mercados de capitais não sejam totalmente eficientes, eles são eficientes o suficiente para que os profissionais da contabilidade

possam ser guiados por implicações de eficiência e teoria da decisão racional subjacente a eles, sendo que a medida em que os mercados não são totalmente eficientes, isso aumenta ainda mais a importância dos relatórios financeiros.

2.2. *Value relevance* – relevância da informação contábil

Se a hipótese de eficiência do mercado é descrição razoável da realidade, deveríamos observar os preços das ações responderem de maneira previsível aos anúncios de novas informações. Quando tais preços respondem às informações contábeis dizemos que essas informações tem relevância (*value relevance*), sendo que a primeira evidência significativa dessa reação do mercado aos anúncios de informações é creditada pela literatura como sendo o estudo de Ball & Brown (1968) e desde então, um grande número de estudos empíricos documentaram aspectos adicionais de *value relevance* (SCOTT, 2014, p.153).

A informação contábil sendo relevante, é útil aos investidores, ajudando-os a estimar os valores e riscos esperados dos retornos das ações. Em essência, a informação é útil se leva os investidores a mudarem suas crenças e ações. Além disso, o grau de utilidade para investidores pode ser medido pela extensão da mudança de volume ou preço após o anúncio da informação (SCOTT, 2014, p.154).

A *value relevance* implica que as pesquisas empíricas podem ajudar os profissionais da contabilidade aumentarem ainda mais a utilidade da informação contábil ao permitir que as respostas do mercado os guiem sobre quais informações são ou não valorizadas pelos investidores e ainda propiciar uma melhora no funcionamento dos mercados em alocar capital se os preços das ações forem bons indicadores de oportunidades de investimento e desempenho futuro das companhias, contudo, os profissionais da contabilidade não necessariamente podem alegar que a melhor política contábil é a que produz a melhor resposta do mercado (SCOTT, 2014, p.154).

A relevância da informação contábil pode ser medida pela extensão da mudança no preço das ações motivada por sua divulgação e isso ocorre em função da racionalidade dos investidores que revisam suas expectativas sobre o desempenho futuro da empresa e com suas crenças revisadas tomam suas decisões de compra e venda. Se não houvesse relevância na informação contábil, não aconteceria revisão de crenças e, portanto, nenhuma alteração de preço associada. Os estudos empíricos, especialmente a partir de Ball & Brown (1968), demonstraram uma resposta diferencial do mercado, apoiando a hipótese de eficiência do mercado e a relevância da informação contábil (SCOTT, 2014, p.172).

Holthausen & Watts (2001) ao relatarem que numerosos estudos contábeis têm investigado os aspectos empíricos da relação entre valores de mercado de ações e números contábeis, preocupam-se em seu estudo com aqueles artigos de *value relevance* que tem por motivação a definição de padrões contábeis, sob o argumento de que sem teorias descritivas para interpretar as associações empíricas, estas teriam implicações limitadas para o estabelecimento de padrões contábeis. Holthausen & Watts (2001, p. 5-6) dividem os estudos de *value relevance* em três categorias: estudos de associação relativa que, por exemplo, examina se a associação de um lucro calculado por um determinado padrão é mais altamente associado com retorno de ações do que outro calculado sob outro padrão; estudos de associação incremental que investigam se um número contábil de interesse é útil na explicação de valores de mercado ou retornos, normalmente em combinação com variáveis especificadas no modelo; e estudos de conteúdo de informação marginal que investigam se um determinado número contábil agrega informação adicional ao conjunto de informações já disponível para os investidores.

Barth, Beaver & Landsman (2001) em seu trabalho publicado no mesmo volume do periódico rebatem Holthausen & Watts (2001), apresentando outra perspectiva ao explicar que a pesquisa de *value relevance* ao refletir as informações usadas pelos investidores também fornecem *insights* para o estabelecimento de padrões contábeis, destacando que apesar de o foco principal dos relatórios contábeis serem os investidores, outros usos desses relatórios, como condições de acompanhamento de contratos, por exemplo, não diminuem a importância da pesquisa em *value relevance*, destacando a importância para o avanço da contabilidade representada por pesquisa de *value relevance* que além da utilização de modelos de avaliação existentes, realizam estudos econométricos que em suas inferências exigem do pesquisador tempo e esforço para desenvolver desenhos de pesquisas e para aprender sobre várias questões de interesse relacionados aos relatórios contábeis.

2.2.1. *Value relevance* e diferenciação pelo mercado do desempenho de longo prazo

Seguindo as pesquisas sobre relevância da informação contábil para o valor da empresa, Rosa & Lustosa (2014) verificaram que o mercado diferenciava as empresas de alto, médio e baixo desempenho operacional de longo prazo, medido por informações contábeis de rentabilidade, variação de vendas e endividamento.

O estudo de Rosa & Lustosa (2014) foi realizado em amostra composta por 142 empresas brasileiras de diferentes setores listadas na Bolsa de Valores e de Mercadorias e

Futuros de São Paulo (BM&FBOVESPA), no período de 1996 a 2009, utilizando janelas móveis de cinco anos que resultaram em dez períodos quinquenais. Sem realizar previsões em amostra *out-of-sample*, os autores avaliaram se os coeficientes de respostas das variáveis contábeis das empresas eram estatisticamente significantes em sua associação com o retorno das respectivas ações a fim de verificar se o mercado diferenciava o desempenho operacional de longo prazo das empresas, medido por informações contábeis de diferentes naturezas, que sintetizavam o desempenho quinquenal em empresas de alto, médio e baixo desempenho.

Os resultados do trabalho de Rosa & Lustosa (2014) demonstraram que o mercado acionário brasileiro diferenciou no período analisado as empresas de alto e baixo desempenho operacional de longo prazo e que essa distinção não foi percebida por completo entre as empresas de alto e médio desempenho operacional.

Esses achados demonstraram que o conjunto das variáveis contábeis utilizadas naquele estudo para denotar o desempenho operacional de longo prazo das empresas era significativamente relevante para explicar o retorno das ações, corroborando e inspirando a realização deste trabalho que visa verificar se as informações de longo prazo seriam mais úteis para explicar a relação entre informações contábeis e retornos de ações do que as de curto prazo.

2.2.2. *Value relevance*, demonstrações financeiras e indicadores contábeis

Estudos que testam a *value relevance* das informações contidas em demonstrações financeiras e seus indicadores contábeis são comuns na literatura e serviram de suporte a elaboração desta pesquisa. Ebaid (2012) ao examinar e comparar a *value relevance* relativa e incremental de um conjunto de medidas de desempenho de empresas baseadas em informações contábeis, apresentou resultados indicando que a relevância relativa e incremental do valor tendem a aumentar ao descer na demonstração de resultados, com o lucro líquido tendo a maior relevância relativa e incremental, enquanto o total de vendas tem a menor.

Ebaid (2012) verificou em seu estudo que as medidas de desempenho baseadas no regime de competência teriam *value relevance* estatisticamente superiores às dos fluxos de caixa operacionais, em linha com o estudo de Francis, Schipper & Vincent (2003) que buscaram responder se indicadores de desempenho diferentes do lucro contábil estariam melhor refletidos nos retornos das ações, segmentando sua análise por tipo de indústrias e concluindo, como eles esperavam, que o lucro é superior a fluxo de caixa operacional e EBITDA (*earnings before interest, taxes, depreciation and amortization*) na explicação de retornos, apesar de apresentarem como limitação do trabalho que se a metodologia fosse por regressões não

lineares dos indicadores de desempenho, poderiam ser produzidas inferências diferentes. Similar ao estudo de Biddle, Bowen & Wallace (1997) que ao compararem associações do EVA (*Economic Value Added*) e do lucro contábil com retorno de ações afirmaram que a maior *value relevance* era a do lucro contábil para os participantes do mercado.

Utilizando-se de um modelo logit binomial, Ou & Penman (1989) investigaram inicialmente 68 descritores contábeis, que ao longo do trabalho foram reduzidos para 28¹, no sentido de analisar um grande conjunto de itens e reduzir tal conjunto para o modelo final de predição, tendo por referência o poder explicativo dessas variáveis. Em certa medida parecido com o desenho desta pesquisa, Ou & Penman (1989) também analisaram seu problema sob as visões de curto e longo prazos, ao diferenciar as posições tomadas em ações em curtas e longas que possuem estratégias de investimento diferentes. Também similar a esta pesquisa, Ou & Penman (1989) realizaram previsões em amostra *out-of-sample* e o período de estimação utilizado por eles foi de 1965-1972 (utilizando 16 descritores) para predição de 1973-1977 e 1973-1977 (utilizando 18 descritores) para predição do período 1978-1983².

As evidências encontradas por Ou & Penman (1989) sugerem que as demonstrações financeiras capturam fundamentos que não se refletem nos preços, assim, apontaram limitações no processo tradicional de abordagem na análise empírica da contabilidade de fazer inferências com base em associações contemporâneas entre números contábeis e preços de ações, ainda assim, concluíram que seus resultados indicam que as associações preditivas entre lucro e retorno de ações capturam boa parte da associação contemporânea entre lucro e retorno relatada por Ball & Brown (1968).

Holthausen & Larcker (1992), utilizando os 68 descritores contábeis investigados por Ou & Penman (1989), examinaram a lucratividade de uma estratégia de investimento baseada em um modelo de logit projetado para prever o sinal de retornos anormais em doze meses a partir dos índices contábeis e obtiveram retornos anormais entre 4,3% e 9,5%, dependendo da medida específica de excesso de retorno comparada, entretanto, ao implementarem a estratégia de negociação de Ou e Penman (1989), com os 28 índices contábeis utilizados no modelo final, afirmaram não terem obtido retornos excedentes.

As duas razões oferecidas por Holthausen & Larcker (1992) para seus resultados diferirem dos de Ou & Penman (1989) tem a ver com o lapso temporal:

¹ As quatro categorias que serão utilizadas nos MLM deste trabalho – alavancagem, liquidez, giro do ativo e lucratividade – estão entre os descritores contábeis utilizados por Ou & Penman (1989) em seu modelo final de predição.

² Dos 16 descritores utilizados na primeira série de predições, 6 se repetiam entre os 18 descritores da segunda série, por isso o total de 28 descritores diferentes.

- a) estratégia de negociação de Ou & Penman (1989) apresenta desempenho ruim no período subsequente a 1983, que não fez parte do estudo de Holthausen & Larcker (1992);
- b) os resultados da estratégia de negociação de Holthausen & Larcker (1992) não incorporaram os resultados de 1973-1977, período em que Ou & Penman (1989) relataram que a estratégia de negociação funcionava bem.

Face a esses resultados, Holthausen & Larcker (1992) afirmaram que a rentabilidade da estratégia de negociação de Ou & Penman (1989) parecia mais frágil do que o sugerido pelo estudo. De qualquer forma, Holthausen & Larcker (1992) concluíram ser surpreendente um modelo estatístico que não considerou quaisquer fundamentos econômicos, ter obtido retornos anormais da magnitude relatada considerando o ambiente competitivo da bolsa de Nova Iorque, de onde foram extraídos os dados da amostra. Por isso, Holthausen & Larcker (1992) apresentam quatro razões pelas quais a lucratividade por eles relatada pudesse estar exagerada:

- a) os resultados não incluíram custos de transação;
- b) o benchmarking para retorno normal poderia estar mal especificado;
- c) poderia ter existido mudanças temporárias de risco que o modelo não tenha ajustado adequadamente;
- d) os dados contábeis de algumas empresas poderiam não estar disponíveis no final do terceiro mês subsequente ao final do ano fiscal, particularmente as empresas com problemas financeiros.

Tal discussão corrobora com o escopo deste estudo que adota o pressuposto de que o alargamento da janela temporal demonstraria que no longo prazo os desempenhos contábil e de mercado se alinhariam, melhorando a qualidade preditiva dos modelos, pois, a explicação para a diferenciação entre os resultados alcançados por Holthausen & Larcker (1992) e Ou & Penman (1989) poderia estar na definição de janela temporal utilizada por cada um dos modelos que se estendido no tempo poderiam alinhar seus resultados graças à suavização, tanto na empresa quanto no mercado, dos impactos de fatores externos a essas empresas.

Em trabalhos como o seminal de Beaver (1966) o desempenho contábil, na forma de indicadores contábeis, foi utilizado como base para predição de insolvência, destacando-se que sua preocupação primária não eram os indicadores contábeis e sim os dados contábeis contidos nesses indicadores, sob a premissa de que os dados contábeis podem ser avaliados em termos de sua utilidade e que a utilidade poderia ser definida em termos de sua capacidade preditiva. Outro trabalho seminal neste campo (ALTMAN, 1968) também se utiliza do desempenho

contábil como base de predição, agregando a MDA (Análise Discriminante Múltipla) como melhoria à análise convencional.

Shumway (2001) inova em seu estudo pela combinação de desempenho contábil com desempenho de mercado e descobre que a combinação melhora a acurácia do modelo, demonstrando que alguns dados contábeis têm baixo poder preditivo e que dados de mercado fortemente relacionados não costumavam ser incluídos em modelos de predição, como os de Beaver (1966) e Altman (1968).

Nesse sentido, segundo He et al. (2005) ao analisar o trabalho de Shumway (2001) ressalta que os resultados daquele trabalho apoiam a premissa de que índices de mercado e contábeis não são informações que competem como melhores estimadores, pelo contrário, afirma que a combinação de ambos em um mesmo modelo auxilia na melhoria da capacidade preditiva.

2.3. Modelos logit multinomial (MLM)

MLM aplicam-se a situações em que as escolhas envolvem mais de duas alternativas onde o pesquisador deseja relacionar a escolha observada a um conjunto de variáveis explanatórias, de forma a explicar e predizer a probabilidade de que tal escolha ocorra (HILL, GRIFFITHS & JUDGE, 2003, p.437).

Segundo Ohlson (1995), o valor da empresa relaciona-se com as variáveis contábeis (patrimônio líquido e lucro líquido), de forma que as informações contábeis apresentariam *value relevance* e afetaria os preços das ações. Ao modificar o patrimônio líquido, o lucro líquido estabelece uma correlação entre eles, sendo que a construção do lucro líquido é resultado das diversas escolhas realizadas pela administração da empresa durante o exercício, tais como: decisões administrativas, de compras, de vendas, de financiamento, de investimento, além de escolhas sobre práticas e políticas contábeis que afetam a divulgação de suas informações e, conseqüentemente, da percepção e expectativas do investidor que por sua vez afetam os preços das ações.

Se o MLM se destina a explicar e predizer a probabilidade de uma escolha ocorrer, por meio da relação com variáveis explanatórias, espera-se que tal modelo represente uma poderosa ferramenta para verificar a relevância da informação contábil na medida em que um constructo de desempenho de mercado transforma a variável dependente em categorias de nível de retorno das ações das empresas, como se as diversas escolhas que a administração da empresa fosse realizando ao longo do período estivesse conduzindo a empresa à escolha do nível de retorno

de suas ações que o mercado reconhecerá ao alinhar as expectativas dos investidores com as informações que fluem da empresa para o mercado reportando o nível de assertividade das decisões e atividades das empresas na direção de uma maior lucratividade e liquidez.

Em estudos contábeis os modelos logísticos tem sido utilizados para predição, como os modelos binomiais de Ohlson (1980) e Shumway (2001) a partir de desempenho contábil ou o multinomial de Johnsen & Melicher (1994) que adicionou uma terceira categoria a fim de verificar o efeito sobre a qualidade de suas predições em comparação às pesquisas realizadas por Beaver (1966) e Altman et al. (1977) que utilizavam modelos de duas categorias. Os resultados verificados por Johnsen & Melicher (1994) demonstraram que a adição de uma terceira categoria melhorava o poder preditivo dos modelos de Beaver (1966) e Altman et al. (1977) que já eram considerados bons. Segundo Ward (1999, p.167), Lau (1987) foi o primeiro nesse campo a estender a metodologia binomial dos modelos de predição para um modelo multinomial, usando cinco categorias.

Garson (2014, p. 12) explica que a regressão logística, base do MLM, pode ser usada para prever uma variável dependente categórica a partir de variáveis independentes contínuas e/ou categóricas, mensurando o efeito das variáveis independentes na dependente e classificando a importância relativa das variáveis independentes. Realizar tais predições, mensurações e classificações foi fundamental para cumprir os objetivos do presente estudo apontando o nível de relevância da informação contábil para a predição do retorno de ações, como medida de desempenho de mercado das empresas.

Este estudo adotou o MLM para analisar a relevância da informação contábil e para tanto o modelo tem por objeto relacionar o retorno das ações com variáveis contábeis, macroeconômicas e de mercado, no entanto, ao invés de uma relação direta entre retorno e tais variáveis, o modelo prevê a criação de 3 categorias que classificam tais retornos em baixo, médio e alto. Outras informações do MLM serão exploradas nas seções 4 e 5 deste trabalho em que, respectivamente, serão discutidos a metodologia da pesquisa e os resultados obtidos.

2.3.1. Detalhando o MLM

Os MLM estimam pelo método de máxima verossimilhança com variáveis dependentes discretas e tais variáveis podem assumir mais de dois valores (STATACORP, 2017, p.1658).

De acordo com Greene (2018, *apud* STATACORP, 2017), Hosmer, Lemeshow & Sturdivant (2013, *apud* STATACORP, 2017), Long (1997, *apud* STATACORP, 2017), Long & Freese (2014, *apud* STATACORP, 2017), Treiman (2009, *apud* STATACORP, 2017) e

Davidson & MacKinnon (1993, *apud* STATACORP, 2017), o MLM estima uma série de coeficientes $\beta^{(1)}$, $\beta^{(2)}$ e $\beta^{(3)}$ correspondentes aos resultados:

$$\Pr(y = 1) = \frac{e^{X\beta^{(1)}}}{e^{X\beta^{(1)}} + e^{X\beta^{(2)}} + e^{X\beta^{(3)}}}$$

$$\Pr(y = 2) = \frac{e^{X\beta^{(2)}}}{e^{X\beta^{(1)}} + e^{X\beta^{(2)}} + e^{X\beta^{(3)}}}$$

$$\Pr(y = 3) = \frac{e^{X\beta^{(3)}}}{e^{X\beta^{(1)}} + e^{X\beta^{(2)}} + e^{X\beta^{(3)}}}$$

O modelo é indefinido no sentido de que existem mais de uma solução para os coeficientes $\beta^{(1)}$, $\beta^{(2)}$ e $\beta^{(3)}$ que levam às mesmas probabilidades para as categorias $y = 1$, $y = 2$ e $y = 3$. Para definir o modelo é escolhido arbitrariamente como zero um dos coeficientes $\beta^{(1)}$, $\beta^{(2)}$ ou $\beta^{(3)}$. Ou seja, definido um dos coeficientes como zero, os restantes medirão a variação relativa à respectiva categoria. Os coeficientes diferem porque têm interpretações diferentes, mas as probabilidades previstas para $y=1$, $y=2$ e $y=3$ ainda serão as mesmas. Portanto, qualquer uma das parametrizações será uma solução para o mesmo modelo³ (STATACORP, 2017, p.1659).

Definido $\beta^{(1)} = 0$, a equação torna-se:

$$\Pr(y = 1) = \frac{1}{1 + e^{X\beta^{(2)}} + e^{X\beta^{(3)}}}$$

$$\Pr(y = 2) = \frac{e^{X\beta^{(2)}}}{1 + e^{X\beta^{(2)}} + e^{X\beta^{(3)}}}$$

$$\Pr(y = 3) = \frac{e^{X\beta^{(3)}}}{1 + e^{X\beta^{(2)}} + e^{X\beta^{(3)}}}$$

E a probabilidade relativa de $y = 2$, por exemplo, seria:

³ Na metodologia deste trabalho optou-se por definir como base de cálculo do MLM a categoria “baixo retorno” e foram calculadas as probabilidades de cada uma das três categorias a fim de definir a categoria predita pelo modelo como aquela cuja probabilidade era dominante em relação às outras duas categorias.

$$\frac{\Pr(y = 2)}{\Pr(y = 1)} = e^{X\beta^{(2)}}$$

Denominando essa equação de razão de risco relativo (RRR – *relative-risk ratios*) e assumindo que X e $\beta_k^{(2)}$ são vetores iguais a (x_1, x_2, \dots, x_i) e $(\beta_1^{(2)}, \beta_2^{(2)}, \dots, \beta_k^{(2)})$, respectivamente. A RRR para a alteração de uma unidade em x_i é então:

$$\frac{e^{\beta_1^{(2)}x_1 + \dots + \beta_i^{(2)}(x_i+1) + \dots + \beta_k^{(2)}x_k}}{e^{\beta_1^{(2)}x_1 + \dots + \beta_i^{(2)}x_i + \dots + \beta_k^{(2)}x_k}} = e^{\beta_i^{(2)}}$$

Assim, a exponencial de um coeficiente é a RRR para a variação de uma unidade na variável correspondente e o risco é mensurado como o risco de uma categoria em relação a categoria base (STATACORP, 2017, p.1660).

3. DESENVOLVIMENTO DA HIPÓTESE DA PESQUISA

De acordo com Lev (1989), a correlação entre resultado contábil e retorno de ações é muito baixa e, além do mais, a relação entre essas variáveis é instável ao longo do tempo, sugerindo que informações trimestrais e anuais teriam utilidade limitadas para tomada de decisão de investidores.

Sob a premissa de que os demais indicadores contábeis da empresa são correlacionados com o resultado contábil, esperava-se que o constructo de desempenho contábil com informações anuais tivesse baixa correlação com o constructo de desempenho de mercado, derivado de retornos anuais de ações. Assim, o modelo de predição que relacionasse desempenhos contábil e de mercado de curto prazo (informações anuais), apresentaria pouco ou nenhum poder preditivo.

Por outro lado, os resultados do estudo de Rosa & Lustosa (2014) sinalizaram fortes evidências de que o mercado efetivamente diferencia as empresas de alto desempenho operacional de longo prazo, com o conjunto das variáveis contábeis utilizadas denotando o desempenho operacional das empresas significativamente relevante para explicar o retorno das ações, corroborando com o pressuposto desta pesquisa de que ao alargar a janela temporal, as tendências do desempenho contábil melhor se alinhariam às do desempenho de mercado.

Esperava-se que com informações de longo prazo (5 anos) fosse verificada uma maior relação entre os desempenhos contábil e de mercado, pois no longo prazo a tendência da gestão

da empresa e sua interação com o mercado seria destacada e ao mesmo tempo suavizados tanto na empresa quanto no mercado os impactos de fatores externos a essas empresas, tais como: crises ou bolhas econômicas, movimentos pontuais de crescimento ou de recessão da economia nacional ou mundial etc. Nesse cenário, um modelo baseado em informações quinquenais apresentaria maior capacidade preditiva do que o de informações anuais, graças ao melhor ajuste entre as tendências de desempenhos contábil e de mercado no longo prazo, dando sustentação à formulação da seguinte hipótese da pesquisa a ser testada empiricamente:

H1: Os desempenhos contábil e de mercado de longo prazo são mais relacionados do que os desempenhos contábil e de mercado de curto prazo.

4. METODOLOGIA DA PESQUISA

4.1. Justificativa da aplicação do MLM

Os estudos relacionados a *value relevance* como Dechow, Hutton & Sloan (1999), Core, Guay & Verdi (2008) costumam utilizar-se de modelos de regressões lineares como base metodológica para estabelecer as relações entre as variáveis, sendo que Kothari (2001) ao fazer uma revisão de pesquisas empíricas em Contabilidade sobre mercado de capitais não cita qualquer trabalho nessa área que tenha utilizado o MLM como base metodológica, apesar de citar alguns que se utilizaram de metodologias não lineares.

Ou & Penman (1989) e Holthausen & Larcker (1992), desenvolveram estudos de *value relevance* com modelos logit binomiais em que testaram estratégias de investimento em ações a partir de informações contábeis das empresas.

Por outro lado, como evolução das pesquisas de predição de insolvência, a partir de Ohlson (1980) com seu modelo binário de logit que já apresentava vantagens sobre os modelos clássicos, a aplicação do MLM nesse campo tem sido comum e tem apresentado boa qualidade preditiva ao estabelecer relação entre informações contábeis e saúde financeira das empresas, como os trabalhos de Lau (1987), Johnsen & Melicher (1994), Pinder (1996), Mselmi, Lahiani & Hamza (2017).

Além das vantagens do logit sobre os modelos clássicos indicadas por Ohlson (1980), o modelo logit se equipara aos métodos de análise mais modernos, como demonstraram Mselmi, Lahiani & Hamza (2017) ao compará-lo com outros quatro modelos de metodologias mais modernas, como rede neural e inteligência artificial, concluindo que todos os cinco métodos

utilizados atingiam as expectativas de poder de predição, superando em eficiência estudos anteriores com modelos lineares clássicos e sem conseguir provar a superioridade dos quatro métodos mais modernos sobre o modelo logit.

Daí a ideia de se utilizar tal metodologia para aplicação num estudo de *value relevance*, como será detalhado nas próximas seções deste trabalho.

4.2. Amostra e fonte de dados

Buscou-se neste trabalho conferir robustez à análise por meio da coleta de informações que redundasse em uma amostra que trouxesse relevância global ao trabalho, assim, a base de dados apresenta informações contábeis e preços de ações de 50.998 empresas pertencentes a 18 países diferentes, distribuídos por todos os continentes e englobando os cinco países do BRICS.

As informações contábeis das empresas e os preços de suas ações são anuais e foram obtidas pelo acesso à base Thomson Reuters com a amostra apresentando 500.479 observações no período de 23 anos (1995 a 2017), permitindo gerar no MLM de curto prazo 296.625 observações válidas e no MLM de longo prazo 153.929. Já as informações de índices de mercado são oriundas das bolsas de valores dos 18 países, conforme detalhado na seção 4.1.1.

Foram excluídas as empresas do sistema financeiro dada a peculiaridade da atividade e o período de séries históricas da amostra compreendeu o intervalo de 1995 a 2017. São consideradas em cada ano de análise as empresas com dados disponíveis naquele ano, assim as informações da base não receberam qualquer tratamento quanto ao viés de sobrevivência.

As 500.479 observações válidas foram organizadas em dados agrupados em painel (*pooled*) por empresa e ano, tipificando uma amostra de volume robusto de informações que se espera oferecer qualidade aos cálculos e ajustes aos modelos.

4.2.1. Extração de dados

As informações contábeis das empresas foram obtidas por meio de acesso à base Thomson Reuters, *Worldscope - Fundamentals Annual* (WFA) por meio do WRDS (2018).

Também da base Thomson Reuters, foram extraídas as informações de mercado (preço das ações) das empresas, obtidas por meio de acesso à base *Worldscope - Stock Data* (WSD) por meio do WRDS (2018).

Os dados macroeconômicos foram obtidos por meio de consulta ao site do Banco Mundial (2019), base de dados da qual foram extraídos os seguintes dados macroeconômicos

dos 18 países que compõem a amostra base deste trabalho: taxa de crescimento anual do PIB ([GDP growth (annual %)]; inflação de preços ao consumidor – taxa de variação dos preços na economia como um todo [Inflation, consumer prices (annual %)]; e taxa de juros para empréstimos – taxa bancária que geralmente atende às necessidades de financiamento de curto e médio prazo do setor privado, normalmente diferenciada de acordo com a capacidade creditícia dos tomadores e os objetivos do financiamento [Lending interest rate (%)].

Os índices de mercado dos 18 países presentes neste estudo foram obtidos por meio de consulta ao site Investing.com (2019), sendo utilizados os índices de cada país, conforme Quadro 1 a seguir.

Quadro 1 – Índices de mercado por país

País	Índices de mercado
África do Sul	FTSE/JSE Africa All Share Index
Alemanha	DAX-30
Austrália	All Ordinaries
Brasil	Ibovespa
Canadá	S&P/TSX Composite Index
China	SSE Composite Index
Cingapura	Straits Times Index (STI)
Coréia do Sul	KOSPI
EUA	S&P 500
França	SBF 120
Grécia	Athex 20
Índia	National Stock Exchange of India (NSE)
Japão	Nikkei 225
Malásia	FTSE Bursa Malaysia Index
Polônia	WIG
Reino Unido	FTSE All-Share Index
Rússia	MICEX (Moscow Interbank Currency Exchange)
Suécia	OMX Stockholm PI (OMXSPI)

Fonte: elaboração própria.

4.2.2. Variáveis utilizadas

A partir dos dados extraídos foram compostas as variáveis primárias que serão base deste estudo, bem como aquelas criadas para execução dos modelos anual e quinquenal, apresentadas no APÊNDICE C, Quadro 6.

Das informações de mercado das empresas, extraídas da base WSD, é utilizada neste trabalho a variável que representa preço de mercado da ação, no final de cada ano base, a fim

de calcular os retornos normais anuais e quinquenais de cada empresa (item5001 – *Market Price - Year End*)⁴ e que é referenciada neste estudo pelo termo “preço da ação”.

4.2.2.1. Informações contábeis

Das informações contábeis das empresas, extraídas da base WFA, serão utilizadas nos modelos anual (curto prazo) e quinquenal (longo prazo) as variáveis detalhadas ao final desta subseção. Além dessas variáveis, encontram-se detalhadas no Quadro 2 (APÊNDICE A) todas as variáveis utilizadas neste trabalho, de acordo com a ordem em que aparecem no *Do-File*⁵ (APÊNDICE B), sendo que os indicadores contábeis foram selecionados da base Thomson Reuters por estarem disponíveis nos formatos anual e quinquenal. Para facilidade de verificação na base Thomson Reuters da qual foram extraídas, foi mantido no Quadro 2 (APÊNDICE A) como símbolo da variável a nomenclatura utilizada pela Thomson Reuters disponível no Guia de Definições (THOMSON FINANCIAL, 2007).

Tendo por referência os indicadores contábeis utilizados nos trabalhos de Ou & Penman (1989) e Holthausen & Larcker (1992), em estudos de *value relevance* que se utilizaram de modelos logit binomiais, bem como os trabalhos seminais em predição de insolvência de Beaver (1966) e Altman et al. (1977), o constructo de desempenho contábil para os MLM deste trabalho será composto por quatro categorias de indicadores: alavancagem, liquidez, lucratividade e giro do ativo.

Para representar cada categoria, foi selecionado no âmbito dos dados da amostra, o indicador daquela categoria de maior correlação com o retorno das ações das empresas que compõem a amostra.

Tais indicadores estavam disponíveis na base Thomson Reuters nos formatos anuais e quinquenais e comporão o constructo de desempenho contábil para os modelos anual e quinquenal, conforme Quadro 3 a seguir.

⁴ Nomenclatura utilizada pela Thomson Reuters disponível no Guia de Definições (THOMSON FINANCIAL, 2007).

⁵ Rotina de execução do trabalho utilizada pelo aplicativo STATA.

Quadro 3 – MLM de curto e longo prazos: variáveis que representam o constructo de desempenho contábil

Símbolo	Variável	Descrição
alav_1ano ⁶	Indicador de alavancagem	$alav_1ano = \frac{\text{Patrimônio Líquido}}{\text{Ativo Total}} * 100$
alav_5anos		média de 5 anos do indicador anual de alavancagem
liq_1ano ⁷	Indicador de liquidez	$liq_1ano = \frac{\text{Caixa e Equivalentes}}{\text{Ativo Circulante}} * 100$
liq_5anos		média de 5 anos do indicador anual de liquidez
lucr_1ano ⁸	Indicador de lucratividade	$lucr_1ano = \frac{\{Luc.Liq.antes de Div.pref.+ (Jr.pag-Jr.Cap.)*(1-Alíq.IR)\}}{\text{Méd.ano ant.e ano corr.da soma de Cap.total,Dív.de cp e parcela de cp das Dív.de lp do ano ant.}} * 100$
lucr_5anos		média de 5 anos do indicador anual de lucratividade
ga_1ano ⁹	Indicador de giro do ativo	$ga_1ano = \frac{\text{Vendas Líquidas}}{\text{Ativo Total}} * 100$
ga_5anos		média de 5 anos do indicador anual de giro do ativo

Fonte: elaboração própria.

4.2.2.2. Dados macroeconômicos e índices de mercado

Dos dados obtidos por meio de consulta ao Banco Mundial (2019) e ao Investing.com (2019) são utilizadas neste trabalho as variáveis detalhadas no Quadro 4, a seguir, para o MLM de curto prazo.

⁶ Leverage Ratio: $\text{Common Equity} / \text{Total Assets} * 100$

⁷ Liquidity Ratio: $(\text{Cash \& Equivalentes} / \text{Current Liabilities-Total}) * 100$

⁸ Profitability Ratio: $(\text{Net Income before Preferred Dividends} + ((\text{Interest Expense on Debt} - \text{Interest Capitalized}) * (1-\text{Tax Rate}))) / \text{Average of Last Year's and Current Year's (Total Capital} + \text{Last Year's Short Term Debt} \& \text{Current Portion of Long Term Debt}) * 100$

⁹ Asset Utilization Ratio: $\text{Net Sales or Revenues} / \text{Total Assets} * 100$

Quadro 4 – MLM de curto prazo: variáveis que representam os dados macroeconômicos e índices de mercado por país

Símbolo	Variável	Descrição
PIB	GDP growth (annual %)	Taxa anual de crescimento percentual do PIB de cada país.
Inflação	Inflation, consumer prices (annual %)	Taxa percentual anual de inflação do respectivo país.
Juros	Lending interest rate (%)	Taxa anual percentual média de juros bancária que geralmente atende às necessidades de financiamento de curto e médio prazo do setor privado, diferenciando-se de acordo com a capacidade creditícia dos tomadores e os objetivos do financiamento.
Ind_mercado	índice de mercado de cada país analisado	Lista dos índices e seus respectivos países mencionados no Quadro 1 deste trabalho.

Fonte: elaboração própria.

Estão indicadas no Quadro 5 as variáveis utilizadas no MLM de longo prazo, médias aritméticas dos cinco anos anteriores calculadas a partir de cada uma das variáveis macroeconômicas e do índice de mercado utilizados no MLM de curto prazo.

Quadro 5 – MLM de longo prazo: variáveis que representam os dados macroeconômicos e índices de mercado por país

Símbolo	Variável	Descrição
PIB5a	GDP growth (annual %) - 5 Yr Avg	Média aritmética dos cinco anos anteriores do PIB de cada país.
Inflação5a	Inflation, consumer prices (annual %) - 5 Yr Avg	Média aritmética dos cinco anos anteriores da inflação do respectivo país.
Juros5a	Lending interest rate (%) - 5 Yr Avg	Média aritmética dos cinco anos anteriores da taxa de juros em percentual anual do respectivo país.
Ind_mercado5a	ind_mercado - 5 Yr Avg	Média aritmética dos cinco anos anteriores do índice de mercado do respectivo país.

Fonte: elaboração própria.

4.3. Estruturação dos modelos de análises

A fim de comparar o efeito do alargamento da janela temporal sobre a qualidade das predições, foram criados dois MLM, um deles com variáveis anuais (curto prazo) e outro com

variáveis quinquenais (longo prazo), cujas variáveis dependentes e independentes serão detalhadas nas subseções a seguir. O Quadro 7 apresenta um resumo das variáveis dos MLM.

Quadro 7 – Resumo das variáveis dos MLM de curto e de longo prazos

MLM	Variável dependente	Variáveis independentes
Curto prazo (informações anuais)	Variável categórica que em termos de retorno de um ano pode assumir os seguintes valores: 1 – baixo; 2 – médio; e 3 – alto.	Valores anuais de: Indicador de alavancagem; Indicador de liquidez; Indicador de lucratividade; Indicador de giro do ativo; PIB; Inflação; Taxa de juros; e Índice de mercado.
Longo prazo (informações quinquenais)	Variável categórica que em termos de retorno de cinco anos pode assumir os seguintes valores: 1 – baixo; 2 – médio; e 3 – alto.	Médias aritméticas de cinco anos dos valores anuais de: Indicador de alavancagem; Indicador de liquidez; Indicador de lucratividade; Indicador de giro do ativo; PIB; Inflação; Taxa de juros; e Índice de mercado.

Fonte: elaboração própria.

Formulada na seção 3 deste trabalho, a hipótese **H₁** é testada por meio da comparação entres os dois MLM, de curto e longo prazos, que utilizarão todas as informações disponíveis na amostra para as empresas dos 18 países.

4.3.1. Modelos para relação entre desempenhos contábil e de mercado de curto e longo prazos

Quanto à metodologia do modelo de análise, este estudo parte de um constructo de desempenho de mercado que ordena o retorno das ações das empresas que compõem a amostra, dividindo-os em tercis de forma a estabelecer categorias (baixo, médio e alto retornos), constituindo-se na variável dependente dos MLM de curto e longo prazos. Já as variáveis independentes dos modelos são representadas por indicadores contábeis que se configuram no constructo de desempenho contábil das empresas, acrescidas de variáveis macroeconômicas e de índices de mercado.

4.3.2. Variáveis dependentes

São variáveis categóricas que podem assumir valores 1, 2 e 3, representando, respectivamente, um dos seguintes tercis em retornos anual ou quinquenal: 1 – baixo, 2 – médio, e 3 – alto. São variáveis representadas por dados em painel (matriz empresa X ano) em que as empresas podem pertencer a 18 países diferentes.

Para o MLM de curto prazo, essa variável (**tret1a_ano_a_ano**¹⁰) é a categoria indicada pelo tercil dos retornos normais anuais, com os anos variando de 1996 a 2017, sendo que para efeito do modelo de estimação foi estabelecida uma janela temporal de 1996 a 2012 para se prever as categorias de 2013 a 2017.

Já para o MLM de longo prazo, essa variável (**tret5a_ano_a_ano**¹¹) é a categoria indicada pelo tercil dos retornos normais quinquenais, sendo que os anos variam de 2000 a 2017 e para efeito do modelo de estimação foi estabelecida uma janela temporal de 2000 a 2012 para se prever as categorias de 2013 a 2017.

4.3.3. Variáveis independentes

Realizados testes de correlação entre os retornos normais anual e quinquenal calculados para cada empresa/ano da amostra (variáveis **ret1a** e **ret5a**) e as variáveis de informações contábeis detalhadas no Quadro 2 no APÊNDICE A deste trabalho, foi selecionada a de maior correlação combinada entre retorno normal anual e quinquenal dentro de cada uma das categorias de indicadores contábeis previamente definidas como componentes do constructo de desempenho contábil deste trabalho para os MLM de curto e longo prazos:

- a) Alavancagem: variáveis **alav_1ano** e **alav_5anos**;
- b) Liquidez: variáveis **liq_1ano** e **liq_5anos**;
- c) Lucratividade: variáveis **lucr_1ano** e **lucr_5anos**;
- d) Giro do ativo: variáveis **ga_1ano** e **ga_5anos**.

Às variáveis de informações contábeis foram acrescentadas as variáveis com dados macroeconômicos e índices de mercado de cada um dos 18 países:

¹⁰ Criada a partir das informações acumuladas ano a ano das variáveis **tret1aYYYY**, variável detalhada no Quadro 6 no APÊNDICE C.

¹¹ Criada a partir das informações acumuladas ano a ano das variáveis **tret5aYYYY**, variável detalhada no Quadro 6 no APÊNDICE C.

- a) PIB: variáveis **PIB** e **PIB5a**;
- b) Inflação: variáveis **inflação** e **inflação5a**;
- c) Taxa de juros: variáveis **juros** e **juros5a**;
- d) Índice de mercado: variáveis **ind_mercado** e **ind_mercado5a**.

Os indicadores contábeis já estavam disponíveis na base de dados em suas formas anuais e quinquenais (médias aritméticas de 5 anos), enquanto os dados macroeconômicos e índices de mercado dos 18 países estavam disponíveis somente em sua forma anual e foram calculados na forma quinquenal por meio de média aritmética.

4.3.4. Modelo Econométrico

$$P_n(i) = \frac{e^{\beta_k X_{jnk}}}{\sum_{j \in C} e^{\beta_k X_{jnk}}}$$

Em que $P_n(i)$ é a probabilidade de a ação da empresa n pertencer a categoria i dentro de um conjunto de possibilidades C representado pela matriz de dados que combina na amostra as empresas e a série temporal; e é a base do logaritmo neperiano; β_k são os parâmetros a serem estimados no modelo em que i varia k varia vezes ($i = 1, 2, \dots, k$) e X_{jnk} são atributos que caracterizam a função.

No MLM desta tese há três categorias a serem testadas na variável dependente ($i = 1, 2$ ou 3 , que correspondem respectivamente a baixo retorno, médio retorno e alto retorno); os coeficientes a serem estimados para cada conjunto de categorias são denominados: β_1 , β_2 e β_3 ; são utilizadas oito variáveis independentes (alavancagem, giro do ativo, liquidez, lucratividade, PIB, inflação, juros e índice de mercado) para testar o modelo e que correspondem ao índice j que varia de 1 a 8.

As probabilidades de ocorrência das 3 categorias são escritas da seguinte forma:

$$P_n(1) = \frac{e^{\beta_1 X_{j1}}}{e^{\beta_1 X_{j1}} + e^{\beta_2 X_{j2}} + e^{\beta_3 X_{j3}}}$$

$$P_n(2) = \frac{e^{\beta_2 X_{j2}}}{e^{\beta_1 X_{j1}} + e^{\beta_2 X_{j2}} + e^{\beta_3 X_{j3}}}$$

$$P_n(3) = \frac{e^{\beta_3 X_{j3}}}{e^{\beta_1 X_{j1}} + e^{\beta_2 X_{j2}} + e^{\beta_3 X_{j3}}}$$

4.4. Procedimentos

A partir da amostra com 500.479 observações no período de 23 anos (1995 a 2017), envolvendo 50.998 empresas de 18 países, com informações anuais de seus indicadores contábeis e preços de suas ações, foram calculados os retornos normais de mercado anual e quinquenal (variáveis **ret1a** e **ret5a**) de cada empresa/ano que dispusesse de informação na base de dados.

O cálculo do retorno normal anual (variável **ret1a**) é realizado por meio da fórmula:

$$\text{Retorno normal anual}_{ano\ t} = \frac{\text{Preço da ação}_{ano\ t} - \text{Preço da ação}_{ano\ t-1}}{\text{Preço da ação}_{ano\ t-1}}$$

E o cálculo do retorno normal quinquenal (variável **ret5a**) é realizado por meio da fórmula:

$$\text{Retorno normal quinquenal}_t = \frac{\text{Preço da ação}_{ano\ t} - \text{Preço da ação}_{ano\ t-5}}{\text{Preço da ação}_{ano\ t-5}}$$

Em seguida, foi classificado em tercís os retornos anuais em cada ano base, dos anos 1996 a 2017 (22 períodos) a fim de gerar as informações para a variável categórica que será a dependente do MLM de curto prazo. Também se classificou em tercís os retornos quinquenais dos anos 2000 a 2017 (18 períodos) para gerar as informações da variável dependente do MLM de longo prazo. A variável categórica dos modelos anual e quinquenal pode assumir valor 1, 2 ou 3, representando respectivamente os tercís de baixo retorno, médio retorno e alto retorno.

Com base nas informações anuais das variáveis macroeconômicas e índices de mercado utilizadas no MLM de curto prazo, foram calculadas as variáveis quinquenais a serem utilizadas no MLM de longo prazo por meio de média aritmética de 5 anos:

$$\text{Média aritmética de 5 anos} = \frac{X_t + X_{t-1} + X_{t-2} + X_{t-3} + X_{t-4}}{5}$$

Tomando por referência os indicadores contábeis utilizados nos trabalhos de Ou & Penman (1989) e Holthausen & Larcker (1992), em estudos de *value relevance* que se utilizaram de modelos logit binomiais, bem como a lista de indicadores contábeis utilizados nos trabalhos de Beaver (1966) e Altman et al. (1977) foram escolhidas quatro categorias de indicadores contábeis para compor o constructo de desempenho contábil dos MLM: alavancagem, liquidez, lucratividade e giro do ativo. Dentre os indicadores contábeis disponíveis na base Thomson Reuters nos formatos anuais e quinquenais, foi selecionado o indicador de maior correlação com retorno dentro de cada uma dessas categorias. Tais indicadores foram detalhados na seção 4.2.2.1. (Informações contábeis).

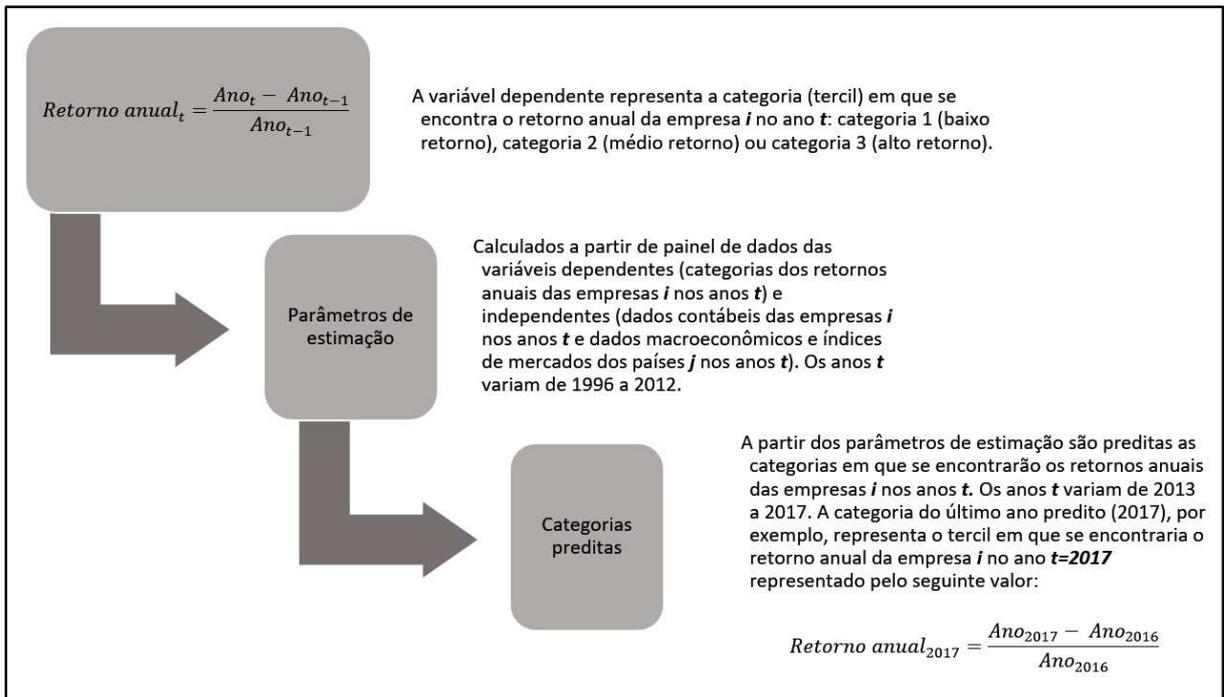
A seguir, foi executado o MLM de curto prazo no STATA (função *mlogit*¹²), utilizando o 1º tercil como base (baixo retorno) e considerando como período de estimação o intervalo entre 1996 e 2012 (17 anos).

Utilizando-se os parâmetros calculados pelo modelo, por meio da função *predict*, foi realizada a predição das probabilidades de as ações das empresas estarem nas categorias 1 (baixo retorno), 2 (médio retorno) e 3 (alto retorno) para o período de 2013 a 2017 (5 anos). A partir da comparação das probabilidades preditas, são criadas variáveis de categorias preditas para os anos de 2013, 2014, 2015, 2016 e 2017 para cada empresa que teve dados calculados pelo MLM com as variáveis assumindo o valor da categoria com maior probabilidade de ocorrer (probabilidade dominante). Então, são comparados os resultados previstos e realizados e relatado o nível de acertos do MLM de curto prazo.

A Figura 1 a seguir esquematiza os procedimentos no MLM de curto prazo desde o cálculo do retorno normal anual que será a base para a determinação dos tercís, ano a ano, que definirá o valor da variável categórica, passando pelo processo de estimação dos parâmetros e até a predição das categorias que serão comparadas com as categorias observadas a fim de apresentar o percentual de predições corretas. A fórmula do *Retorno anual*₂₀₁₇ apresentada na Figura 1 é meramente ilustrativa do racional da previsão, pois, o modelo não calcula o valor do retorno futuro da ação, ele assume como previsão o valor da categoria com maior probabilidade de ocorrer (probabilidade dominante) dentre as probabilidades de ocorrências preditas para as 3 categorias, estimadas de acordo com os parâmetros calculados no período 1996-2012.

¹² Apesar de a variável categórica ser calculada em tercís de baixo, médio e alto retorno, optou-se por utilizar a função *mlogit* (logit multinomial), em detrimento da função *ologit* (logit multinomial ordenado) porque as empresas que compõem a amostra mudam de posições de ano base a ano base, deslocando-se entre as categorias que compõem o painel de dados. Ainda assim o autor testou a utilização do logit multinomial ordenado e o resultado das predições foram similares, considerando o desenho metodológico adotado nesta pesquisa.

Figura 1 – Procedimentos esquematizados do MLM de curto prazo



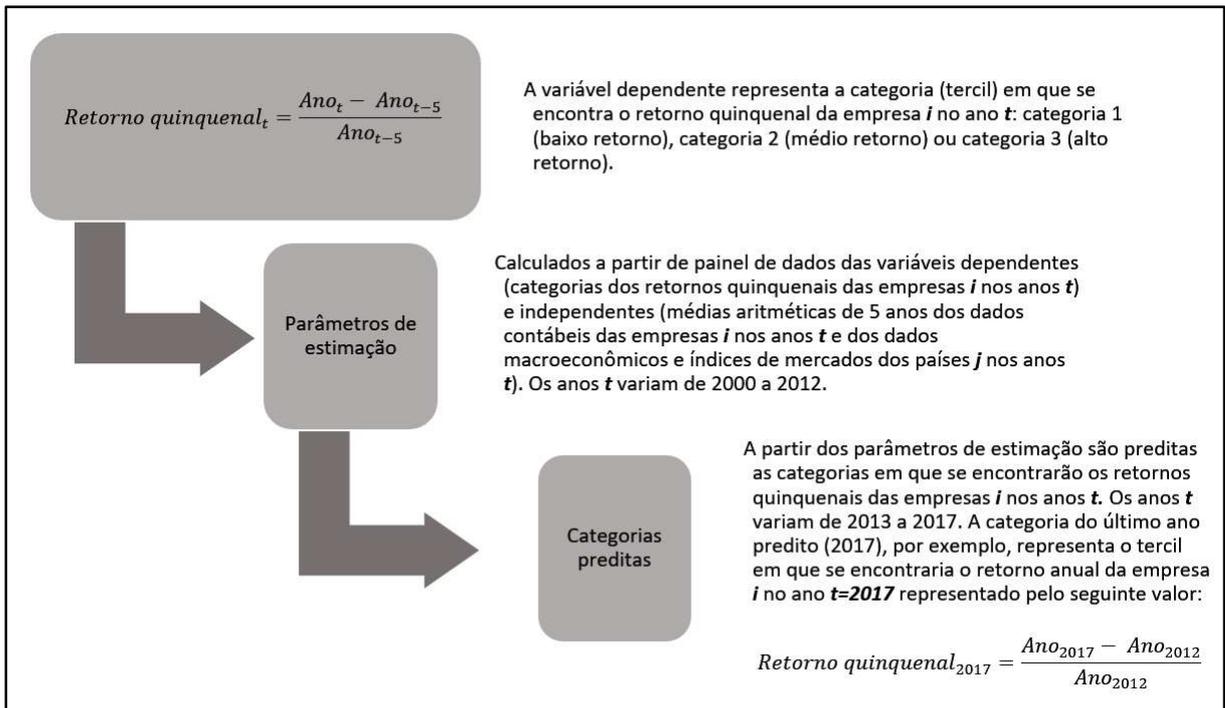
Fonte: elaboração própria.

Na execução do MLM de longo prazo no STATA (função `mlogit`) também foi utilizado o 1º tercil como base (baixo retorno), sendo considerado como período de estimação o intervalo entre 2000 e 2012 (13 anos).

A exemplo do MLM de curto prazo, com base nos parâmetros do MLM de longo prazo são criadas variáveis de categorias preditas para os anos de 2013, 2014, 2015, 2016 e 2017, comparados os resultados previstos e realizados e relatado o nível de acertos do MLM de longo prazo.

A Figura 2 a seguir esquematiza os procedimentos no MLM de longo prazo desde o cálculo do retorno normal quinquenal que será a base para a determinação dos tercís, ano a ano, que definirá o valor da variável categórica, passando pelo processo de estimação dos parâmetros e até a predição das categorias que serão comparadas com as categorias observadas a fim de apresentar o percentual de predições corretas. A fórmula do **Retorno quinquenal**₂₀₁₇ apresentada na Figura 2, apenas ilustra a lógica da previsão, pois, como afirmado anteriormente, o modelo não prediz o valor do retorno futuro da ação, ele assume como previsão o valor da categoria com probabilidade dominante de ocorrer dentre as probabilidades de ocorrências preditas para as 3 categorias, estimadas com os parâmetros calculados no período 2000-2012.

Figura 2 – Procedimentos esquematizados do MLM de longo prazo



Fonte: elaboração própria.

A partir da comparação entre os resultados dos acertos de predição dos MLM de curto e longo prazos, testou-se a hipótese H_1 formulada na seção 3 deste trabalho que apresentou a hipótese desta pesquisa, cujos resultados e considerações serão realizados nas seções seguintes.

Repetindo-se os procedimentos numa amostra que continha informações de empresas sediadas em um único país, foi testada a robustez dos MLM de curto e longo prazos por meio da comparação entre os resultados obtidos com a amostra desse país e as predições realizadas com a amostra dos 18 países, cujos procedimentos foram detalhados nesta seção.

5. RESULTADOS

A partir dos dados da amostra, de toda a metodologia da pesquisa apresentada na seção 4 e em especial os procedimentos detalhados na seção 4.4, são apresentados nesta seção os resultados alcançados e suas análises dentro do escopo da hipótese da pesquisa apresentada na seção 3 e dos objetivos deste trabalho elencados na seção 1.3.

5.1. Estatística descritiva

A Tabela 1 apresenta o descritivo da amostra, incluindo a distribuição das observações por países. Destaca-se o tamanho da amostra, com 500.479 observações, de 50.998 empresas sediadas em 18 países de todos os continentes do mundo, sendo um da África, três das Américas, seis da Ásia, seis da Europa, um Transcontinental (Rússia) e um da Oceania. Com essa configuração de amostra estão presentes todos os países do BRICS.

Tabela 1 – Distribuição amostral

Distribuição da Amostra	
Total de empresas	50.998
Países selecionados	18
Total de observações	500.479
Distribuição da amostra por país	
África do Sul	6.528
Alemanha	14.310
Austrália	27.975
Brasil	7.007
Canadá	39.972
China	42.091
Cingapura	10.493
Coréia do Sul	27.210
EUA	127.475
França	14.955
Grécia	5.292
Índia	31.232
Japão	76.182
Malásia	17.169
Polônia	5.700
Reino Unido	34.114
Rússia	4.682
Suécia	8.092

Fonte: elaboração própria.

As empresas dos EUA têm a maior participação na amostra, com 127.475 observações.

A Rússia tem a menor participação, com 4.682 observações e foi mantida na amostra por pertencer ao BRICS, apesar de não figurar entre os 18 países com maior quantidade de observações na base de dados bruta extraída da base Thomson Reuters que originalmente contava com 126 países.

O Brasil apresentou 7.007 observações e figurava como o 15º maior país em número de observações da base de dados bruta.

Os retornos de mercado anuais e quinquenais foram calculados e a seguir classificados usando os tercís (divisão do intervalo de uma distribuição de frequência em três classes de igual ou aproximado número de observações) para a criação de três categorias: baixo, médio e alto retorno de ações, conforme Tabela 2.

Tabela 2 – Distribuição em tercís dos retornos anuais e quinquenais

Tercís do retorno de mercado	Anual		Quinquenal	
	Freq.	%	Freq.	%
1 – Baixo	149.854	33,34	93.457	33,34
2 – Médio	149.834	33,33	93.450	33,33
3 – Alto	149.793	33,33	93.442	33,33
Total	449.481	100	280.349	100

Fonte: elaboração própria.

Os indicadores contábeis (alavancagem, liquidez, lucratividade e giro do ativo) e as variáveis macroeconômicas (PIB, inflação e taxa de juros) encontram-se em números percentuais na base de dados e dessa forma são apresentados nas Tabelas 3 a 6 a seguir. Já os índices de mercado, encontram-se na base de dados na forma original pela qual são reportados ao mercado em seus respectivos países.

As Tabelas 3 e 4, a seguir, apresentam as estatísticas descritivas das variáveis utilizadas no estudo, segmentadas pelos modelos de curto e longo prazos. Considerando a série histórica de 22 anos para o modelo de curto prazo (anual) e de 18 anos para o modelo de longo prazo (quinquenal), as médias das variáveis macroeconômicas em ambos os modelos possuem pouca dispersão. Já os índices de mercado apresentam grande dispersão porque são apresentadas pelo conjunto de 18 países cujas bolsas de valores tem índices de grandezas variadas, entretanto, essa dispersão não impacta os modelos econométricos uma vez que os dados são rodados em painel por país, de forma que o índice de mercado de cada país interage com as variáveis de empresas daquele mesmo país.

Tabela 3 – Curto Prazo – toda amostra utilizada (1996 a 2017) – estatística descritiva

Variáveis	Estatísticas Descritivas				
	Obs.	Média	Desvio Padrão	Mínimo	Máximo
Curto prazo					
alav_1ano	488.250	-8.399	1.939.702	-1.050.000.000	4.127.716
ga_1ano	487.806	2	222	-7	127.843
liq_1ano	366.835	35	281	-216	169.789
lucr_1ano	449.879	-57	5.780	-2.206.967	124.863
PIB	489.547	3	3	-9	15
Inflação	489.547	2	2	-2	86
Juros	489.547	5	7	0	86
ind_mercado	476.011	6.611	8.885	45	76.402

Fonte: elaboração própria.

Tabela 4 – Longo prazo – toda amostra utilizada (2000 a 2017) – estatística descritiva

Variáveis	Estatísticas Descritivas				
	Obs.	Média	Desvio Padrão	Mínimo	Máximo
Longo prazo					
alav_5anos	342.462	-4.287	702.801	-210.000.000	842.731
ga_5anos	341.378	2	129	-1	57.833
liq_5anos	333.465	36	295	-216	169.789
lucr_5anos	283.614	-6	778	-177.696	56.898
PIB5a	308.218	3	3	-6	12
inflação5a	308.218	2	2	-1	39
juros5a	308.218	5	6	0	72
ind_mercado5a	297.946	6.695	8.422	133	62.839

Fonte: elaboração própria.

A grande dispersão que se percebe nas estatísticas descritivas dos indicadores contábeis, em especial do indicador de alavancagem, deriva da presença de outliers na amostra que não foram retirados. A decisão do autor em não tratar tais outliers se justifica pela amplitude da amostra que abastece os modelos de curto prazo, com 296.625 observações válidas, e de longo prazo, 153.929 observações válidas, conforme resultados das regressões logísticas realizadas. Como tais informações são oriundas de 18 países diferentes, com realidades econômicas e sociais distintas ao longo dos 22 anos de série histórica, além da diversidade de seus sistemas contábeis e legais, isto naturalmente implicaria em variabilidade das informações, em especial dos indicadores contábeis, tornando arbitraria a decisão pela retirada dos outliers.

Entretanto, dada a dispersão verificada em algumas variáveis, foram criadas as tabelas de percentis para se buscar visualizar o impacto dessa dispersão sobre os modelos econométricos, conforme Tabelas 5 e 6 a seguir.

Tabela 5 – Curto Prazo – toda amostra utilizada (1996 a 2017) – distribuição em percentis

Variáveis	Percentis									
	Obs.	1%	5%	10%	25%	50%	75%	90%	95%	99%
Curto prazo										
alav_1ano	488.250	-1.877	-49	8	29	49	70	87	94	99
ga_1ano	487.806	0	0	0	0	1	1	2	2	4
liq_1ano	366.835	0	1	4	12	28	57	86	96	100
lucr_1ano	449.879	-457	-97	-43	-5	4	10	18	26	58
PIB	489.547	-5	-1	0	2	3	4	7	9	11
Inflação	489.547	-1	-0	-0	1	2	3	5	7	11
Juros	489.547	0	0	1	2	5	7	9	11	44
ind_mercado	476.011	352	880	1.112	1.426	3.104	8.988	16.111	19.114	51.284

Fonte: elaboração própria.

Tabela 6 – Longo prazo – toda amostra utilizada (2000 a 2017) – distribuição em percentis

Variáveis	Percentis									
	Obs.	1%	5%	10%	25%	50%	75%	90%	95%	99%
Longo prazo										
alav_5anos	342.462	-2.049	-42	11	30	48	67	82	89	97
ga_5anos	341.378	0	0	0	0	1	1	2	2	4
liq_5anos	333.465	1	4	7	15	30	53	77	88	97
lucr_5anos	283.614	-194	-57	-26	-2	4	10	16	22	40
PIB5a	308.218	-0	0	1	1	3	4	7	9	11
inflação5a	308.218	-1	-0	-0	1	2	3	4	7	10
juros5a	308.218	0	0	0	2	5	6	9	11	41
ind_mercado5a	297.946	320	1.120	1.186	1.381	3.073	11.183	14.162	16.457	49.676

Fonte: elaboração própria.

A análise da distribuição em percentis dos valores das variáveis dos modelos de curto e longo prazos demonstra que as médias e desvios padrões estão refletindo a presença de outliers inferiores concentrados nos percentis de 1% a 5% e superiores a partir de 99%, não prejudicando o desempenho dos modelos logit uma vez que a característica metodológica dessas regressões afasta a influência desses valores discrepantes e sem conteúdo informacional.

Nos APÊNDICES D e E encontram-se um conjunto de tabelas suplementares que discriminam a estatística descritiva e distribuição em percentis das variáveis independentes dos modelos de curto e longo prazos. Essas estatísticas foram segregadas em 8 tabelas, quatro para cada um dos modelos e apresentam as variáveis da amostra total subdivididas em amostra de estimação dos parâmetros e amostra de predição, considerando que as previsões são realizadas em amostra *out-of-sample*.

Foi realizado teste de correlação entre as variáveis que compõem os MLM, aparentemente as correlações são aceitáveis e não há indícios de problemas de multicolinearidade entre as variáveis independentes, conforme Tabelas 15 (MLM de curto prazo) e 16 (MLM de longo prazo).

Tabela 15 – Correlação das variáveis do MLM de curto prazo

Correlação MLM de curto prazo	tret1a_ano_ano	alav_1ano	ga_1ano	liq_1ano	lucr_1ano	PIB	inflação	juros	ind_mercado
tret1a_ano_ano	1,0000								
alav_1ano	0,0020	1,0000							
ga_1ano	-0,0023	-0,0000	1,0000						
liq_1ano	-0,0151	0,0015	0,0027	1,0000					
lucr_1ano	0,0033	0,0002	-0,0010	-0,0042	1,0000				
PIB	0,0106	0,0017	-0,0017	0,0760	0,0030	1,0000			
Inflação	-0,0091	-0,0010	0,0003	-0,1056	0,0009	0,3335	1,0000		
Juros	0,0112	0,0001	0,0003	-0,0516	0,0018	0,1966	0,5012	1,0000	
ind_mercado	0,0176	0,0011	-0,0012	-0,0002	0,0024	-0,1733	-0,0523	0,2059	1,0000

Nota: Total de 410.468 observações, excluindo os *missings values* nos cruzamentos das variáveis

Fonte: elaboração própria.

Tabela 16 – Correlação das variáveis do MLM de longo prazo

Correlação MLM de longo prazo	tret5a_ano_ano	alav_5anos	ga_5anos	liq_5anos	lucr_5anos	PIB5a	Inflação5a	Inte rest5a	ind_mercado5a
tret5a_ano_ano	1,0000								
alav_5anos	0,0027	1,0000							
ga_5anos	0,0380	-0,0007	1,0000						
liq_5anos	-0,0827	0,0018	-0,1552	1,0000					
lucr_5anos	0,0201	0,0007	0,0248	-0,0158	1,0000				
PIB5a	0,1418	0,0005	-0,0541	-0,0780	0,0079	1,0000			
inflação5a	0,0594	-0,0009	-0,0248	-0,1231	0,0048	0,4771	1,0000		
juros5a	0,0567	-0,0004	-0,0223	-0,0425	0,0046	0,2614	0,5735	1,0000	
ind_mercado5a	-0,0356	0,0014	0,0014	-0,0047	0,0006	-0,2346	-0,0623	0,2648	1,0000

Nota: Total de 238.069 observações, excluídos os *missings values* nos cruzamentos das variáveis

Fonte: elaboração própria.

Percebe-se na comparação entre o MLM de curto e o de longo prazo que as correlações entre a variável dependente (constructo de desempenho de mercado) e todas as variáveis independentes (constructos de desempenho contábil, variáveis macroeconômicas e índice de mercado), aumentaram quando se alargou a janela temporal de um para cinco anos. Essa maior correlação identificada na comparação entre curto e longo prazos aponta de forma preliminar na direção de confirmação da hipótese desta pesquisa, a ser verificada na análise dos resultados das regressões logísticas dos MLM de curto e longo prazos e das predições das categorias de desempenho de mercado:

H₁: Os desempenhos contábil e de mercado de longo prazo são mais relacionados do que os desempenhos contábil e de mercado de curto prazo.

5.2. Análises dos MLM

Wulff (2015) explica que duas questões complicam a interpretação dos coeficientes em um MLM:

- a) os resultados representam contrastes entre as categorias, dificultando a visualização das implicações para cada categoria a partir dos coeficientes e diferentemente dos modelos binários, um sinal positivo de um coeficiente não significa necessariamente que um aumento na variável independente corresponda a um aumento na probabilidade da ocorrência de uma categoria específica;
- b) a relação entre as variáveis explicativas e a probabilidade de um determinado resultado de escolha é não linear e pode até mudar de sinal na distribuição de um único preditor.

Assim, um dos dispositivos que Wulff (2015) sugere para que pesquisadores realizem suas conclusões é a análise dos efeitos marginais que são apresentados ao longo desta seção. Para efeito de análise dos MLM nesta tese também será utilizada a RRR (*relative-risk ratio* ou razão de risco relativo), detalhada matematicamente na seção 2.3.1. Ainda assim, são apresentadas nesta seção as tabelas 17 e 18 com os coeficientes obtidos nos MLM de curto e longo prazos.

Os resultados apresentados na Tabela 17, a seguir, indicam a importância relativa de cada variável do MLM de curto prazo, onde a categoria “baixo retorno” é a base das estimativas em comparação com as categorias “médio e alto retornos”. Enquanto a categoria de médio retorno apontou que os parâmetros das variáveis giro do ativo (*ga_1ano*) e taxa de juros (*juros*) não foram significativos, a categoria de alto retorno apontou como não significativos os parâmetros das variáveis alavancagem (*alav_1ano*), giro do ativo (*ga_1ano*), lucratividade (*lucr_1ano*) e da constante do modelo.

Tabela 17 – Estimativas dos Parâmetros do MLM de curto prazo

MLM de curto prazo	B	Erro padrão	Z	P> z 	[95% Conf.Interval]	
Baixo	(base outcome)					
Médio						
alav_1ano	0,0000	0,0000	2,75	0,01	0,0000	0,0001
ga_1ano	-0,0013	0,0012	-1,15	0,25	-0,0036	0,0009
liq_1ano	-0,0050	0,0002	-31,06	0,00	-0,0054	-0,0047
lucr_1ano	0,0000	0,0000	3,62	0,00	0,0000	0,0000
PIB	-0,0257	0,0015	-16,66	0,00	-0,0287	-0,0226
Inflação	-0,0795	0,0023	-33,85	0,00	-0,0841	-0,0749
Juros	-0,0012	0,0008	-1,49	0,14	-0,0029	0,0004
ind_mercado	0,0000	0,0000	18,42	0,00	0,0000	0,0000
Constante	0,4186	0,0105	39,87	0,00	0,3981	0,4392
Alto						
alav_1ano	0,0000	0,0000	1,58	0,12	0,0000	0,0000
ga_1ano	-0,0003	0,0004	-0,63	0,53	-0,0011	0,0006
liq_1ano	-0,0008	0,0002	-4,95	0,00	-0,0011	-0,0005
lucr_1ano	0,0000	0,0000	1	0,32	0,0000	0,0000
PIB	0,0060	0,0015	3,97	0,00	0,0030	0,0089
Inflação	-0,0189	0,0020	-9,33	0,00	-0,0229	-0,0150
Juros	0,0077	0,0008	10,23	0,00	0,0063	0,0092
ind_mercado	0,0000	0,0000	7,37	0,00	0,0000	0,0000
Constante	0,0166	0,0105	1,59	0,11	-0,0039	0,0371

Fonte: elaboração própria.

A seguir, na Tabela 18 é indicada a importância relativa das variáveis do MLM de longo prazo e percebe-se uma melhora na significância dos coeficientes em comparação com o MLM de curto prazo (Tabela 17). Nas estimativas para a categoria médio retorno apenas o parâmetro da variável giro do ativo (ga_5anos) não se apresenta significativo, enquanto a variável alavancagem (alav_5anos) não é significativa na categoria alto retorno. Tal melhora corrobora a percepção de que o alargamento da janela temporal de um para cinco anos aumenta a relação entre o desempenho contábil e o desempenho de mercado.

Tabela 18 – Estimativas dos Parâmetros do MLM de longo prazo

MLM de longo prazo	B	Erro padrão	z	P> z	[95% Conf.Interval]	
Baixo	(base outcome)					
Médio						
alav_5anos	0,0018	0,0002	8,08	0,00	0,0014	0,0022
ga_5anos	0,0117	0,0092	1,28	0,20	-0,0063	0,0296
liq_5anos	-0,0035	0,0003	-11,49	0,00	-0,0041	-0,0029
lucr_5anos	0,0124	0,0003	44,37	0,00	0,0118	0,0129
PIB5a	0,0294	0,0032	9,25	0,00	0,0232	0,0356
inflação5a	-0,1135	0,0050	-22,63	0,00	-0,1233	-0,1036
juros5a	0,0100	0,0015	6,62	0,00	0,0071	0,0130
ind_mercado5a	0,0000	0,0000	2,62	0,01	0,0000	0,0000
cons	0,2570	0,0232	11,06	0,00	0,2115	0,3026
Alto						
alav_5anos	0,0002	0,0001	1,6	0,11	0,0000	0,0004
ga_5anos	0,0419	0,0091	4,63	0,00	0,0242	0,0597
liq_5anos	0,0009	0,0003	2,97	0,00	0,0003	0,0014
lucr_5anos	0,0151	0,0003	51,96	0,00	0,0145	0,0157
PIB5a	0,1247	0,0029	42,74	0,00	0,1190	0,1304
inflação5a	0,0073	0,0043	1,71	0,09	-0,0011	0,0157
juros5a	0,0245	0,0013	18,2	0,00	0,0219	0,0271
ind_mercado5a	0,0000	0,0000	-8,66	0,00	0,0000	0,0000
cons	-0,4232	0,0222	-19,04	0,00	-0,4668	-0,3796

Fonte: elaboração própria.

As estatísticas de RRR apresentadas nas Tabelas 19 e 20 demonstram que todas as variáveis independentes dos MLM de curto e longo prazos estão muito próximas de 1, com exceção da variável PIB no MLM de longo prazo.

Ainda assim, percebe-se um crescimento no valor das RRR das variáveis contábeis quando se compara os MLM de curto em direção ao de longo prazo, corroborando, mesmo que de forma tênue, a hipótese de que o alargamento da janela temporal eleva o poder explicativo do constructo de desempenho contábil em relação ao do desempenho de mercado.

A interpretação da RRR se refere sempre à categoria base, neste caso a de baixo retorno, sendo que quanto maior o valor acima de 1 para RRR, menor serão as chances de que a cada variação positiva no valor daquela variável, a empresa se movimente em direção a baixo retorno, se ela estiver em médio ou alto retorno ($RRR > 1$). Por outro lado, quanto mais abaixo de 1 estiver o valor da RRR, maior serão as chances de que a cada variação positiva no valor daquela variável, a empresa se movimente em direção de baixo retorno ($RRR < 1$).

Analisando a Tabela 19, na categoria médio retorno do MLM de curto prazo existem 3 variáveis independentes significativas e com algum efeito negativo ($RRR < 1$), ou seja, a probabilidade de as ações das empresas pertencerem a categoria médio retorno, em comparação

com a categoria base – baixo retorno, tende a diminuir quando há aumento da liquidez, do PIB ou da inflação.

Tabela 19 – Razão de Risco Relativo (RRR) do MLM de curto prazo

MLM de curto prazo	RRR	Erro padrão	Z	P> z 	[95% Conf.Interval]	
Baixo	(base outcome)					
Médio						
alav_lano	1,0000	0,0000	2,75	0,01	1,0000	1,0001
ga_lano	0,9987	0,0012	-1,15	0,25	0,9964	1,0009
liq_lano	0,9950	0,0002	-31,06	0,00	0,9947	0,9953
lucr_lano	1,0000	0,0000	3,62	0,00	1,0000	1,0000
PIB	0,9747	0,0015	-16,66	0,00	0,9717	0,9776
Inflação	0,9236	0,0022	-33,85	0,00	0,9194	0,9279
Juros	0,9988	0,0008	-1,49	0,14	0,9971	1,0004
ind_mercado	1,0000	0,0000	18,42	0,00	1,0000	1,0000
Constante	1,5199	0,0160	39,87	0,00	1,4889	1,5515
Alto						
alav_lano	1,0000	0,0000	1,58	0,12	1,0000	1,0000
ga_lano	0,9997	0,0004	-0,63	0,53	0,9989	1,0006
liq_lano	0,9992	0,0002	-4,95	0,00	0,9989	0,9995
lucr_lano	1,0000	0,0000	1,00	0,32	1,0000	1,0000
PIB	1,0060	0,0015	3,97	0,00	1,0030	1,0089
Inflação	0,9812	0,0020	-9,33	0,00	0,9773	0,9851
Juros	1,0078	0,0008	10,23	0,00	1,0063	1,0093
ind_mercado	1,0000	0,0000	7,37	0,00	1,0000	1,0000
Constante	1,0167	0,0106	1,59	0,11	0,9961	1,0378

Fonte: elaboração própria.

Na categoria alto retorno do MLM de curto prazo (Tabela 19) existem 2 variáveis independentes significativas e com algum efeito negativo ($RRR < 1$), ou seja, a probabilidade de as ações das empresas pertencerem a categoria alto retorno, em comparação com a categoria base – baixo retorno, tende a diminuir quando há aumento da liquidez ou da inflação.

Analisando a Tabela 20, na categoria médio retorno do MLM de longo prazo existem 2 variáveis independentes significativas e com algum efeito negativo ($RRR < 1$), ou seja, a probabilidade de as ações das empresas pertencerem a categoria médio retorno, em comparação com a categoria base – baixo retorno, tende a diminuir quando há aumento da liquidez ou da inflação.

Tabela 20 – Razão de Risco Relativo (RRR) do MLM de longo prazo

MLM de longo prazo	RRR	Erro padrão	Z	P> z 	[95% Conf.Interval]	
Baixo	(base outcome)					
Médio						
alav_5anos	1,0018	0,0002	8,08	0,00	1,0014	1,0022
ga_5anos	1,0118	0,0093	1,28	0,20	0,9938	1,0301
liq_5anos	0,9965	0,0003	-11,49	0,00	0,9959	0,9971
lucr_5anos	1,0124	0,0003	44,37	0,00	1,0119	1,0130
PIB5a	1,0298	0,0033	9,25	0,00	1,0234	1,0363
inflação5a	0,8927	0,0045	-22,63	0,00	0,8840	0,9015
juros5a	1,0101	0,0015	6,62	0,00	1,0071	1,0131
ind_mercado5a	1,0000	0,0000	2,62	0,01	1,0000	1,0000
cons	1,2931	0,0301	11,06	0,00	1,2355	1,3533
Alto						
alav_5anos	1,0002	0,0001	1,6	0,11	1,0000	1,0004
ga_5anos	1,0428	0,0094	4,63	0,00	1,0245	1,0615
liq_5anos	1,0009	0,0003	2,97	0,00	1,0003	1,0014
lucr_5anos	1,0152	0,0003	51,96	0,00	1,0146	1,0158
PIB5a	1,1328	0,0033	42,74	0,00	1,1263	1,1393
inflação5a	1,0073	0,0043	1,71	0,09	0,9989	1,0158
juros5a	1,0248	0,0014	18,2	0,00	1,0221	1,0275
ind_mercado5a	1,0000	0,0000	-8,66	0,00	1,0000	1,0000
cons	0,6550	0,0146	-19,04	0,00	0,6270	0,6841

Fonte: elaboração própria.

Na categoria alto retorno do MLM de longo prazo (Tabela 20) todas as variáveis independentes tem algum efeito positivo ($RRR > 1$), ou seja, a probabilidade de as ações das empresas pertencerem a categoria alto retorno, em comparação com a categoria base – baixo retorno, tende a aumentar se há aumento em qualquer das variáveis significativas do modelo.

Analisando os indicadores de desempenho dos MLM de curto e longo prazos (Tabela 21), verifica-se que o pseudo coeficiente de determinação (Pseudo R^2), apresenta valores modestos para ambos os modelos, indicando baixo poder explicativo. Entretanto, percebe-se melhora significativa no MLM de longo prazo em comparação com o de curto prazo, aumentando seu valor de 0,008 para 0,045, revelando que o alargamento da janela temporal melhora o poder explicativo do modelo. Em relação à qualidade de ajuste dos modelos, os indicadores Critério de Informação de Akaike (AIC) e Critério de Informação Bayesiano (BIC), disponíveis na Tabela 21, tiveram redução em suas estimativas para o MLM de longo prazo em comparação com o de curto prazo, também sinalizando no sentido de maior poder explicativo do MLM de longo prazo. A razão de verossimilhança ou likelihood ratio (LR) apresenta o valor de 5.219,58 para o MLM de curto prazo e de 15.354,60 para o de longo prazo, sendo ambos significantes quanto às variáveis associadas aos modelos.

Tabela 21 – Indicadores de desempenho dos MLM

MLM	Curto prazo	Longo prazo
Número de observações	296.625	153.929
LR chi2(16)	5.219,58	15.354,60
Prob > chi2	0,0000	0,0000
Pseudo R ²	0,0080	0,0455
AIC	646.473,40	322.495,30
BIC	646.664,20	322.674,30

Fonte: elaboração própria.

5.3. Análises das predições realizadas

Utilizando-se os parâmetros calculados pelos modelos, por meio da função *predict* do STATA foi realizada a predição das probabilidades de as ações das empresas estarem nas categorias 1 (baixo retorno), 2 (médio retorno) e 3 (alto retorno) para o período de 2013 a 2017 (5 anos). A partir da comparação das probabilidades preditas, são criadas variáveis de categorias preditas para os anos de 2013, 2014, 2015, 2016 e 2017 para cada empresa que teve dados calculados pelo MLM com as variáveis de categorias previstas assumindo o valor da categoria com maior probabilidade de ocorrer (probabilidade dominante).

Então, são comparados os resultados previstos e realizados, relatando-se o nível de acertos do MLM de curto prazo (Tabela 22):

Tabela 22 – MLM de curto prazo: acertos de predição

MLM de curto prazo (1 ano)						
Anos	Acertos				Previsões realizadas	% Acertos
	1	2	3	Total		
2013	2.455	4.552	1.708	8.715	23.532	37,03%
2014	1.483	4.493	3.230	9.206	23.784	38,71%
2015	804	5.492	3.173	9.469	23.587	40,14%
2016	763	4.763	2.408	7.934	23.558	33,68%
2017	1.310	4.765	2.092	8.167	23.989	34,04%
Total	6.815	24.065	12.611	43.491	118.450	36,72%

Fonte: elaboração própria.

A exemplo do MLM de curto prazo, com base nos parâmetros do MLM de longo prazo são criadas variáveis de categorias preditas para os anos de 2013, 2014, 2015, 2016 e 2017, comparados os resultados previstos e realizados, relatando-se o nível de acertos do MLM de longo prazo (Tabela 23):

Tabela 23 – MLM de longo prazo: acertos de predição

MLM de longo prazo (5 anos)						
Anos	Acertos				Previsões realizadas	% Acertos
	1	2	3	Total		
2013	2.153	3.308	2.326	7.787	17.292	45,03%
2014	2.061	2.880	2.121	7.062	17.314	40,79%
2015	2.049	3.362	2.175	7.586	17.375	43,66%
2016	1.976	3.828	2.436	8.240	17.622	46,76%
2017	1.859	3.685	2.363	7.907	17.803	44,41%
Total	10.098	17.063	11.421	38.582	87.406	44,14%

Fonte: elaboração própria.

Os resultados das predições dos MLM de curto e longo prazos corroboram com a hipótese da pesquisa, pois:

- a) o percentual total de acertos no MLM de curto prazo é de 36,72%, sendo apenas 10,15% superior ao limite de 33,33% que são as chances de o retorno de uma ação estar aleatoriamente em qualquer um dos tercís;
- b) o total de acertos no MLM de longo prazo é de 44,14%, superando em 20,22% o total de acertos do modelo de curto prazo e distanciando-se em 32,43% do limite de 33,33% que representa as chances de o retorno de uma ação estar aleatoriamente em qualquer um dos tercís.

Assim, dada a diferença entre os níveis de acertos do modelo de longo prazo e o de curto prazo, verifica-se que o alargamento da janela temporal aumenta a relação entre desempenho de mercado e desempenho contábil, aderente a hipótese H_1 de que os desempenhos contábil e de mercado de longo prazo são mais relacionados do que os desempenhos contábil e de mercado de curto prazo.

Comparando as Tabelas 24 e 25 que apresentam os efeitos marginais dos MLM de curto e longo prazos, respectivamente, percebe-se que os efeitos marginais das variáveis contábeis estão muito próximos do valor 0,33 no MLM de curto prazo, ratificando o baixo poder explicativo dessas variáveis nas mudanças de tercís que aconteceriam com os retornos das ações, demonstrado nos acertos preditivos informados na Tabela 22 que informou 36,72% de acertos. Já no MLM de longo prazo, os efeitos marginais se distanciam do valor 0,33, com destaque para as variáveis giro do ativo e lucratividade que apresentaram as maiores variações em comparação com MLM de curto prazo, redundando em maior poder explicativo ao MLM de longo prazo que se traduz em 44,14% de acertos reportado na Tabela 23, resultado que permite afirmar que o alargamento da janela temporal do desempenho contábil aumenta o relacionamento deste com o desempenho de mercado das empresas.

Tabela 24 – Efeito marginal dos valores preditos – MLM curto prazo

Variáveis	Baixo Retorno		Médio Retorno		Alto Retorno	
	Efeito Marginal	Erro padrão*	Efeito Marginal	Erro padrão*	Efeito Marginal	Erro padrão*
alav_1ano	0,3252	0,0009	0,3394	0,0009	0,3354	0,0009
ga_1ano	0,3337	0,0069	0,3265	0,0122	0,3398	0,0072
liq_1ano	0,3351	0,0010	0,3231	0,0010	0,3418	0,0010
lucr_1ano	0,3250	0,0009	0,3398	0,0009	0,3352	0,0009
PIB	0,3731	0,0152	0,1177	0,0065	0,5092	0,0159
Inflação	0,6923	0,0200	0,0166	0,0017	0,2911	0,0196
Juros	0,2908	0,0063	0,2874	0,0065	0,4219	0,0070
ind_mercado	0,3356	0,0011	0,3272	0,0011	0,3372	0,0011

Nota: Método Delta - efeitos marginais utilizou a média, * p<0,001

Fonte: elaboração própria.

Tabela 25 – Efeito marginal dos valores preditos – MLM longo prazo

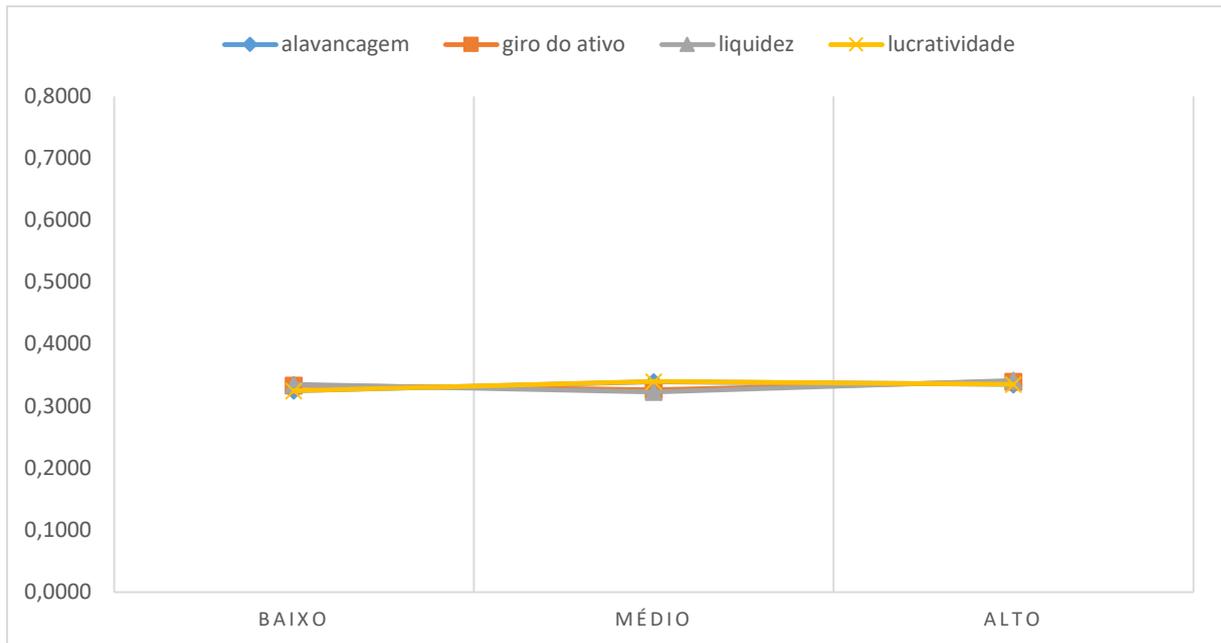
Variáveis	Baixo Retorno		Médio Retorno		Alto Retorno	
	Efeito Marginal	Erro padrão*	Efeito Marginal	Erro padrão*	Efeito Marginal	Erro padrão*
alav_5anos	0,3088	0,0011	0,3472	0,0012	0,3440	0,0012
ga_5anos	0,0984	0,0331	0,1803	0,0577	0,7213	0,0710
liq_5anos	0,3140	0,0015	0,3307	0,0016	0,3552	0,0015
lucr_5anos	0,1794	0,0020	0,3877	0,0032	0,4329	0,0033
PIB5a	0,0060	0,0005	0,0120	0,0015	0,9820	0,0017
inflação5a	0,3973	0,0452	0,0019	0,0004	0,6008	0,0452
juros5a	0,1738	0,0077	0,2939	0,0108	0,5322	0,0112
ind_mercado5a	0,3062	0,0016	0,3360	0,0017	0,3578	0,0016

Nota: Método Delta - efeitos marginais utilizou a média, * p<0,001

Fonte: elaboração própria.

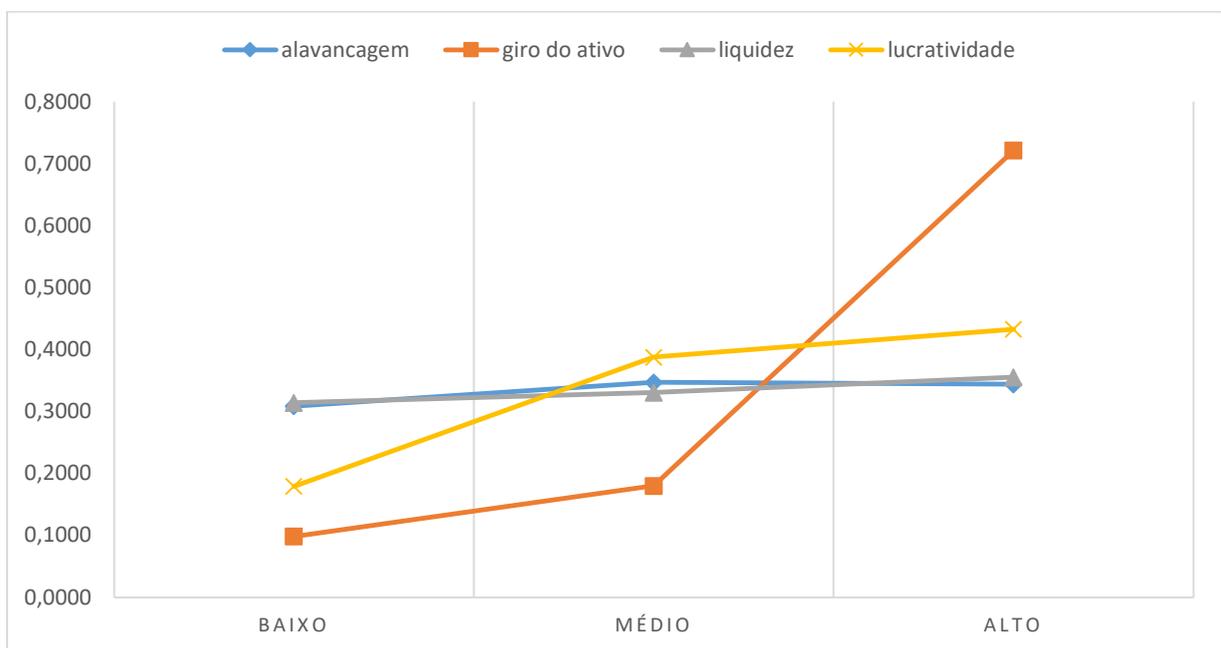
A seguir, os Gráficos 1 e 2 demonstram as variações em termos de efeitos marginais do constructo de desempenho contábil verificadas entre os MLM de curto e longo prazos, tornando visível o efeito que o alargamento da janela temporal exerceu sobre a relação entre desempenho de mercado e desempenho contábil, em especial os indicadores de giro do ativo e lucratividade.

Gráfico 1 – Efeitos marginais do desempenho contábil – MLM de curto prazo



Fonte: elaboração própria.

Gráfico 2 – Efeitos marginais do desempenho contábil – MLM de longo prazo



Fonte: elaboração própria.

Como o escopo desta pesquisa concentra-se na verificação de se as informações de longo prazo seriam mais úteis do que as de curto prazo para explicar a relação entre informações contábeis e retornos de ações, a análise dos resultados foi direcionada para a Tabela 26, a seguir, em que são comparados os níveis de predição dos MLM de curto e longo prazos, notadamente

em sua última linha que demonstra a melhora significativa do percentual total de acertos do modelo de longo prazo em comparação com o de curto prazo.

Tabela 26 – Comparação entre acertos de predição dos MLM de curto e longo prazos – Total geral

País	MLM de curto prazo			MLM de longo prazo			Variação Percentual
	Acertos	Previsões realizadas	% Acertos	Acertos	Previsões realizadas	% Acertos	
África do Sul	421	1.177	35,77%	418	912	45,83%	28,14%
Alemanha	1.089	2.580	42,21%	931	2.139	43,53%	3,12%
Austrália	2.913	7.796	37,37%	3.346	5.788	57,81%	54,71%
Brasil	378	1.503	25,15%	280	1.233	22,71%	-9,71%
Canadá	2.620	9.608	27,27%	3.860	7.144	54,03%	98,14%
China	5.925	17.599	33,67%	5.014	12.484	40,16%	19,30%
Cingapura	885	2.736	32,35%	566	1.922	29,45%	-8,96%
Coréia do Sul	3.030	8.729	34,71%	2.875	6.902	41,65%	20,00%
EUA	6.395	19.128	33,43%	5.697	13.308	42,81%	28,04%
França	1.212	2.937	41,27%	957	2.201	43,48%	5,36%
Grécia	298	973	30,63%	461	913	50,49%	64,86%
Índia	4.973	11.020	45,13%	2.694	7.080	38,05%	-15,68%
Japão	8.147	17.279	47,15%	7.501	15.121	49,61%	5,21%
Malásia	1.322	4.161	31,77%	829	2.414	34,34%	8,09%
Polônia	694	2.131	32,57%	513	1.433	35,80%	9,92%
Reino Unido	2.100	5.777	36,35%	1.946	4.207	46,26%	27,25%
Rússia	409	1.211	33,77%	233	896	26,00%	-23,00%
Suécia	680	2.105	32,30%	461	1.309	35,22%	9,02%
Total	43.491	118.450	36,72%	38.582	87.406	44,14%	20,22%

Fonte: elaboração própria.

Além da Tabela 26, acima, que totaliza o percentual de acertos nos 5 anos de previsão, no APÊNDICE F deste trabalho estão detalhados os resultados por países, ano a ano de previsão (2013 a 2017), onde é destacada a variação percentual entre os níveis de acertos dos modelos de curto e de longo prazos (Tabelas 27 a 31). Essas tabelas comparativas são ricas em informações e poderão ser exploradas em pesquisas futuras que busquem explicar, por exemplo:

- o comportamento de alguns mercados nacionais que apresentam maior percentual de acertos no modelo de curto do que no de longo prazo, contrariando a média geral da amostra, como é o caso de Brasil, Cingapura, Índia e Rússia (Tabela 26, acima);
- o alto percentual de acertos do modelo de longo prazo em mercados nacionais como Austrália, Canadá, Grécia e Japão, muito acima da média geral da amostra (Tabela 26, acima);

- c) o comportamento volátil no percentual de acertos dos modelos no mercado chinês ao longo dos anos de 2014 a 2017 (Tabelas 28, 29, 30 e 31 – APÊNDICE F), em comparação com a média da amostra nesses anos;
- d) a proximidade do percentual de acertos dos modelos de curto e longo prazos apresentados pelo mercado alemão nos anos de 2013 a 2016 (Tabelas 27, 28, 29 e 30 – APÊNDICE F), muito diferente do comportamento médio da amostra naqueles 4 anos;
- e) a média geral de acertos do modelo de curto prazo que tem seu ápice no ano de 2015 (9.469 acertos, 40,14%), 3º ano da série de 5 anos de predição, e seu mínimo no ano de 2016 (7.934 acertos, 33,68%);
- f) a média geral de acertos do modelo de longo prazo que tem seu ápice no ano de 2016 (8.240 acertos, 46,76%), 4º ano da série de 5 anos de predição, e seu mínimo no ano de 2014 (7.062 acertos, 40,79%).

5.4. Teste e procedimentos de robustez

Os modelos analisados nas seções anteriores apresentaram resultados que permitem afirmar que o alargamento da janela temporal de um para cinco anos aumenta a relação entre os desempenhos contábil e de mercado, relação esta que é capturada pelo MLM de longo prazo, melhorando seu poder explicativo e, conseqüentemente, aumentando a quantidade de acertos em suas predições em comparação com o MLM de curto prazo.

A metodologia aplicada nesses modelos reúne as informações dos 18 países da amostra para estimar os parâmetros e depois prever o desempenho de mercado de todas as empresas da amostra, sem realizar cortes por países, setores econômicos, sistemas legais ou políticos, sistemas contábeis ou quaisquer outras características que pudessem segmentar a amostra, ressaltando, conforme delimitação deste estudo, que não há a intenção neste trabalho em determinar os retornos das ações e sim verificar se as informações de longo prazo seriam mais úteis para explicar a relação entre informações contábeis e retornos de ações do que as de curto prazo. Dessa forma, os modelos são simplificados e não buscam abranger ou analisar as inúmeras informações e variáveis que afetam o retorno de ações que é a base da definição das variáveis categóricas dos modelos testados.

A fim de verificar se essa simplificação metodológica poderia trazer algum viés para os resultados das predições, foi realizado teste de robustez na forma de aplicação dos MLM de curto e longo prazos em amostra de um país específico cujos resultados pudessem ser

comparados às predições realizadas com a amostra dos 18 países. Dados o tamanho e a importância de seu mercado de capitais, os EUA foram escolhidos para esta comparação.

Assim, repetiu-se os procedimentos numa amostra que continha apenas informações de empresas sediadas nos EUA, conforme *Do-File* específico (APÊNDICE G), sendo os parâmetros de desempenho dos MLM de curto e longo prazos com empresas dos EUA apresentados na Tabela 32:

Tabela 32 – Indicadores de desempenho dos MLM – amostra EUA

MLM	Curto prazo	Longo prazo
Número de observações	81.488	41.746
LR chi2(16)	705,02	3.352,87
Prob > chi2	0,0000	0,0000
Pseudo R ²	0,0039	0,0367

Fonte: elaboração própria.

Os resultados de predições da amostra EUA estão resumidos nas Tabelas 33 e 34, a seguir, e estão em linha com o total geral alcançado pelas predições da amostra com 18 países.

Tabela 33 – MLM de curto prazo: acertos de predição EUA

MLM de curto prazo						
Anos	Acertos				Previsões realizadas	% Acertos
	1	2	3	Total		
2013	1	1.201	270	1.472	3.972	37,06%
2014		1.168	209	1.377	4.026	34,20%
2015	2	1.126	266	1.394	3.891	35,83%
2016		1.165	141	1.306	3.687	35,42%
2017		1.190	47	1.237	3.552	34,83%
Total	3	5.850	933	6.786	19.128	35,48%

Fonte: elaboração própria.

Tabela 34 – MLM de longo prazo: acertos de predição EUA

MLM de longo prazo						
Anos	Acertos				Previsões realizadas	% Acertos
	1	2	3	Total		
2013	195	613	433	1.241	2.791	44,46%
2014	214	747	241	1.202	2.717	44,24%
2015	222	727	184	1.133	2.627	43,13%
2016	241	759	129	1.129	2.579	43,78%
2017	234	789	68	1.091	2.594	42,06%
Total	1.106	3.635	1.055	5.796	13.308	43,55%

Fonte: elaboração própria.

Percebe-se que ao isolar as empresas dos EUA em amostra específica e usar os próprios parâmetros da sua amostra para realizar as previsões, seus resultados melhor se alinharam aos resultados da amostra 18 países, em comparação às suas previsões estimadas com os parâmetros da amostra geral (18 países), conforme Tabela 35.

Tabela 35 – Comparação de acertos totais entre as amostras EUA e 18 países

Total Geral	MLM de curto prazo			MLM de longo prazo			Variação Percentual
	Acertos	Previsões realizadas	% Acertos	Acertos	Previsões realizadas	% Acertos	
Amostra 18 países	43.491	118.450	36,72%	38.582	87.406	44,14%	20,22%
EUA previsão com parâmetros dos 18 países	6.395	19.128	33,43%	5.697	13.308	42,81%	28,04%
EUA previsão com seus próprios parâmetros	6.786	19.128	35,48%	5.796	13.308	43,55%	22,75%

Fonte: elaboração própria.

Diante desses resultados é possível afirmar que a metodologia é robusta o suficiente para ser replicada em um único país e ratifica a percepção de que o alargamento da janela temporal aumenta a relação entre desempenhos contábil e de mercado, como demonstra a melhora do poder explicativo do MLM de longo prazo sobre o de curto prazo refletido no maior percentual de acertos de previsão.

5.5. Síntese dos resultados em relação às hipóteses da pesquisa

Na comparação das correlações entre a variável dependente que representa o desempenho de mercado e as independentes das quais fazem parte o desempenho contábil, nos MLM de curto e longo prazos (Tabelas 15 e 16), a maior correlação identificada no modelo de longo prazo já apontava de forma preliminar na direção de confirmação da hipótese desta pesquisa.

A importância relativa das variáveis do MLM de longo prazo com maior significância dos coeficientes em comparação com o MLM de curto prazo (Tabela 18 versus Tabela 17) também corroborava para a percepção de que o alargamento da janela temporal de um para cinco anos aumentaria a relação entre os desempenhos contábil e de mercado, em linha com a hipótese da pesquisa. Tal percepção é reforçada pelas RRR do MLM de longo prazo (Tabela

20) que apresentavam melhor desempenho nos efeitos das variáveis em comparação com o MLM de curto prazo (Tabela 19).

Apesar de os indicadores de desempenho dos MLM de curto e longo prazos (Tabela 21) apresentarem valores modestos no Pseudo R^2 , o de longo prazo (0,045) era muito superior ao de curto prazo (0,008), sinalizando na direção de confirmação da hipótese da pesquisa, assim como também apontavam os critérios de qualidade de ajuste dos modelos (AIC e BIC) para o maior poder explicativo do MLM de longo prazo em comparação com o de curto prazo.

Já os resultados das predições dos MLM de curto e longo prazos, principal etapa dos procedimentos, corroboram com a hipótese da pesquisa na medida que confirma o baixo poder preditivo do modelo de curto prazo com percentual total de acertos de 36,72%, apenas 10,15% superior ao limite de 33,33% que são as chances de o retorno de uma ação estar aleatoriamente em qualquer um dos tercis, enquanto o de longo prazo teve 44,14%, superando em 20,22% o total de acertos do modelo de curto prazo e distanciando-se em 32,43% do limite de 33,33% que representa as chances de o nível do retorno de uma ação estar aleatoriamente em uma categoria.

Assim, verifica-se que o alargamento da janela temporal aumenta a relação entre desempenho de mercado e desempenho contábil, aderente a hipótese H_1 de que os desempenhos contábil e de mercado de longo prazo são mais relacionados do que os desempenhos contábil e de mercado de curto prazo.

Essa constatação é reforçada pelos resultados dos efeitos marginais (Tabelas 24 e 25) em que os das variáveis contábeis de curto prazo estão muito próximos do valor 0,33, ratificando o baixo poder explicativo dessas variáveis, enquanto as de longo prazo se distanciam do valor 0,33, com destaque para as variáveis giro do ativo e lucratividade que apresentaram as maiores variações.

Por fim, ao se aplicar os modelos em amostra específica de empresas dos EUA, usando os próprios parâmetros dessa amostra para realizar as predições, diferentemente da metodologia base desta pesquisa que reunia as informações dos 18 países da amostra para estimar os parâmetros e depois previa o desempenho de mercado de todas as empresas dessa amostra, os resultados permitem afirmar que a metodologia é robusta o suficiente para ser replicada em um único país e ratifica a percepção de que o alargamento da janela temporal aumenta a relação entre desempenhos contábil e de mercado, como demonstra a melhora do poder explicativo do MLM de longo prazo sobre o de curto prazo refletido no maior percentual de acertos de predição. Ou seja, os resultados do teste de robustez dos modelos, alinhados aos resultados da

amostra de 18 países (Tabela 35) permitem confirmar a hipótese H_1 discriminada nesta pesquisa.

6. CONSIDERAÇÕES FINAIS

Prever o nível de retorno das ações permitiria ao usuário da informação mitigar o risco de perda e potencializar suas chances de ganho, bem como orientar a atuação da empresa e de seus *stakeholders* a fim de melhorar sua performance. O desenvolvimento de um modelo que fosse aplicável à realidade despertaria o interesse do mercado e contribuiria em muito para o avanço dos estudos contábeis uma vez que seria aderente à principal função da Contabilidade que é gerar informação para a tomada de decisão de seus usuários (CPC, 2011).

Construir um modelo que possa processar as inúmeras informações e variáveis que afetam o retorno de ações a fim de prever seu desempenho com segurança, provavelmente não será uma tarefa viável. Entretanto, estabelecer um mínimo de relação entre os desempenhos contábil e de mercado já contribuiria para o processo de tomada de decisão do investidor, mitigando em algum percentual as incertezas que cercam sua decisão de investimento, em especial a de longo prazo.

É nessa ideia de se buscar evidências de relações entre informações contábeis e retornos futuros de ações que se insere esta pesquisa a qual adotou como pressuposto que ao alargar a janela temporal da análise, as tendências do desempenho contábil se alinhariam às do desempenho de mercado, permitindo no longo prazo identificar relações que pudessem ser utilizadas para prever a um determinado nível de probabilidade a direção do desempenho de mercado das empresas.

Fala-se em tendências porque se espera que o preço das ações de uma empresa seja reflexo das diversas escolhas realizadas pela administração da empresa durante o exercício, tais como: decisões administrativas, de compras, de vendas, de financiamento, de investimento, além de escolhas sobre práticas e políticas contábeis que afetam a divulgação de suas informações e, conseqüentemente, da percepção e expectativas do investidor que por sua vez afetam os preços das ações.

Como se as diversas escolhas que a administração da empresa fosse realizando ao longo dos períodos estivessem conduzindo a empresa à escolha do nível de retorno de suas ações que o mercado reconhecerá ao alinhar as expectativas dos investidores com as informações que fluem da empresa para o mercado reportando o nível de assertividade de suas decisões e atividades na direção de um maior giro do ativo e lucratividade.

E ao se referir a escolhas, insere-se a metodologia adotada nesta pesquisa ao utilizar-se dos MLM para prever desempenho de mercado, transformando o retorno de ações em variável categórica de nível de desempenho, habilitando os MLM a explicarem e preverem a probabilidade de um nível de desempenho de mercado ocorrer com base em variáveis explanatórias representadas por indicadores contábeis.

Dessa forma, este trabalho utilizou-se do MLM para analisar o comportamento da relação entre desempenhos contábil e de mercado no curto e no longo prazo, tendo por objetivo verificar se as informações de longo prazo seriam mais úteis para explicar a relação entre informações contábeis e retornos de ações do que as de curto prazo no intuito de embasar a hipótese de que as informações contábeis têm poder preditivo para o desempenho de mercado das empresas, representado pelo preço de suas ações, quando observadas no longo prazo.

Para tanto, foram analisados e comparados os impactos no poder preditivo dos modelos ao se utilizar dados anuais e quinquenais, cujos resultados encontrados oferecem indícios de que o desempenho de mercado no curto prazo não tem relação com o desempenho contábil de curto prazo das empresas, ou se tiver é bastante limitado, e que por outro lado essa relação existe no longo prazo, conforme hipótese desenvolvida nesta pesquisa.

A amostra com 500.479 observações de 50.998 empresas, de 18 países diferentes, distribuídas pelos cinco continentes (África, América, Ásia, Europa e Oceania) e englobando os cinco países do BRICS, além de conferir um caráter global à pesquisa, abre as portas para futuras pesquisas que busquem explicar, por exemplo, os diferentes níveis de acertos por anos e/ou países da amostra ou a concentração de acertos em determinada categoria, dentre os três níveis de retorno dos modelos; o comportamento de mercados nacionais que apresentam maior percentual de acertos no modelo de curto do que no de longo prazo, aparentemente sinalizando que nesses países existiria *value relevance* no curto prazo; os motivos do alto percentual de acertos do modelo de longo prazo em alguns países; a volatilidade no percentual de acertos dos modelos observados em alguns anos do mercado chinês; a aparente similaridade de nível de *value relevance* no curto e longo prazos apresentados pelo mercado alemão em alguns dos anos preditos; e a identificação de qual seria o melhor prazo de ajuste para as predições, considerando que em determinados anos do período de cinco anos de predições, os resultados apontaram variações no percentual de acertos nos modelos de curto e longo prazos.

A análise e comparação, em recorte especial da amostra com empresas sediadas nos EUA, contra o resultado global alcançado com os dados de empresas de 18 países, apresentou resultados que permitem afirmar que a metodologia é robusta o suficiente para ser replicada em outras situações.

Os resultados das previsões dos MLM de curto e longo prazos corroboram com a hipótese da pesquisa, pois o percentual total de acertos no curto prazo é de 36,72%, sendo apenas 10,15% superior ao limite de 33,33% que são as chances de o retorno de uma ação estar aleatoriamente em qualquer um dos tercios, enquanto no longo prazo foi de 44,14%, superando em 20,22% o total de acertos do modelo de curto prazo e distanciando-se em 32,43% do limite de 33,33%, sendo importante ressaltar que os indicadores responsáveis por esse desempenho do modelo de longo prazo são giro do ativo e lucratividade, conforme demonstraram as variações em termos de efeitos marginais.

Esta pesquisa por meio de seus resultados espera ter contribuído para reforçar a importância relativa da Contabilidade em explicar desempenho econômico futuro das empresas e o preço de suas ações no mercado que, de certa forma, é o objetivo e principal função da Contabilidade.

A metodologia aplicada nos modelos de previsão desenvolvidos nesta pesquisa tem por base informações públicas e facilmente disponíveis ao mercado, cujos resultados sinalizam sua utilidade para investidores em ações que tenham objetivos de longo prazo na formação de carteiras com maior probabilidade de alcançar no horizonte de 5 anos desempenhos acima da média do mercado.

Esta pesquisa tem limitações em função de que a intenção não foi realizar a previsão de valores futuros de retornos das ações e sim verificar se as informações de longo prazo seriam mais úteis para explicar a relação entre informações contábeis e retornos de ações do que as de curto prazo. Assim, os modelos são simplificados e não buscam abranger ou analisar as inúmeras informações e variáveis que afetam o retorno de ações, pois, para tal verificação não era necessário prever o valor do retorno futuro e sim em qual categoria de nível de retorno a ação da empresa estaria no futuro dentre as categorias baixo, médio e alto retorno, sendo confirmada a hipótese pela comparação entre os resultados de previsão dos dois modelos, um de curto e outro de longo prazo.

REFERÊNCIAS

- ALTMAN, E. I. *Financial ratios, discriminant analysis and the prediction of corporate bankruptcy*. The Journal of Finance, vol. 23, nº 4, p.589-609. 1968.
- ALTMAN, E. I.; HALDEMAN, R. G.; NARAYANAN, P. *ZETA Analysis: A new model to identify bankruptcy risk of corporations*. Journal of Banking and Finance, vol. 1, p.29-54. 1977.
- BALL, R.; BROWN, P. *An empirical evaluation of accounting income numbers*. Journal of Accounting Research, vol.6, nº 2 (Autumn, 1968), p.159-178. 1968.
- BANCO MUNDIAL, disponível em <https://data.worldbank.org/indicador>, acesso em 07/12/2019.
- BARTH, M. E.; BEAVER, W. H.; LANDSMAN, W. R. *The relevance of the value relevance literature for financial accounting standard setting: another view*. Journal of Accounting and Economics, v. 31, p. 77–104, 2001.
- BEAVER, W. H. *Financial ratios as predictors of failure*. Journal of Accounting Research. Empirical Research in Accounting: Selected Studies, vol.4, p.71-111. 1966.
- _____. *The information content of annual earnings announcements*. Journal of Accounting Research, vol.6, Empirical Research in Accounting: Selected Studies 1968, p.67-92. 1968.
- BENSTON, G. J. *Published corporate accounting data and stock prices*. Journal of Accounting Research. Empirical Research in Accounting: Selected Studies, vol. 5, p. 22-54. 1967.
- BIDDLE, G.; BOWEN, R.M.; WALLACE, J.S. *Does EVA beat earnings? Evidence on associations with stock returns and firm values*. Journal of Accounting and Economics, vol. 24 nº 3, p. 301-336. 1997.
- CORE, J. E.; GUAY, W. R.; VERDI, R. *Is accruals quality a priced risk factor?* Journal of Accounting and Economics, v. 46, n. 1, p. 2-22, 2008.
- CPC – COMITÊ DE PRONUNCIAMENTOS CONTÁBEIS. CPC 00 (R1) – Estrutura Conceitual para Elaboração e Divulgação de Relatório Contábil-Financeiro. Disponível em: <<http://www.cpc.org.br/CPC/Documentos-Emitidos/Pronunciamentos/Pronunciamento?Id=80>>, divulgado em 15/12/2011.
- DAVIDSON, R.; MACKINNON, J. G. *Estimation and inference in econometrics*. New York: Oxford University Press. 1993.
- DECHOW, P. M.; HUTTON, A. P.; SLOAN, R. G. *An empirical assessment of the residual income valuation model*. Journal of Accounting and Economics, v. 26, n. 1-3, p. 1-34, 1999.
- EBAID, I. E. *The value relevance of accounting-based performance measures in emerging economies: the case of Egypt*. Management Research Review, vol. 35, nº 1, pp. 69-88. 2012.

FAMA, E.F.; FISHER, L.; JENSEN, M. C.; ROLL, R. *The adjustment of stock prices to new information*. International Economic Review, vol. 10, nº1, p. 1-21. 1969.

FAMA, E. F. *Efficient capital markets: a review of theory and empirical work*. Journal of Finance, vol. 25, nº 2, p. 383-417. 1970.

FASB – FINANCIAL ACCOUNTING STANDARDS BOARD. Exposure Draft. Disponível em: <https://www.fasb.org/jsp/FASB/Document_C/DocumentPage?cid=1176168366904&acceptedDisclaimer=true>, divulgado em 11/08/2016.

FRANCIS, J.; SCHIPPER, K.; VINCENT, L. *The relative and incremental explanatory power of earnings and alternative (to earnings) performance measure of returns*. Contemporary Accounting Research, vol. 20, nº 1, p. 121-164. 2003.

GARSON, G. D. *Logistic regression: binary & multinomial*. 2014 ed. Asheboro, NC: Statistical Associates Publishers, 2014.

GREENE, W. H. *Econometric analysis*. 8ª ed. New York: Pearson, 2018.

HE, Y.; KAMATH, R.; MEIER, H. H. *An empirical evaluation of bankruptcy prediction models for small firms: an Over-The-Counter (OTC) market experience*. Academy of Accounting and Financial Studies Journal, vol. 9, nº 1, p. 1-23. 2005.

HILL, R. C.; GRIFFITHS, W. E.; JUDGE, G. G. *Econometria*. Tradução de Alfredo Alves de Farias. 2. ed. São Paulo: Saraiva, 2003. Título original: *Undergraduate econometrics*.

HORRIGAN, J. O. *The determination of long-term credit standing with financial ratios*. Journal of Accounting Research. Empirical Research in Accounting: Selected Studies, vol.4, p.44-62. 1966.

HOSMER Jr., D. W.; LEMESHOW, S.; STURDIVANT, R. X. *Applied logistic regression*. 3ª ed. Hoboken, New Jersey: Wiley, 2013.

HOLTHAUSEN, R. W.; LARCKER, D. F. *The prediction of stock returns using financial statement information*. Journal of Accounting and Economics, v. 15, p. 373-411, 1992.

HOLTHAUSEN, R. W.; WATTS, R. L. *The relevance of the value-relevance literature for financial accounting standard setting*. Journal of Accounting and Economics, v. 31, p. 3-75, 2001.

IASB – INTERNATIONAL ACCOUNTING STANDARDS BOARD. Conceptual Framework for Financial Reporting. Disponível em: <<http://eifrs.ifrs.org/eifrs/PDFArchive?viewFile=19705>>, divulgado em março/2018.

INVESTING.COM, disponível em <https://www.investing.com/indices/>, acesso em 13/12/2019.

JOHNSEN, T.; MELICHER, R. W. *Predicting corporate bankruptcy and financial distress: information value added by multinomial logit models*. Journal of Economics and Business, vol. 46, p. 269-286. 1994

- KOTHARI, S. P. *Capital markets research in accounting*. Journal of Accounting and Economics, vol. 31 p. 105-231. 2001.
- KOTHARI, S. P.; SO, E.; VERDI, R. *Analysts' forecasts and asset pricing: a survey*. Annual Review of Economics, vol. 8, p. 197-219. 2016.
- LAU, A. H. L. *A five-state financial distress prediction model*. Journal of Accounting Research, vol. 25, p.127-138. 1987.
- LEV, B. *On the usefulness of earnings and earnings research: lessons and directions from two decades of empirical research*. Journal of Accounting Research, vol. 27, p.153-192. 1989.
- LONG, J. S. *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage, 1997.
- LONG, J. S.; FREESE, J. *Regression models for categorical dependent variables using STATA*. 3. ed. College Station, TX: Stata Press, 2014.
- MSELMI, N.; LAHIANI, A.; HAMZA, T. *Financial distress prediction: the case of French small and medium-sized firms*. International Review of Financial Analysis, vol. 50, p.67-80. 2017.
- OHLSON, J. A. *Financial ratios and the probabilistic prediction of bankruptcy*. Journal of Accounting Research, vol. 18, p. 109-131. 1980.
- _____. *Earnings, book values, and dividends in equity valuation*. Contemporary Accounting Research, v. 11, p. 661-687, 1995.
- OU, J. A.; PENMAN, S. H. *Financial statement analysis and the prediction of stock returns*. Journal of Accounting and Economics, v. 11, p. 295-329. 1989
- PINDER, J. P. *Decisions analysis using multinomial logit models: mortgage portfolio valuation*. Journal of Economics and Business, v. 48, p. 67-77, 1996.
- ROSA, M. S. S.; LUSTOSA, P. R. B. *Mercado e desempenho operacional contábil de longo prazo*. BASE – Revista de Administração e Contabilidade da Unisinos, São Leopoldo-RS, v. 11, n. 1, p. 34-46, jan/mar. 2014.
- SANTOS, J.G.C.; CALÍOPE, T.S.; COELHO, A.C. *Teoria da Firma como Fundamento para formulação de Teorias Contábeis*. Revista de Educação e Pesquisa em Contabilidade – REPeC, vol. 9, nº 1, art. 6, p.101-116. 2015.
- SCOTT, W. R. *Financial accounting theory*. 7. ed. Toronto: Pearson, 2014.
- SHUMWAY, T. *Forecasting bankruptcy more accurately: a simple hazard model*. Journal of Business, vol. 74, p. 101-124. 2001.
- STATACORP. *Stata 15 Base Reference Manual*. College Station, TX: Stata Press. 2017.

THOMSON FINANCIAL. *Worldscope Database Datatype Definitions Guide*. Thomson Financial 2003-2007 , issue 6, abril. 2007. Disponível em: < <https://www-app.uniregensburg.de/Fakultaeten/WiWi/roeder/DownloadsGeneral/Datastream%20Worldscope.pdf> >, acessado em 01/12/2019.

TREIMAN, D. J. *Quantitative data analysis: doing social research to test ideas*. San Francisco: Jossey-Bass, 2009.

WARD, T. J. *A review of financial distress research methods and recommendations for future research*. Academy of Accounting and Financial Studies Journal, vol. 3, nº 1, p. 160-178. 1999.

WRDS – Wharton Research Data Services, University of Pennsylvania, disponível em <https://wrds-www.wharton.upenn.edu/>, acesso em 25/08/2018.

WULFF, J. N. *Interpreting results from the multinomial logit model: demonstrated by foreign market entry*. Organizational Research Methods, vol. 18, nº 2, p. 300-325. 2015.

APÊNDICE A – Lista das variáveis utilizadas no trabalho

Quadro 2 – Lista das variáveis utilizadas no trabalho

Símbolo	Variável	Descrição
item6105	Worldscope Permanent I.D.	Identificador de cada empresa, utilizado para ordenar o painel de dados.
year_	ano base da informação	Ano base da informação, variando de 1995 a 2017 e utilizado para ordenar o painel de dados.
item5001	Market Price – Year End	Preço da ação no final do ano civil, utilizada como base para o cálculo dos retornos normais anuais e quinquenais.
item8101	Quick Ratio	Indicador de Liquidez (Caixa & Equivalentes + Recebíveis (Líquido)) / Passivo Circulante * 100
item8106	Current Ratio	Indicador de Liquidez Ativo Circulante / Passivo Circulante * 100
item8111 (liq_1ano)	Cash & Equivalents % Total Current Assets (Variável independente do modelo anual)	Indicador de Liquidez (Caixa & Equivalentes / Passivo Circulante) * 100
item8141	Net Sales / Working Capital	Indicador de Giro do Ativo Vendas Líquidas ou Receitas / (Ativo Circulante - Passivo Circulante) * 100
item8221	Total Debt % Total Capital	Indicador de Alavancagem (Dívida de Longo Prazo + Dívida de Curto Prazo & Parcela Circulante da Dívida de Longo Prazo) / (Capital Total + Dívida de Curto Prazo & Parcela Circulante da Dívida de Longo Prazo) * 100
item8226	Long Term Debt % Common Equity	Indicador de Alavancagem Dívida de Longo Prazo / Patrimônio Líquido * 100
item8231	Total Debt % Common Equity	Indicador de Alavancagem (Dívida de Longo Prazo + Dívida de Curto Prazo & Parcela Circulante da Dívida de Longo Prazo) / Patrimônio Líquido * 100
item8236	Total Debt % Total Assets	Indicador de Alavancagem (Dívida de Curto Prazo & Parcela Circulante da Dívida de Longo Prazo + Dívida de Longo Prazo) / Ativo Total * 100
item8241 (alav_1ano)	Common Equity % Total Assets (Variável independente do modelo anual)	Indicador de Alavancagem Patrimônio Líquido / Ativo Total * 100
item8271	Working Capital % Total Capital	Indicador de Alavancagem (Ativo Circulante - Passivo Circulante) / Capital Total * 100

continua

continuação

Símbolo	Variável	Descrição
item8301	Return On Equity - Total (%)	Indicador de Lucratividade (Lucro Líquido antes de Dividendos Preferenciais - Dividendos Preferenciais Exigidos) / Média do Ano Anterior e Ano Corrente do Patrimônio Líquido * 100
item8326	Return On Assets	Indicador de Lucratividade (Lucro Líquido antes de Dividendos Preferenciais + ((Juros Pagos-Juros Capitalizados) * (1-Alíquota de Imposto))) / Média do Ano Anterior e Ano Corrente do Ativo Total * 100
item8361	Operating Income Return On Total Capital (%)	Indicador de Lucratividade Lucro Operacional / Média do Ano Anterior e Ano Corrente do Capital Total * 100
item8366	Net Margin	Indicador de Lucratividade Lucro Líquido antes de Dividendos Preferenciais / Vendas ou Receitas Líquidas * 100
item8376 (lucr_1ano)	Return On Invested Capital (Variável independente do modelo anual)	Indicador de Lucratividade (Lucro Líquido antes de Dividendos Preferenciais + ((Juros Pagos - Juros Capitalizados) * (1-Alíquota de Imposto))) / Média do Ano Anterior e Ano Corrente de (Capital Total + Ano Anterior Dívida de Curto Prazo & Parcela Circulante da Dívida de Longo Prazo) * 100
item8401 (ga_1ano)	Total Asset Turnover (Variável independente do modelo anual)	Indicador de Giro do Ativo Vendas ou Receitas Líquidas / Ativo Total
item8911	Retained Earnings % Equity	Indicador de Lucratividade Lucros Retidos / Patrimônio Líquido * 100
item15121	Total Capital % Total Assets	Indicador de Alavancagem Capital Total / Ativo Total * 100
item8105	Quick Ratio - 5 Yr Avg	Indicador de Liquidez Média aritmética dos cinco anos anteriores de Liquidez Imediata
item8110	Current Ratio - 5 Yr Avg	Indicador de Liquidez Média aritmética dos cinco anos anteriores de Liquidez Corrente
item8115 (liq_5anos)	Cash & Short Term Investments % Total Current Liabilities - 5 Yr Avg. (Variável independente do modelo quinquenal)	Indicador de Liquidez Média aritmética dos cinco anos anteriores de Caixa & Equivalentes / Passivo Circulante * 100
item8145	Net Sales / Working Capital - 5 Yr Avg	Indicador de Giro do Ativo Média aritmética dos cinco anos anteriores de Vendas ou Receitas Líquidas / Capital de Giro

Continua

continuação

Símbolo	Variável	Descrição
item8225	Total Debt % Total Capital - 5 Yr Avg	Indicador de Alavancagem Média aritmética dos cinco anos anteriores de Dívida Total % Total Capital
item8230	Long Term Debt % Common Equity - 5 Yr Avg	Indicador de Alavancagem Média aritmética dos cinco anos anteriores de Dívida de Longo Prazo % Patrimônio Líquido
item8235	Total Debt % Common Equity - 5 Yr Avg	Indicador de Alavancagem Média aritmética dos cinco anos anteriores de Dívida Total % Patrimônio Líquido
item8240	Total Debt % Total Assets - 5 Yr Avg	Indicador de Alavancagem Média aritmética dos cinco anos anteriores de Dívida Total % Ativo Total
item8245 (alav_5anos)	Common Equity % Total Assets - 5 Yr Avg. (Variável independente do modelo quinquenal)	Indicador de Alavancagem Média aritmética dos cinco anos anteriores de Patrimônio Líquido % Ativo Total
item8275	Working Capital % Total Capital - 5 Yr Avg	Indicador de Alavancagem Média aritmética dos cinco anos anteriores de Capital de Giro % Capital Total
item8305	Return On Equity - Total (%) - 5 Yr Avg	Indicador de Lucratividade Média aritmética dos cinco anos anteriores de Retorno sobre Patrimônio Líquido
item8330	Return On Assets - 5 Yr Avg	Indicador de Lucratividade Média aritmética dos cinco anos anteriores de Retorno sobre Ativos
item8365	Operating Income Return On Total Capital (%) - 5 Yr Avg	Indicador de Lucratividade Média aritmética dos cinco anos anteriores de Lucro Operacional/ Total Capital
item8370	Net Margin - 5 Yr Avg	Indicador de Lucratividade Média aritmética dos cinco anos anteriores de Margem Líquida
item8380 (lucr_5anos)	Return On Invested Capital - 5 Yr Avg. (Variável independente do modelo quinquenal)	Indicador de Lucratividade Média aritmética dos cinco anos anteriores de Retorno sobre o Capital Investido
item8405 (ga_5anos)	Total Asset Turnover - 5 Yr Avg. (Variável independente do modelo quinquenal)	Indicador de Giro do Ativo Média aritmética dos cinco anos anteriores de Giro do Ativo Total
item8915	Retained Earnings % Equity - 5 Yr Avg	Indicador de Lucratividade Média aritmética dos cinco anos anteriores Lucros Retidos % Patrimônio Líquido.

Continua

Conclusão

Símbolo	Variável	Descrição
item15124	Total Capital % Total Assets - 5 Yr Avg	Indicador de Alavancagem Média aritmética dos cinco anos anteriores de Capital Total % Ativo Total
item6026	Nation	Representa o país no qual a empresa está domiciliada.

Fonte: elaboração própria.

APÊNDICE B – Do-file – Amostra 18 países

* Project: Minha Tese - última versão

* Last version: 22/02/2020

* Author: Sérgio Carlos dos Santos

*1) Limpa a memória e cria o arquivo log

clear all

log using "C:\Users\sergi\Documents\Sergio\Doutorado\Projeto_Doutorado\2018_1\Base de dados\Thomson Reuters\36c12743dc84c86b_dta\minha_tese.log", text replace

*2) Carrega arquivo base de dados

use "C:\Users\sergi\Documents\Sergio\Doutorado\Projeto_Doutorado\2018_1\Base de dados\Thomson Reuters\36c12743dc84c86b_dta\BASE_última versão_22022020.dta"

*3) Classifica anos do painel por código da empresa (WORLDSCOPE PERM I.D.)

bysort item6105 year_ : gen year=year_[_n-1]

*4) Deleta a variável year criada com missing values durante a classificação dos anos do painel

drop year

*5) Verifica observações (linhas) duplicadas por empresa antes de definir o painel por código da empresa e ano

quietly by item6105 year_ : gen dup=cond(_N==1,0,_n)

tab dup

*6) Deleta observações duplicadas

drop if dup>0

tab dup

*7) Deleta a variável dup criada para checar observações duplicadas

drop dup

*8) Cria a variável id (chave primária a partir da variável item6105 - WORLDSCOPE PERM I.D.) transformando em número a string item6105

encode item6105, gen (id)

*9) Define o painel de dados por código da empresa e ano

xtset id year_

*10) Calcula retornos de mercado de cada empresa, anual e quinquenal, criando as variáveis ret1a e ret5a

by id: gen ret1a=(item5001-item5001[_n-1])/item5001[_n-1]

by id: gen ret5a=(item5001-item5001[_n-5])/item5001[_n-5]

*11) Classifica os retornos anuais e quinquenais em tercís, de toda a base, criando as variáveis tret1a e tret5a, e checa se a divisão dos retornos por tercís estão igualmente distribuídos

xtile tret1a = ret1a, nq(3)

tab tret1a

xtile tret5a = ret5a, nq(3)

tab tret5a

*12) Classifica em tercís os retornos anuais em cada ano da base, dos anos de 1996 a 2017, criando as variáveis `tret1aYYYY` de cada ano, e checa se a divisão dos retornos por tercís está igualmente distribuída

```
xtile tret1a1996 = ret1a if year_==1996, nq(3)
```

```
tab tret1a1996
```

```
xtile tret1a1997 = ret1a if year_==1997, nq(3)
```

```
tab tret1a1997
```

```
xtile tret1a1998 = ret1a if year_==1998, nq(3)
```

```
tab tret1a1998
```

```
xtile tret1a1999 = ret1a if year_==1999, nq(3)
```

```
tab tret1a1999
```

```
xtile tret1a2000 = ret1a if year_==2000, nq(3)
```

```
tab tret1a2000
```

```
xtile tret1a2001 = ret1a if year_==2001, nq(3)
```

```
tab tret1a2001
```

```
xtile tret1a2002 = ret1a if year_==2002, nq(3)
```

```
tab tret1a2002
```

```
xtile tret1a2003 = ret1a if year_==2003, nq(3)
```

```
tab tret1a2003
```

```
xtile tret1a2004 = ret1a if year_==2004, nq(3)
```

```
tab tret1a2004
```

```
xtile tret1a2005 = ret1a if year_==2005, nq(3)
```

```
tab tret1a2005
```

```
xtile tret1a2006 = ret1a if year_==2006, nq(3)
```

```
tab tret1a2006
```

```
xtile tret1a2007 = ret1a if year_==2007, nq(3)
```

```
tab tret1a2007
```

```
xtile tret1a2008 = ret1a if year_==2008, nq(3)
```

```
tab tret1a2008
```

```
xtile tret1a2009 = ret1a if year_==2009, nq(3)
```

```
tab tret1a2009
```

```
xtile tret1a2010 = ret1a if year_==2010, nq(3)
```

```
tab tret1a2010
```

```
xtile tret1a2011 = ret1a if year_==2011, nq(3)
```

```
tab tret1a2011
```

```
xtile tret1a2012 = ret1a if year_==2012, nq(3)
```

```
tab tret1a2012
```

```
xtile tret1a2013 = ret1a if year_==2013, nq(3)
```

```
tab tret1a2013
```

```
xtile tret1a2014 = ret1a if year_==2014, nq(3)
```

```
tab tret1a2014
```

```
xtile tret1a2015 = ret1a if year_==2015, nq(3)
```

```
tab tret1a2015
```

```
xtile tret1a2016 = ret1a if year_==2016, nq(3)
```

```
tab tret1a2016
```

```
xtile tret1a2017 = ret1a if year_==2017, nq(3)
```

```
tab tret1a2017
```

*13) Classifica em tercís os retornos quinquenais dos anos de 2000 a 2017, criando as variáveis `tret5aYYYY` de cada ano, e checa se a divisão dos retornos por tercís está igualmente distribuída

```

xtile tret5a2000 = ret5a if year_==2000, nq(3)
tab tret5a2000
xtile tret5a2001 = ret5a if year_==2001, nq(3)
tab tret5a2001
xtile tret5a2002 = ret5a if year_==2002, nq(3)
tab tret5a2002
xtile tret5a2003 = ret5a if year_==2003, nq(3)
tab tret5a2003
xtile tret5a2004 = ret5a if year_==2004, nq(3)
tab tret5a2004
xtile tret5a2005 = ret5a if year_==2005, nq(3)
tab tret5a2005
xtile tret5a2006 = ret5a if year_==2006, nq(3)
tab tret5a2006
xtile tret5a2007 = ret5a if year_==2007, nq(3)
tab tret5a2007
xtile tret5a2008 = ret5a if year_==2008, nq(3)
tab tret5a2008
xtile tret5a2009 = ret5a if year_==2009, nq(3)
tab tret5a2009
xtile tret5a2010 = ret5a if year_==2010, nq(3)
tab tret5a2010
xtile tret5a2011 = ret5a if year_==2011, nq(3)
tab tret5a2011
xtile tret5a2012 = ret5a if year_==2012, nq(3)
tab tret5a2012
xtile tret5a2013 = ret5a if year_==2013, nq(3)
tab tret5a2013
xtile tret5a2014 = ret5a if year_==2014, nq(3)
tab tret5a2014
xtile tret5a2015 = ret5a if year_==2015, nq(3)
tab tret5a2015
xtile tret5a2016 = ret5a if year_==2016, nq(3)
tab tret5a2016
xtile tret5a2017 = ret5a if year_==2017, nq(3)
tab tret5a2017

```

*14) Calcula e cria as variáveis das médias aritméticas de 5 anos das variáveis macroeconômicas e índice de mercado

by id: gen PIB5a = (PIB[_n]+PIB[_n-1]+PIB[_n-2]+PIB[_n-3]+PIB[_n-4])/5

by id: gen inflação5a = (inflação[_n]+inflação[_n-1]+inflação[_n-2]+inflação[_n-3]+inflação[_n-4])/5

by id: gen juros5a = (juros[_n]+juros[_n-1]+juros[_n-2]+juros[_n-3]+juros[_n-4])/5

by id: gen ind_mercado5a = (ind_mercado[_n]+ind_mercado[_n-1]+ind_mercado[_n-2]+ind_mercado[_n-3]+ind_mercado[_n-4])/5

*15) Verifica a correlação entre os retornos anuais e quinquenais de cada empresa e as variáveis macroeconômicas e índice de mercado anuais e quinquenais (média 5 anos)

correlate ret1a PIB inflação juros ind_mercado

correlate ret5a PIB5a inflação5a juros5a ind_mercado5a

*16) Verifica a correlação entre os retornos anuais e quinquenais de cada empresa e os indicadores contábeis disponíveis na base Thomson Reuters em ambos os formatos: anuais e quinquenais (média 5 anos)

correlate ret1a alav_1ano item8141 item15121 item8236 item8221 item8271 lucr_1ano item8226

correlate ret5a alav_5anos item8145 item15124 item8240 item8225 item8275 lucr_5anos item8230

correlate ret1a ga_1ano item8231 item8326 item8106 item8361 item8301 item8366 item8101 liq_1ano item8911

correlate ret5a ga_5anos item8235 item8330 item8110 item8365 item8305 item8370 item8105 liq_5anos item8915

*17) Verifica a correlação entre os retornos anuais e quinquenais de cada empresa e todos os indicadores testados nos passos 15 e 16

correlate ret1a PIB inflação juros ind_mercado alav_1ano item8141 item15121 item8236 item8221 item8271 lucr_1ano item8226 ga_1ano item8231 item8326 item8106 item8361 item8301 item8366 item8101 liq_1ano item8911

correlate ret5a PIB5a inflação5a juros5a ind_mercado5a alav_5anos item8145 item15124 item8240 item8225 item8275 lucr_5anos item8230 ga_5anos item8235 item8330 item8110 item8365 item8305 item8370 item8105 liq_5anos item8915

*18) Verifica a correlação entre os retornos anuais e quinquenais de cada empresa e todos os indicadores que compõem os modelos logit

correlate ret1a alav_1ano ga_1ano liq_1ano lucr_1ano PIB inflação juros ind_mercado
correlate ret5a alav_5anos ga_5anos liq_5anos lucr_5anos PIB5a inflação5a juros5a ind_mercado5a

** Retorno Anual

*19) Usa dentre as variáveis analisadas no passo 17, as variáveis de maior correlação com retorno anual (sendo apenas um indicador contábil de Giro do Ativo, Liquidez, Lucratividade e Alavancagem; as variáveis macroeconômicas; e o índice de mercado de capitais do país),
* para rodar o logit multinomial (mlogit), 1º tercil como base e 15 iterações para forçar o logit a convergir para uma determinada função de verossimilhança, caso a convergência não seja normalmente alcançada

* utiliza como variável dependente tret1a_ano_a_ano, criada a seguir:

gen tret1a_ano_a_ano=.

replace tret1a_ano_a_ano=tret1a1996 if year_==1996

replace tret1a_ano_a_ano=tret1a1997 if year_==1997

replace tret1a_ano_a_ano=tret1a1998 if year_==1998

replace tret1a_ano_a_ano=tret1a1999 if year_==1999

replace tret1a_ano_a_ano=tret1a2000 if year_==2000

replace tret1a_ano_a_ano=tret1a2001 if year_==2001

replace tret1a_ano_a_ano=tret1a2002 if year_==2002

replace tret1a_ano_a_ano=tret1a2003 if year_==2003

replace tret1a_ano_a_ano=tret1a2004 if year_==2004

replace tret1a_ano_a_ano=tret1a2005 if year_==2005

replace tret1a_ano_a_ano=tret1a2006 if year_==2006

replace tret1a_ano_a_ano=tret1a2007 if year_==2007

replace tret1a_ano_a_ano=tret1a2008 if year_==2008

replace tret1a_ano_a_ano=tret1a2009 if year_==2009

replace tret1a_ano_a_ano=tret1a2010 if year_==2010

```

replace tret1a_ano_a_ano=tret1a2011 if year_==2011
replace tret1a_ano_a_ano=tret1a2012 if year_==2012
replace tret1a_ano_a_ano=tret1a2013 if year_==2013
replace tret1a_ano_a_ano=tret1a2014 if year_==2014
replace tret1a_ano_a_ano=tret1a2015 if year_==2015
replace tret1a_ano_a_ano=tret1a2016 if year_==2016
replace tret1a_ano_a_ano=tret1a2017 if year_==2017
mlogit tret1a_ano_a_ano alav_1ano ga_1ano liq_1ano lucr_1ano PIB inflação juros
ind_mercado if inrange(year_,1995,2012), base(1) iterate(15)
*20) Prediz a probabilidade de ser categoria 1 (baixo retorno) para os retornos anuais nos anos
2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do modelo rodado no passo 18,
criando-se as variáveis pla_1_2013, pla_1_2014, pla_1_2015, pla_1_2016 e pla_1_2017
predict pla_1_2013 if year_==2013, outcome(1) p
predict pla_1_2014 if year_==2014, outcome(1) p
predict pla_1_2015 if year_==2015, outcome(1) p
predict pla_1_2016 if year_==2016, outcome(1) p
predict pla_1_2017 if year_==2017, outcome(1) p
*21) Detalha a distribuição das variáveis criadas para mostrar a probabilidade de ser categoria
1 (baixo retorno)
sum pla_1_2013, detail
sum pla_1_2014, detail
sum pla_1_2015, detail
sum pla_1_2016, detail
sum pla_1_2017, detail
*22) Prediz a probabilidade de ser categoria 2 (médio retorno) para os retornos anuais nos
anos 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do modelo rodado no passo 18,
criando-se as variáveis pla_2_2013, pla_2_2014, pla_2_2015, pla_2_2016 e pla_2_2017
predict pla_2_2013 if year_==2013, outcome(2) p
predict pla_2_2014 if year_==2014, outcome(2) p
predict pla_2_2015 if year_==2015, outcome(2) p
predict pla_2_2016 if year_==2016, outcome(2) p
predict pla_2_2017 if year_==2017, outcome(2) p
*23) Detalha a distribuição das variáveis criadas para mostrar a probabilidade de ser categoria
2 (médio retorno)
sum pla_2_2013, detail
sum pla_2_2014, detail
sum pla_2_2015, detail
sum pla_2_2016, detail
sum pla_2_2017, detail
*24) Prediz a probabilidade de ser categoria 3 (alto retorno) para os retornos anuais nos anos
2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do modelo rodado no passo 18,
criando-se as variáveis pla_3_2013, pla_3_2014, pla_3_2015, pla_3_2016 e pla_3_2017
predict pla_3_2013 if year_==2013, outcome(3) p
predict pla_3_2014 if year_==2014, outcome(3) p
predict pla_3_2015 if year_==2015, outcome(3) p
predict pla_3_2016 if year_==2016, outcome(3) p
predict pla_3_2017 if year_==2017, outcome(3) p
*25) Detalha a distribuição das variáveis criadas para mostrar a probabilidade de ser categoria
3 (alto retorno)
sum pla_3_2013, detail

```

```
sum pla_3_2014, detail
sum pla_3_2015, detail
sum pla_3_2016, detail
sum pla_3_2017, detail
```

*26) Cria variáveis de categorias previstas para os anos de 2013, 2014, 2015, 2016 e 2017 para cada empresa que teve dados calculados pelo modelo logit multinomial (mlogit) no passo 18, a variável assumirá o valor da categoria com maior probabilidade de ocorrer (probabilidade dominante)

```
gen pla_cat_2013=.
replace pla_cat_2013=1 if pla_1_2013>pla_2_2013 & pla_1_2013>pla_3_2013
replace pla_cat_2013=2 if pla_2_2013>pla_1_2013 & pla_2_2013>pla_3_2013
replace pla_cat_2013=3 if pla_3_2013>pla_1_2013 & pla_3_2013>pla_2_2013
gen pla_cat_2014=.
replace pla_cat_2014=1 if pla_1_2014>pla_2_2014 & pla_1_2014>pla_3_2014
replace pla_cat_2014=2 if pla_2_2014>pla_1_2014 & pla_2_2014>pla_3_2014
replace pla_cat_2014=3 if pla_3_2014>pla_1_2014 & pla_3_2014>pla_2_2014
gen pla_cat_2015=.
replace pla_cat_2015=1 if pla_1_2015>pla_2_2015 & pla_1_2015>pla_3_2015
replace pla_cat_2015=2 if pla_2_2015>pla_1_2015 & pla_2_2015>pla_3_2015
replace pla_cat_2015=3 if pla_3_2015>pla_1_2015 & pla_3_2015>pla_2_2015
gen pla_cat_2016=.
replace pla_cat_2016=1 if pla_1_2016>pla_2_2016 & pla_1_2016>pla_3_2016
replace pla_cat_2016=2 if pla_2_2016>pla_1_2016 & pla_2_2016>pla_3_2016
replace pla_cat_2016=3 if pla_3_2016>pla_1_2016 & pla_3_2016>pla_2_2016
gen pla_cat_2017=.
replace pla_cat_2017=1 if pla_1_2017>pla_2_2017 & pla_1_2017>pla_3_2017
replace pla_cat_2017=2 if pla_2_2017>pla_1_2017 & pla_2_2017>pla_3_2017
replace pla_cat_2017=3 if pla_3_2017>pla_1_2017 & pla_3_2017>pla_2_2017
```

*27) Verifica a distribuição das variáveis (valores previstos): pla_cat_2013, pla_cat_2014, pla_cat_2015, pla_cat_2016 e pla_cat_2017

```
tab pla_cat_2013
tab pla_cat_2014
tab pla_cat_2015
tab pla_cat_2016
tab pla_cat_2017
```

*28) Compara os resultados previstos e realizados. Observar, na matriz gerada, a quantidade da diagonal principal em relação à quantidade total:

```
tab tret1a2013 pla_cat_2013
tab tret1a2014 pla_cat_2014
tab tret1a2015 pla_cat_2015
tab tret1a2016 pla_cat_2016
tab tret1a2017 pla_cat_2017
```

*29) Cria a variável prev_corr1a para as previsões dos anos 2013, 2014, 2015, 2016 e 2017 para mostrar a categoria prevista corretamente (valor previsto pla_cat_20YY é igual ao valor realizado tret1a20YY)

```
gen prev_corr1a2013=pla_cat_2013 if year==2013 & tret1a2013==pla_cat_2013
gen prev_corr1a2014=pla_cat_2014 if year==2014 & tret1a2014==pla_cat_2014
gen prev_corr1a2015=pla_cat_2015 if year==2015 & tret1a2015==pla_cat_2015
gen prev_corr1a2016=pla_cat_2016 if year==2016 & tret1a2016==pla_cat_2016
gen prev_corr1a2017=pla_cat_2017 if year==2017 & tret1a2017==pla_cat_2017
```

*30) Verifica a distribuição das variáveis (prev_corr1a): anos 2013, 2014, 2015, 2016 e 2017

```
tab prev_corr1a2013
```

```
tab prev_corr1a2014
```

```
tab prev_corr1a2015
```

```
tab prev_corr1a2016
```

```
tab prev_corr1a2017
```

*31) Cria a variável pr_acerto1a2013 e reporta nessa variável a probabilidade informada no predict, conforme passos 19, 21 e 23 anteriores

```
gen pr_acerto1a2013=.
```

```
replace pr_acerto1a2013=p1a_1_2013 if p1a_1_2013>p1a_2_2013 &  
p1a_1_2013>p1a_3_2013
```

```
replace pr_acerto1a2013=p1a_2_2013 if p1a_2_2013>p1a_1_2013 &  
p1a_2_2013>p1a_3_2013
```

```
replace pr_acerto1a2013=p1a_3_2013 if p1a_3_2013>p1a_1_2013 &  
p1a_3_2013>p1a_2_2013
```

*32) Cria a variável pr_acerto1a2014 e reporta nessa variável a probabilidade informada no predict, conforme passos 19, 21 e 23 anteriores

```
gen pr_acerto1a2014=.
```

```
replace pr_acerto1a2014=p1a_1_2014 if p1a_1_2014>p1a_2_2014 &  
p1a_1_2014>p1a_3_2014
```

```
replace pr_acerto1a2014=p1a_2_2014 if p1a_2_2014>p1a_1_2014 &  
p1a_2_2014>p1a_3_2014
```

```
replace pr_acerto1a2014=p1a_3_2014 if p1a_3_2014>p1a_1_2014 &  
p1a_3_2014>p1a_2_2014
```

*33) Cria a variável pr_acerto1a2015 e reporta nessa variável a probabilidade informada no predict, conforme passos 19, 21 e 23 anteriores

```
gen pr_acerto1a2015=.
```

```
replace pr_acerto1a2015=p1a_1_2015 if p1a_1_2015>p1a_2_2015 &  
p1a_1_2015>p1a_3_2015
```

```
replace pr_acerto1a2015=p1a_2_2015 if p1a_2_2015>p1a_1_2015 &  
p1a_2_2015>p1a_3_2015
```

```
replace pr_acerto1a2015=p1a_3_2015 if p1a_3_2015>p1a_1_2015 &  
p1a_3_2015>p1a_2_2015
```

*34) Cria a variável pr_acerto1a2016 e reporta nessa variável a probabilidade informada no predict, conforme passos 19, 21 e 23 anteriores

```
gen pr_acerto1a2016=.
```

```
replace pr_acerto1a2016=p1a_1_2016 if p1a_1_2016>p1a_2_2016 &  
p1a_1_2016>p1a_3_2016
```

```
replace pr_acerto1a2016=p1a_2_2016 if p1a_2_2016>p1a_1_2016 &  
p1a_2_2016>p1a_3_2016
```

```
replace pr_acerto1a2016=p1a_3_2016 if p1a_3_2016>p1a_1_2016 &  
p1a_3_2016>p1a_2_2016
```

*35) Cria a variável pr_acerto1a2017 e reporta nessa variável a probabilidade informada no predict, conforme passos 19, 21 e 23 anteriores

```
gen pr_acerto1a2017=.
```

```
replace pr_acerto1a2017=p1a_1_2017 if p1a_1_2017>p1a_2_2017 &  
p1a_1_2017>p1a_3_2017
```

```
replace pr_acerto1a2017=p1a_2_2017 if p1a_2_2017>p1a_1_2017 &  
p1a_2_2017>p1a_3_2017
```

```
replace pr_acerto1a2017=pla_3_2017 if pla_3_2017>pla_1_2017 &
pla_3_2017>pla_2_2017
```

*36) Detalha a distribuição das probabilidades dos acertos

```
sum pr_acerto1a2013, detail
```

```
sum pr_acerto1a2014, detail
```

```
sum pr_acerto1a2015, detail
```

```
sum pr_acerto1a2016, detail
```

```
sum pr_acerto1a2017, detail
```

```
*****
```

```
** Retorno Quinquenal
```

```
*****
```

*37) Usa dentre as variáveis analisadas no passo 17, as variáveis quinquenais recíprocas às de maior correlação com retorno anual (utilizadas no passo 18 e sendo apenas um indicador contábil de Giro do Ativo, Liquidez, Lucratividade e Alavancagem; as variáveis macroeconômicas; e o índice de mercado de capitais do país),

* para rodar o logit multinomial (mlogit), 1º tercil como base e 15 iterações para forçar o logit a convergir para uma determinada função de verossimilhança, caso a convergência não seja normalmente alcançada

* utiliza como variável dependente tret5a_ano_a_ano, criada a seguir:

```
gen tret5a_ano_a_ano=.
```

```
replace tret5a_ano_a_ano=tret5a2000 if year_==2000
```

```
replace tret5a_ano_a_ano=tret5a2001 if year_==2001
```

```
replace tret5a_ano_a_ano=tret5a2002 if year_==2002
```

```
replace tret5a_ano_a_ano=tret5a2003 if year_==2003
```

```
replace tret5a_ano_a_ano=tret5a2004 if year_==2004
```

```
replace tret5a_ano_a_ano=tret5a2005 if year_==2005
```

```
replace tret5a_ano_a_ano=tret5a2006 if year_==2006
```

```
replace tret5a_ano_a_ano=tret5a2007 if year_==2007
```

```
replace tret5a_ano_a_ano=tret5a2008 if year_==2008
```

```
replace tret5a_ano_a_ano=tret5a2009 if year_==2009
```

```
replace tret5a_ano_a_ano=tret5a2010 if year_==2010
```

```
replace tret5a_ano_a_ano=tret5a2011 if year_==2011
```

```
replace tret5a_ano_a_ano=tret5a2012 if year_==2012
```

```
replace tret5a_ano_a_ano=tret5a2013 if year_==2013
```

```
replace tret5a_ano_a_ano=tret5a2014 if year_==2014
```

```
replace tret5a_ano_a_ano=tret5a2015 if year_==2015
```

```
replace tret5a_ano_a_ano=tret5a2016 if year_==2016
```

```
replace tret5a_ano_a_ano=tret5a2017 if year_==2017
```

```
mlogit tret5a_ano_a_ano alav_5anos ga_5anos liq_5anos lucr_5anos PIB5a inflação5a
```

```
juros5a ind_mercado5a if inrange(year_,1995,2012), base(1) iterate(15)
```

*38) Prediz a probabilidade de ser categoria 1 (baixo retorno) para os retornos quinquenais nos anos 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do modelo rodado no passo 36, criando-se as variáveis p5a_1_YYYY para cada um dos 5 anos

```
predict p5a_1_2013 if year_==2013, outcome(1) p
```

```
predict p5a_1_2014 if year_==2014, outcome(1) p
```

```
predict p5a_1_2015 if year_==2015, outcome(1) p
```

```
predict p5a_1_2016 if year_==2016, outcome(1) p
```

```
predict p5a_1_2017 if year_==2017, outcome(1) p
```

*39) Detalha a distribuição das variáveis criadas para mostrar a probabilidade de ser categoria 1 (baixo retorno)

```
sum p5a_1_2013, detail
sum p5a_1_2014, detail
sum p5a_1_2015, detail
sum p5a_1_2016, detail
sum p5a_1_2017, detail
```

*40) Prediz a probabilidade de ser categoria 2 (médio retorno) para os retornos anuais nos anos 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do modelo rodado no passo 36, criando-se as variáveis p5a_2_YYYY para cada um dos 5 anos

```
predict p5a_2_2013 if year_==2013, outcome(2) p
predict p5a_2_2014 if year_==2014, outcome(2) p
predict p5a_2_2015 if year_==2015, outcome(2) p
predict p5a_2_2016 if year_==2016, outcome(2) p
predict p5a_2_2017 if year_==2017, outcome(2) p
```

*41) Detalha a distribuição das variáveis criadas para mostrar a probabilidade de ser categoria 2 (médio retorno)

```
sum p5a_2_2013, detail
sum p5a_2_2014, detail
sum p5a_2_2015, detail
sum p5a_2_2016, detail
sum p5a_2_2017, detail
```

*42) Prediz a probabilidade de ser categoria 3 (alto retorno) para os retornos anuais nos anos 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do modelo rodado no passo 36, criando-se as variáveis p5a_3_YYYY para cada um dos 5 anos

```
predict p5a_3_2013 if year_==2013, outcome(3) p
predict p5a_3_2014 if year_==2014, outcome(3) p
predict p5a_3_2015 if year_==2015, outcome(3) p
predict p5a_3_2016 if year_==2016, outcome(3) p
predict p5a_3_2017 if year_==2017, outcome(3) p
```

*43) Detalha a distribuição das variáveis criadas para mostrar a probabilidade de ser categoria 3 (alto retorno)

```
sum p5a_3_2013, detail
sum p5a_3_2014, detail
sum p5a_3_2015, detail
sum p5a_3_2016, detail
sum p5a_3_2017, detail
```

*44) Cria variáveis de categorias preditas para os anos de 2013, 2014, 2015, 2016 e 2017 para cada empresa que teve dados calculados pelo modelo logit multinomial (mlogit) no passo 36, a variável assumirá o valor da categoria com maior probabilidade de ocorrer (probabilidade dominante)

```
gen p5a_cat_2013=.
replace p5a_cat_2013=1 if p5a_1_2013>p5a_2_2013 & p5a_1_2013>p5a_3_2013
replace p5a_cat_2013=2 if p5a_2_2013>p5a_1_2013 & p5a_2_2013>p5a_3_2013
replace p5a_cat_2013=3 if p5a_3_2013>p5a_1_2013 & p5a_3_2013>p5a_2_2013
gen p5a_cat_2014=.
replace p5a_cat_2014=1 if p5a_1_2014>p5a_2_2014 & p5a_1_2014>p5a_3_2014
replace p5a_cat_2014=2 if p5a_2_2014>p5a_1_2014 & p5a_2_2014>p5a_3_2014
replace p5a_cat_2014=3 if p5a_3_2014>p5a_1_2014 & p5a_3_2014>p5a_2_2014
gen p5a_cat_2015=.
replace p5a_cat_2015=1 if p5a_1_2015>p5a_2_2015 & p5a_1_2015>p5a_3_2015
replace p5a_cat_2015=2 if p5a_2_2015>p5a_1_2015 & p5a_2_2015>p5a_3_2015
```

replace p5a_cat_2015=3 if p5a_3_2015>p5a_1_2015 & p5a_3_2015>p5a_2_2015
gen p5a_cat_2016=.

replace p5a_cat_2016=1 if p5a_1_2016>p5a_2_2016 & p5a_1_2016>p5a_3_2016

replace p5a_cat_2016=2 if p5a_2_2016>p5a_1_2016 & p5a_2_2016>p5a_3_2016

replace p5a_cat_2016=3 if p5a_3_2016>p5a_1_2016 & p5a_3_2016>p5a_2_2016

gen p5a_cat_2017=.

replace p5a_cat_2017=1 if p5a_1_2017>p5a_2_2017 & p5a_1_2017>p5a_3_2017

replace p5a_cat_2017=2 if p5a_2_2017>p5a_1_2017 & p5a_2_2017>p5a_3_2017

replace p5a_cat_2017=3 if p5a_3_2017>p5a_1_2017 & p5a_3_2017>p5a_2_2017

*45) Verifica a distribuição das variáveis (valores previstos): p5a_cat_2013, p5a_cat_2014, p5a_cat_2015, p5a_cat_2016 e p5a_cat_2017

tab p5a_cat_2013

tab p5a_cat_2014

tab p5a_cat_2015

tab p5a_cat_2016

tab p5a_cat_2017

*46) Compara os resultados previstos e realizados. Observar, na matriz gerada, a quantidade da diagonal principal em relação à quantidade total:

tab tret5a2013 p5a_cat_2013

tab tret5a2014 p5a_cat_2014

tab tret5a2015 p5a_cat_2015

tab tret5a2016 p5a_cat_2016

tab tret5a2017 p5a_cat_2017

*47) Cria a variável prev_corr5a para as previsões dos anos 2013, 2014, 2015, 2016 e 2017 para mostrar a categoria prevista corretamente(valor previsto p5a_cat_20YY é igual ao valor realizado tret5a20YY)

gen prev_corr5a2013=p5a_cat_2013 if year==2013 & tret5a2013==p5a_cat_2013

gen prev_corr5a2014=p5a_cat_2014 if year==2014 & tret5a2014==p5a_cat_2014

gen prev_corr5a2015=p5a_cat_2015 if year==2015 & tret5a2015==p5a_cat_2015

gen prev_corr5a2016=p5a_cat_2016 if year==2016 & tret5a2016==p5a_cat_2016

gen prev_corr5a2017=p5a_cat_2017 if year==2017 & tret5a2017==p5a_cat_2017

*48) Verifica a distribuição das variáveis (prev_corr5a): anos 2013, 2014, 2015, 2016 e 2017

tab prev_corr5a2013

tab prev_corr5a2014

tab prev_corr5a2015

tab prev_corr5a2016

tab prev_corr5a2017

*49) Cria a variável pr_acerto5a2013 e reporta nessa variável a probabilidade informada no predict, conforme passos 37, 39 e 41 anteriores

gen pr_acerto5a2013=.

replace pr_acerto5a2013=p5a_1_2013 if p5a_1_2013>p5a_2_2013 &
p5a_1_2013>p5a_3_2013

replace pr_acerto5a2013=p5a_2_2013 if p5a_2_2013>p5a_1_2013 &
p5a_2_2013>p5a_3_2013

replace pr_acerto5a2013=p5a_3_2013 if p5a_3_2013>p5a_1_2013 &
p5a_3_2013>p5a_2_2013

*50) Cria a variável pr_acerto5a2014 e reporta nessa variável a probabilidade informada no predict, conforme passos 37, 39 e 41 anteriores

gen pr_acerto5a2014=.

```
replace pr_acerto5a2014=p5a_1_2014 if p5a_1_2014>p5a_2_2014 &
p5a_1_2014>p5a_3_2014
```

```
replace pr_acerto5a2014=p5a_2_2014 if p5a_2_2014>p5a_1_2014 &
p5a_2_2014>p5a_3_2014
```

```
replace pr_acerto5a2014=p5a_3_2014 if p5a_3_2014>p5a_1_2014 &
p5a_3_2014>p5a_2_2014
```

*51) Cria a variável pr_acerto5a2015 e reporta nessa variável a probabilidade informada no predict, conforme passos 37, 39 e 41 anteriores

```
gen pr_acerto5a2015=.
```

```
replace pr_acerto5a2015=p5a_1_2015 if p5a_1_2015>p5a_2_2015 &
p5a_1_2015>p5a_3_2015
```

```
replace pr_acerto5a2015=p5a_2_2015 if p5a_2_2015>p5a_1_2015 &
p5a_2_2015>p5a_3_2015
```

```
replace pr_acerto5a2015=p5a_3_2015 if p5a_3_2015>p5a_1_2015 &
p5a_3_2015>p5a_2_2015
```

*52) Cria a variável pr_acerto5a2016 e reporta nessa variável a probabilidade informada no predict, conforme passos 37, 39 e 41 anteriores

```
gen pr_acerto5a2016=.
```

```
replace pr_acerto5a2016=p5a_1_2016 if p5a_1_2016>p5a_2_2016 &
p5a_1_2016>p5a_3_2016
```

```
replace pr_acerto5a2016=p5a_2_2016 if p5a_2_2016>p5a_1_2016 &
p5a_2_2016>p5a_3_2016
```

```
replace pr_acerto5a2016=p5a_3_2016 if p5a_3_2016>p5a_1_2016 &
p5a_3_2016>p5a_2_2016
```

*53) Cria a variável pr_acerto5a2017 e reporta nessa variável a probabilidade informada no predict, conforme passos 37, 39 e 41 anteriores

```
gen pr_acerto5a2017=.
```

```
replace pr_acerto5a2017=p5a_1_2017 if p5a_1_2017>p5a_2_2017 &
p5a_1_2017>p5a_3_2017
```

```
replace pr_acerto5a2017=p5a_2_2017 if p5a_2_2017>p5a_1_2017 &
p5a_2_2017>p5a_3_2017
```

```
replace pr_acerto5a2017=p5a_3_2017 if p5a_3_2017>p5a_1_2017 &
p5a_3_2017>p5a_2_2017
```

*54) Detalha a distribuição das probabilidades dos acertos

```
sum pr_acerto5a2013, detail
```

```
sum pr_acerto5a2014, detail
```

```
sum pr_acerto5a2015, detail
```

```
sum pr_acerto5a2016, detail
```

```
sum pr_acerto5a2017, detail
```

```
*****
```

```
** Detalha os acertos de previsões por países
```

```
*****
```

*55) Apresenta a distribuição de observações por países constantes da base de dados

```
tab item6026
```

*56) Apresenta a distribuição de observações por ano de cada um dos 18 países constantes da base de dados

```
bysort item6026:tab year_
```

```
*****
```

```
** Retorno Anual
```

```
*****
```

*57) Apresenta a distribuição de quantidade de previsões realizadas por país e ano constantes da base de dados, na visão de retorno anual

bysort item6026:tab p1a_cat_2013

bysort item6026:tab p1a_cat_2014

bysort item6026:tab p1a_cat_2015

bysort item6026:tab p1a_cat_2016

bysort item6026:tab p1a_cat_2017

*58) Apresenta a distribuição de quantidade de previsões acertadas (tret1aYYYY = p1a_cat_YYYYY) - valor da categoria prevista pelo modelo logit multinomial igual ao valor realizado no respectivo ano e por país constantes da base de dados, na visão de retorno anual

bysort item6026:tab prev_corr1a2013

bysort item6026:tab prev_corr1a2014

bysort item6026:tab prev_corr1a2015

bysort item6026:tab prev_corr1a2016

bysort item6026:tab prev_corr1a2017

** Retorno Quinquenal

*59) Apresenta a distribuição de quantidade de previsões realizadas por país e ano constantes da base de dados, na visão de retorno quinquenal

bysort item6026:tab p5a_cat_2013

bysort item6026:tab p5a_cat_2014

bysort item6026:tab p5a_cat_2015

bysort item6026:tab p5a_cat_2016

bysort item6026:tab p5a_cat_2017

*60) Apresenta a distribuição de quantidade de previsões acertadas (tret5aYYYY = p5a_cat_YYYYY) - valor da categoria prevista pelo modelo logit multinomial igual ao valor realizado no respectivo ano e por país constantes da base de dados, na visão de retorno quinquenal

bysort item6026:tab prev_corr5a2013

bysort item6026:tab prev_corr5a2014

bysort item6026:tab prev_corr5a2015

bysort item6026:tab prev_corr5a2016

bysort item6026:tab prev_corr5a2017

*61) Fecha o arquivo log

log close

APÊNDICE C – Lista das variáveis criadas para execução dos modelos anual e quinquenal

Quadro 6 – Lista das variáveis criadas para execução dos modelos anual e quinquenal

Símbolo	Variável	Descrição
Year	Ano da informação.	Criada para <i>missing values</i> durante a classificação dos anos do painel e depois excluída.
Dup	Observações duplicadas.	Criada para identificar observações duplicadas e depois excluída.
Id	Identificador de empresa.	Criada a partir da variável item6105, define o painel de dados em conjunto com a variável year .
ret1a	Retorno anual de mercado da ação de cada empresa.	Retorno anual de mercado de cada empresa, calculado a partir da variável item5001: (item5001-item5001[n-1])/item5001[n-1].
ret5a	Retorno quinquenal de mercado da ação de cada empresa.	Retorno quinquenal de mercado de cada empresa, calculado a partir da variável item5001: (item5001-item5001[n-5])/item5001[n-5].
tret1a	Tercil do retorno anual.	Classificação em tercils de toda a base, a partir da variável ret1a.
tret5a	Tercil do retorno quinquenal.	Classificação em tercils de toda a base, a partir da variável ret5a.
tret1aYYYY	Tercil do retorno anual no ano YYYY.	Classificação em tercils dos retornos anuais em cada ano da base, onde YYYY varia do ano 1996 ao 2017.
tret5aYYYY	Tercil do retorno quinquenal no ano YYYY.	Classificação em tercils dos retornos quinquenais em cada ano da base, onde YYYY varia do ano 2000 a 2017.
tret1a_ano_a_ano	Categoria indicada pelo tercil dos resultados anuais. (Variável dependente do modelo anual)	Criada a partir das informações acumuladas ano a ano das variáveis tret1aYYYY.
pla_1_YYYY	Probabilidade de ser categoria 1 no ano YYYY, de acordo com modelo de predição, em retorno anual.	Variáveis que armazenam a predição da probabilidade de ser categoria 1 (baixo retorno) para os retornos anuais nos anos de 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do MLM cuja variável dependente reporta-se às categorias de retornos anuais (1 - baixo, 2 - médio e 3 - alto).

continua

continuação

Símbolo	Variável	Descrição
p1a_2_YYYY	Probabilidade de ser categoria 2 no ano YYYY, de acordo com modelo de predição, em retorno anual.	Variáveis que armazenam a predição da probabilidade de ser categoria 2 (médio retorno) para os retornos anuais nos anos de 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do MLM cuja variável dependente reporta-se às categorias de retornos anuais (1 - baixo, 2 - médio e 3 - alto).
p1a_3_YYYY	Probabilidade de ser categoria 3 no ano YYYY, de acordo com modelo de predição, em retorno anual.	Variáveis que armazenam a predição da probabilidade de ser categoria 3 (alto retorno) para os retornos anuais nos anos de 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do MLM cuja variável dependente reporta-se às categorias de retornos anuais (1 - baixo, 2 - médio e 3 - alto).
p1a_cat_YYYY	Categoria predita pelo modelo - retorno anual.	Variáveis que armazenam a categoria predita pelo MLM. A variável assumirá o valor da categoria com maior probabilidade de ocorrer (probabilidade dominante), podendo variar em três categorias: 1 - baixo retorno, 2 - médio retorno e 3 - alto retorno; para os retornos anuais dos anos de 2013, 2014, 2015, 2016 e 2017.
prev_corr1aYYYY	Categoria corretamente predita pelo modelo (acerto) - retorno anual.	Variáveis que armazenam a categoria de retorno anual corretamente predita ($tret1aYYYY = p1a_cat_YYYY$) - valor da categoria prevista pelo MLM igual ao valor realizado no respectivo ano que varia de 2013 a 2017.
pr_acerto1aYYYY	Probabilidade da ocorrência da categoria predita pelo modelo - retorno anual.	Variáveis que armazenam a probabilidade de ocorrência da categoria predita pelo MLM cuja variável dependente reporta-se às categorias de retornos anuais, para os anos de 2013, 2014, 2015, 2016 e 2017.
tret5a_ano_a_ano	Categoria indicada pelo tercil dos resultados quinquenais. (Variável dependente do modelo quinquenal)	Criada a partir das informações acumuladas ano a ano das variáveis tret5aYYYY.

continua

conclusão

Símbolo	Variável	Descrição
p5a_1_YYYY	Probabilidade de ser categoria 1 no ano YYYY, de acordo com modelo de predição, em retorno quinquenal.	Variáveis que armazenam a predição da probabilidade de ser categoria 1 (baixo retorno) para os retornos quinquenais nos anos de 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do MLM cuja variável dependente reporta-se às categorias de retornos anuais (1 - baixo, 2 - médio e 3 - alto).
p5a_2_YYYY	Probabilidade de ser categoria 2 no ano YYYY, de acordo com modelo de predição, em retorno quinquenal.	Variáveis que armazenam a predição da probabilidade de ser categoria 2 (médio retorno) para os retornos quinquenais nos anos de 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do MLM cuja variável dependente reporta-se às categorias de retornos anuais (1 - baixo, 2 - médio e 3 - alto).
p5a_3_YYYY	Probabilidade de ser categoria 3 no ano YYYY, de acordo com modelo de predição, em retorno quinquenal.	Variáveis que armazenam a predição da probabilidade de ser categoria 3 (alto retorno) para os retornos quinquenais nos anos de 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do MLM cuja variável dependente reporta-se às categorias de retornos anuais (1 - baixo, 2 - médio e 3 - alto).
p5a_cat_YYYY	Categoria predita pelo modelo - retorno quinquenal.	Variáveis que armazenam a categoria predita pelo MLM. A variável assumirá o valor da categoria com maior probabilidade de ocorrer (probabilidade dominante), podendo variar em três categorias: 1 - baixo retorno, 2 - médio retorno e 3 - alto retorno; para os retornos quinquenais dos anos de 2013, 2014, 2015, 2016 e 2017.
prev_corr5aYYYY	Categoria corretamente predita pelo modelo (acerto) - retorno quinquenal.	Variáveis que armazenam a categoria de retorno quinquenal corretamente predita ($tret5aYYYY = p5a_cat_YYYY$) - valor da categoria prevista pelo MLM igual ao valor realizado no respectivo ano que varia de 2013 a 2017.
pr_acerto5aYYYY	Probabilidade da ocorrência da categoria predita pelo modelo - retorno quinquenal.	Variáveis que armazenam a probabilidade de ocorrência da categoria predita pelo MLM cuja variável dependente reporta-se às categorias de retornos quinquenais, para os anos de 2013, 2014, 2015, 2016 e 2017.

Fonte: elaboração própria.

APÊNDICE D – Estatísticas descritivas suplementares – MLM curto prazo

Tabela 7 – Curto Prazo – amostra utilizada na estimação (1996 a 2012) – estatística descritiva

Variáveis	Estatísticas Descritivas				
	Obs.	Média	Desvio Padrão	Mínimo	Máximo
Curto prazo					
alav_1ano	361.241	-6587,363	1.383.165,00	-579.000.000	52.021
ga_1ano	360.891	2	221	-7	127.843
liq_1ano	354.915	35	35	-304	10.666
lucr_1ano	330.908	-61	6.602	-2.206.967	124.863
PIB	362.268	3	3	-9	15
Inflação	362.268	2	3	-1	86
Juros	362.268	6	7	0	86
ind_mercado	348.732	5.796	7.753	45	69.305

Fonte: elaboração própria.

Tabela 8 – Curto Prazo – amostra utilizada na estimação (1996 a 2012) – distribuição em percentis

Variáveis	Percentis									
	Obs.	1%	5%	10%	25%	50%	75%	90%	95%	99%
Curto prazo										
alav_1ano	361.241	-1.555	-47	8	29	48	70	86	93	99
ga_1ano	360.891	0	0	0	0	1	1	2	3	5
liq_1ano	354.915	0	1	3	11	27	55	85	96	100
lucr_1ano	330.908	-452	-95	-43	-5	4	11	19	27	60
PIB	362.268	-5	-2	-0	2	3	4	8	10	13
Inflação	362.268	-1	-1	-0	1	2	3	5	7	12
Juros	362.268	0	0	1	2	5	7	9	12	45
ind_mercado	348.732	303	877	970	1.258	2.715	8.579	13.833	17.226	39.986

Fonte: elaboração própria.

Tabela 9 – Curto Prazo – amostra utilizada na previsão (2013 a 2017) – estatística descritiva

Variáveis	Estatísticas Descritivas				
	Obs.	Média	Desvio Padrão	Mínimo	Máximo
Curto prazo					
alav_1ano	127.009	-13552,75	3.003.708,00	-1.050.000.000	4.127.716
ga_1ano	126.915	2	225	-2	80.067
liq_1ano	126.424	39	30	-1.084	1.033
lucr_1ano	118.971	-47	2.255	-456.600	43.372
PIB	127.279	3	2	-4	8
Inflação	127.279	2	2	-2	16
Juros	127.279	4	5	0	52
ind_mercado	127.279	8.844	11.124	424	76.402

Fonte: elaboração própria.

Tabela 10 – Curto Prazo – amostra utilizada na previsão (2013 a 2017) – distribuição em percentis

Variáveis	Percentis									
	Obs.	1%	5%	10%	25%	50%	75%	90%	95%	99%
Curto prazo										
alav_1ano	127.009	-2.952	-56	9	31	52	73	88	94	99
ga_1ano	126.915	0	0	0	0	1	1	2	2	4
liq_1ano	126.424	0	2	5	14	32	60	88	96	100
lucr_1ano	118.971	-469	-103	-45	-4	4	9	16	23	53
PIB	127.279	-0	1	1	2	3	5	7	8	8
Inflação	127.279	-1	-0	0	1	2	2	5	6	11
Juros	127.279	0	0	0	1	3	5	10	10	32
ind_mercado	127.279	505	1.797	1.916	2.116	3.539	13.010	19.034	22.765	59.505

Fonte: elaboração própria.

APÊNDICE E – Estatísticas descritivas suplementares – MLM longo prazo

Tabela 11 – Longo prazo – amostra utilizada na estimação (2000 a 2012) – estatística descritiva

Variáveis	Estatísticas Descritivas				
	Obs.	Média	Desvio Padrão	Mínimo	Máximo
Anual					
alav_5anos	227.244	-2.441	388.886	-116.000.000	10.470
ga_5anos	226.420	2	154	-1	57.833
liq_5anos	220.613	35	362	-12	169.789
lucr_5anos	188.419	-4	693	-177.696	56.898
PIB5a	202.202	3	3	-5	12
inflação5a	202.202	2	2	-1	39
juros5a	202.202	6	7	0	72
ind_mercado5a	191.941	6.033	7.251	133	60.056

Fonte: elaboração própria.

Tabela 12 – Longo prazo – amostra utilizada na estimação (2000 a 2012) – Distribuição em percentis

Variáveis	Obs.	Percentis								
		1%	5%	10%	25%	50%	75%	90%	95%	99%
Longo prazo										
alav_5anos	227.244	-1.631	-35	11	30	47	65	81	88	96
ga_5anos	226.420	0	0	0	0	1	1	2	2	4
liq_5anos	220.613	1	4	6	14	28	51	75	87	97
lucr_5anos	188.419	-181	-53	-24	-2	4	10	17	22	41
PIB5a	202.202	-0	-0	0	1	3	4	6	9	12
inflação5a	202.202	-1	-0	-0	2	2	3	4	6	10
juros5a	202.202	0	0	2	2	5	6	8	12	48
ind_mercado5a	191.941	305	1.120	1.146	1.227	2.865	11.015	13.610	14.473	41.112

Fonte: elaboração própria.

Tabela 13 – Longo prazo – amostra utilizada na previsão (2013 a 2017) – estatística descritiva

Variáveis	Estatísticas Descritivas				
	Obs.	Média	Desvio Padrão	Mínimo	Máximo
Anual					
alav_5anos	115.218	-7929	1.081.580,00	-210.000.000	842.731
ga_5anos	114.958	1	48	-0	16.015
liq_5anos	112.852	38	26	-216	343
lucr_5anos	95.195	-10	924	-177.688	29.198
PIB5a	106.016	3	3	-6	11
inflação5a	106.016	2	2	-1	28
juros5a	106.016	4	5	0	61
ind_mercado5a	106.005	7.894	10.095	257	62.839

Fonte: elaboração própria.

Tabela 14 – Longo prazo – amostra utilizada na previsão (2013 a 2017) – Distribuição em percentis

Variáveis	Obs.	Percentis								
		1%	5%	10%	25%	50%	75%	90%	95%	99%
Anual										
alav_5anos	115.218	-2.984	-61	11	31	51	70	84	91	98
ga_5anos	114.958	0	0	0	0	1	1	2	2	4
liq_5anos	112.852	1	4	8	17	33	56	80	90	98
lucr_5anos	95.195	-215	-65	-29	-2	4	9	15	21	39
PIB5a	106.016	-0	0	1	2	2	5	7	8	9
inflação5a	106.016	-0	0	1	1	2	3	6	8	10
juros5a	106.016	0	0	0	1	3	6	10	10	37
ind_mercado5a	106.005	456	1.381	1.576	1.982	3.398	12.564	16.457	18.931	53.209

Fonte: elaboração própria.

APÊNDICE F – Comparação entre acertos de predição dos MLM de curto e longo prazos, por países e ano a ano

Tabela 27 – Comparação entre acertos de predição dos MLM de curto e longo prazos – ano 2013

País	MLM de curto prazo			MLM de longo prazo			Variação Percentual
	Acertos	Previsões realizadas	% Acertos	Acertos	Previsões realizadas	% Acertos	
África do Sul	97	255	38,04%	86	185	46,49%	22,21%
Alemanha	240	573	41,88%	207	479	43,22%	3,18%
Austrália	692	1.573	43,99%	684	1.167	58,61%	33,23%
Brasil	54	321	16,82%	100	244	40,98%	143,62%
Canadá	404	2.256	17,91%	803	1.576	50,95%	184,52%
China	1.206	3.076	39,21%	919	1.918	47,91%	22,21%
Cingapura	174	542	32,10%	123	358	34,36%	7,02%
Coréia do Sul	577	1.593	36,22%	714	1.337	53,40%	47,44%
EUA	1.123	3.972	28,27%	912	2.791	32,68%	15,58%
França	256	591	43,32%	203	468	43,38%	0,14%
Grécia	74	216	34,26%	140	203	68,97%	101,30%
Índia	1.160	2.140	54,21%	290	1.487	19,50%	-64,02%
Japão	1.576	3.327	47,37%	1.801	3.065	58,76%	24,05%
Malásia	304	810	37,53%	146	344	42,44%	13,09%
Polônia	112	408	27,45%	100	248	40,32%	46,89%
Reino Unido	393	1.151	34,14%	383	893	42,89%	25,61%
Rússia	159	350	45,43%	85	238	35,71%	-21,38%
Suécia	114	378	30,16%	91	291	31,27%	3,69%
Total	8.715	23.532	37,03%	7.787	17.292	45,03%	21,60%

Fonte: elaboração própria.

Tabela 28 – Comparação entre acertos de predição dos MLM de curto e longo prazos – ano 2014

País	MLM de curto prazo			MLM de longo prazo			Variação Percentual
	Acertos	Previsões realizadas	% Acertos	Acertos	Previsões realizadas	% Acertos	
África do Sul	73	239	30,54%	84	183	45,90%	50,28%
Alemanha	240	552	43,48%	190	455	41,76%	-3,96%
Austrália	612	1.576	38,83%	604	1.115	54,17%	39,50%
Brasil	40	312	12,82%	47	247	19,03%	48,42%
Canadá	516	2.055	25,11%	809	1.494	54,15%	115,66%
China	1.719	3.251	52,88%	361	2.322	15,55%	-70,60%
Cingapura	175	557	31,42%	94	363	25,90%	-17,58%
Coréia do Sul	569	1.667	34,13%	472	1.361	34,68%	1,60%
EUA	1.508	4.026	37,46%	1.207	2.717	44,42%	18,60%
França	259	608	42,60%	207	453	45,70%	7,27%
Grécia	52	207	25,12%	135	196	68,88%	174,19%
Índia	767	2.191	35,01%	648	1.425	45,47%	29,90%
Japão	1.611	3.434	46,91%	1.415	3.028	46,73%	-0,39%
Malásia	263	838	31,38%	162	354	45,76%	45,81%
Polônia	136	437	31,12%	98	263	37,26%	19,73%
Reino Unido	410	1.176	34,86%	388	859	45,17%	29,56%
Rússia	136	279	48,75%	36	212	16,98%	-65,16%
Suécia	120	379	31,66%	105	267	39,33%	24,20%
Total	9.206	23.784	38,71%	7.062	17.314	40,79%	5,38%

Fonte: elaboração própria.

Tabela 29 – Comparação entre acertos de predição dos MLM de curto e longo prazos – ano 2015

País	MLM de curto prazo			MLM de longo prazo			Variação Percentual
	Acertos	Previsões realizadas	% Acertos	Acertos	Previsões realizadas	% Acertos	
África do Sul	80	238	33,61%	73	186	39,25%	16,76%
Alemanha	250	535	46,73%	211	435	48,51%	3,80%
Austrália	547	1.568	34,89%	649	1.144	56,73%	62,62%
Brasil	19	294	6,46%	34	248	13,71%	112,14%
Canadá	388	1.849	20,98%	831	1.433	57,99%	176,35%
China	1.736	3.485	49,81%	1.039	2.617	39,70%	-20,30%
Cingapura	174	559	31,13%	73	376	19,41%	-37,63%
Coréia do Sul	593	1.744	34,00%	542	1.388	39,05%	14,84%
EUA	1.413	3.891	36,31%	1.272	2.627	48,42%	33,34%
França	270	605	44,63%	198	436	45,41%	1,76%
Grécia	70	189	37,04%	79	179	44,13%	19,16%
Índia	1.121	2.114	53,03%	500	1.336	37,43%	-29,42%
Japão	1.725	3.473	49,67%	1.317	3.019	43,62%	-12,17%
Malásia	268	827	32,41%	148	409	36,19%	11,66%
Polônia	154	442	34,84%	107	303	35,31%	1,35%
Reino Unido	494	1.151	42,92%	389	826	47,09%	9,73%
Rússia	39	207	18,84%	22	156	14,10%	-25,15%
Suécia	128	416	30,77%	102	257	39,69%	28,99%
Total	9.469	23.587	40,14%	7.586	17.375	43,66%	8,76%

Fonte: elaboração própria.

Tabela 30 – Comparação entre acertos de predição dos MLM de curto e longo prazos – ano 2016

País	MLM de curto prazo			MLM de longo prazo			Variação Percentual
	Acertos	Previsões realizadas	% Acertos	Acertos	Previsões realizadas	% Acertos	
África do Sul	77	222	34,68%	74	176	42,05%	21,22%
Alemanha	212	491	43,18%	176	405	43,46%	0,65%
Austrália	533	1.543	34,54%	710	1.180	60,17%	74,19%
Brasil	162	287	56,45%	46	245	18,78%	-66,74%
Canadá	646	1.712	37,73%	745	1.347	55,31%	46,58%
China	615	3.681	16,71%	1.521	2.776	54,79%	227,94%
Cingapura	169	539	31,35%	129	403	32,01%	2,09%
Coréia do Sul	652	1.818	35,86%	565	1.397	40,44%	12,77%
EUA	1.129	3.687	30,62%	1.180	2.579	45,75%	49,42%
França	218	601	36,27%	188	435	43,22%	19,15%
Grécia	61	185	32,97%	60	174	34,48%	4,58%
Índia	880	2.229	39,48%	495	1.367	36,21%	-8,28%
Japão	1.571	3.489	45,03%	1.532	2.982	51,37%	14,10%
Malásia	272	832	32,69%	173	620	27,90%	-14,65%
Polônia	144	431	33,41%	113	312	36,22%	8,40%
Reino Unido	415	1.163	35,68%	403	821	49,09%	37,56%
Rússia	26	198	13,13%	45	153	29,41%	123,98%
Suécia	152	450	33,78%	85	250	34,00%	0,66%
Total	7.934	23.558	33,68%	8.240	17.622	46,76%	38,84%

Fonte: elaboração própria.

Tabela 31 – Comparação entre acertos de predição dos MLM de curto e longo prazos – ano 2017

País	MLM de curto prazo			MLM de longo prazo			Variação Percentual
	Acertos	Previsões realizadas	% Acertos	Acertos	Previsões realizadas	% Acertos	
África do Sul	94	223	42,15%	101	182	55,49%	31,65%
Alemanha	147	429	34,27%	147	365	40,27%	17,53%
Austrália	529	1.536	34,44%	699	1.182	59,14%	71,71%
Brasil	103	289	35,64%	53	249	21,29%	-40,28%
Canadá	666	1.736	38,36%	672	1.294	51,93%	35,37%
China	649	4.106	15,81%	1.174	2.851	41,18%	160,52%
Cingapura	193	539	35,81%	147	422	34,83%	-2,72%
Coréia do Sul	639	1.907	33,51%	582	1.419	41,01%	22,40%
EUA	1.222	3.552	34,40%	1.126	2.594	43,41%	26,17%
França	209	532	39,29%	161	409	39,36%	0,20%
Grécia	41	176	23,30%	47	161	29,19%	25,31%
Índia	1.045	2.346	44,54%	761	1.465	51,95%	16,62%
Japão	1.664	3.556	46,79%	1.436	3.027	47,44%	1,38%
Malásia	215	854	25,18%	200	687	29,11%	15,64%
Polônia	148	413	35,84%	95	307	30,94%	-13,65%
Reino Unido	388	1.136	34,15%	383	808	47,40%	38,78%
Rússia	49	177	27,68%	45	137	32,85%	18,65%
Suécia	166	482	34,44%	78	244	31,97%	-7,18%
Total	8.167	23.989	34,04%	7.907	17.803	44,41%	30,46%

Fonte: elaboração própria.

APÊNDICE G – Do-file – amostra EUA

* Project: Minha Tese, só com EUA

* Last version: 22/02/2020

* Author: Sérgio Carlos dos Santos

*1) Limpa a memória e cria o arquivo log

clear all

log using "C:\Users\sergi\Documents\Sergio\Doutorado\Projeto_Doutorado\2018_1\Base de dados\Thomson Reuters\36c12743dc84c86b_dta\minha_tese.log", text replace

*2) Carrega arquivo base de dados e deleta os demais países, mantendo apenas um determinado país escolhido (pais_)

use "C:\Users\sergi\Documents\Sergio\Doutorado\Projeto_Doutorado\2018_1\Base de dados\Thomson Reuters\36c12743dc84c86b_dta\BASE_última versão_22022020.dta"
tab item6026

gen pais_="UNITED STATES"

keep if item6026==pais_

tab item6026

*3) Classifica anos do painel por código da empresa (WORLDSCOPE PERM I.D.)

bysort item6105 year_ : gen year=year_[_n-1]

*4) Deleta a variável year criada com missing values durante a classificação dos anos do painel

drop year

*5) Verifica observações (linhas) duplicadas por empresa antes de definir o painel por código da empresa e ano

quietly by item6105 year_ : gen dup=cond(_N==1,0,_n)

tab dup

*6) Deleta observações duplicadas

drop if dup>0

tab dup

*7) Deleta a variável dup criada para checar observações duplicadas

drop dup

*8) Cria a variável id (chave primária a partir da variável item6105 - WORLDSCOPE PERM I.D.) transformando em número a string item6105

encode item6105, gen (id)

*9) Define o painel de dados por código da empresa e ano

xtset id year_

*10) Calcula retornos de mercado de cada empresa, anual e quinquenal, criando as variáveis ret1a e ret5a

by id: gen ret1a=(item5001-item5001[_n-1])/item5001[_n-1]

by id: gen ret5a=(item5001-item5001[_n-5])/item5001[_n-5]

*11) Classifica os retornos anuais e quinquenais em tercís, de toda a base, criando as variáveis tret1a e tret5a, e checa se a divisão dos retornos por tercís estão igualmente distribuídos

xtile tret1a = ret1a, nq(3)

tab tret1a

xtile tret5a = ret5a, nq(3)

tab tret5a

*12) Classifica em tercís os retornos anuais em cada ano da base, dos anos de 1996 a 2017, criando as variáveis `ret1aYYYY` de cada ano, e checa se a divisão dos retornos por tercís está igualmente distribuída

```
xtile ret1a1996 = ret1a if year_==1996, nq(3)
```

```
tab ret1a1996
```

```
xtile ret1a1997 = ret1a if year_==1997, nq(3)
```

```
tab ret1a1997
```

```
xtile ret1a1998 = ret1a if year_==1998, nq(3)
```

```
tab ret1a1998
```

```
xtile ret1a1999 = ret1a if year_==1999, nq(3)
```

```
tab ret1a1999
```

```
xtile ret1a2000 = ret1a if year_==2000, nq(3)
```

```
tab ret1a2000
```

```
xtile ret1a2001 = ret1a if year_==2001, nq(3)
```

```
tab ret1a2001
```

```
xtile ret1a2002 = ret1a if year_==2002, nq(3)
```

```
tab ret1a2002
```

```
xtile ret1a2003 = ret1a if year_==2003, nq(3)
```

```
tab ret1a2003
```

```
xtile ret1a2004 = ret1a if year_==2004, nq(3)
```

```
tab ret1a2004
```

```
xtile ret1a2005 = ret1a if year_==2005, nq(3)
```

```
tab ret1a2005
```

```
xtile ret1a2006 = ret1a if year_==2006, nq(3)
```

```
tab ret1a2006
```

```
xtile ret1a2007 = ret1a if year_==2007, nq(3)
```

```
tab ret1a2007
```

```
xtile ret1a2008 = ret1a if year_==2008, nq(3)
```

```
tab ret1a2008
```

```
xtile ret1a2009 = ret1a if year_==2009, nq(3)
```

```
tab ret1a2009
```

```
xtile ret1a2010 = ret1a if year_==2010, nq(3)
```

```
tab ret1a2010
```

```
xtile ret1a2011 = ret1a if year_==2011, nq(3)
```

```
tab ret1a2011
```

```
xtile ret1a2012 = ret1a if year_==2012, nq(3)
```

```
tab ret1a2012
```

```
xtile ret1a2013 = ret1a if year_==2013, nq(3)
```

```
tab ret1a2013
```

```
xtile ret1a2014 = ret1a if year_==2014, nq(3)
```

```
tab ret1a2014
```

```
xtile ret1a2015 = ret1a if year_==2015, nq(3)
```

```
tab ret1a2015
```

```
xtile ret1a2016 = ret1a if year_==2016, nq(3)
```

```
tab ret1a2016
```

```
xtile ret1a2017 = ret1a if year_==2017, nq(3)
```

```
tab ret1a2017
```

*13) Classifica em tercís os retornos quinquenais dos anos de 2000 a 2017, criando as variáveis `ret5aYYYY` de cada ano, e checa se a divisão dos retornos por tercís está igualmente distribuída

```

xtile tret5a2000 = ret5a if year_==2000, nq(3)
tab tret5a2000
xtile tret5a2001 = ret5a if year_==2001, nq(3)
tab tret5a2001
xtile tret5a2002 = ret5a if year_==2002, nq(3)
tab tret5a2002
xtile tret5a2003 = ret5a if year_==2003, nq(3)
tab tret5a2003
xtile tret5a2004 = ret5a if year_==2004, nq(3)
tab tret5a2004
xtile tret5a2005 = ret5a if year_==2005, nq(3)
tab tret5a2005
xtile tret5a2006 = ret5a if year_==2006, nq(3)
tab tret5a2006
xtile tret5a2007 = ret5a if year_==2007, nq(3)
tab tret5a2007
xtile tret5a2008 = ret5a if year_==2008, nq(3)
tab tret5a2008
xtile tret5a2009 = ret5a if year_==2009, nq(3)
tab tret5a2009
xtile tret5a2010 = ret5a if year_==2010, nq(3)
tab tret5a2010
xtile tret5a2011 = ret5a if year_==2011, nq(3)
tab tret5a2011
xtile tret5a2012 = ret5a if year_==2012, nq(3)
tab tret5a2012
xtile tret5a2013 = ret5a if year_==2013, nq(3)
tab tret5a2013
xtile tret5a2014 = ret5a if year_==2014, nq(3)
tab tret5a2014
xtile tret5a2015 = ret5a if year_==2015, nq(3)
tab tret5a2015
xtile tret5a2016 = ret5a if year_==2016, nq(3)
tab tret5a2016
xtile tret5a2017 = ret5a if year_==2017, nq(3)
tab tret5a2017

```

*14) Calcula e cria as variáveis das médias aritméticas de 5 anos das variáveis macroeconômicas e índice de mercado

by id: gen PIB5a = (PIB[_n]+PIB[_n-1]+PIB[_n-2]+PIB[_n-3]+PIB[_n-4])/5

by id: gen inflação5a = (inflação[_n]+inflação[_n-1]+inflação[_n-2]+inflação[_n-3]+inflação[_n-4])/5

by id: gen juros5a = (juros[_n]+juros[_n-1]+juros[_n-2]+juros[_n-3]+juros[_n-4])/5

by id: gen ind_mercado5a = (ind_mercado[_n]+ind_mercado[_n-1]+ind_mercado[_n-2]+ind_mercado[_n-3]+ind_mercado[_n-4])/5

*15) Verifica a correlação entre os retornos anuais e quinquenais de cada empresa e as variáveis macroeconômicas e índice de mercado anuais e quinquenais (média 5 anos)

correlate ret1a PIB inflação juros ind_mercado

correlate ret5a PIB5a inflação5a juros5a ind_mercado5a

*16) Verifica a correlação entre os retornos anuais e quinquenais de cada empresa e os indicadores contábeis disponíveis na base Thomson Reuters em ambos os formatos: anuais e quinquenais (média 5 anos)

```
correlate ret1a alav_1ano item8141 item15121 item8236 item8221 item8271 lucr_1ano
item8226
```

```
correlate ret5a alav_5anos item8145 item15124 item8240 item8225 item8275 lucr_5anos
item8230
```

```
correlate ret1a ga_1ano item8231 item8326 item8106 item8361 item8301 item8366 item8101
liq_1ano item8911
```

```
correlate ret5a ga_5anos item8235 item8330 item8110 item8365 item8305 item8370
item8105 liq_5anos item8915
```

*17) Verifica a correlação entre os retornos anuais e quinquenais de cada empresa e todos os indicadores testados nos passos 15 e 16

```
correlate ret1a PIB inflação juros ind_mercado alav_1ano item8141 item15121 item8236
item8221 item8271 lucr_1ano item8226 ga_1ano item8231 item8326 item8106 item8361
item8301 item8366 item8101 liq_1ano item8911
```

```
correlate ret5a PIB5a inflação5a juros5a ind_mercado5a alav_5anos item8145 item15124
item8240 item8225 item8275 lucr_5anos item8230 ga_5anos item8235 item8330 item8110
item8365 item8305 item8370 item8105 liq_5anos item8915
```

*18) Verifica a correlação entre os retornos anuais e quinquenais de cada empresa e todos os indicadores que compõem os modelos logit

```
correlate ret1a alav_1ano ga_1ano liq_1ano lucr_1ano PIB inflação juros ind_mercado
correlate ret5a alav_5anos ga_5anos liq_5anos lucr_5anos PIB5a inflação5a juros5a
ind_mercado5a
```

** Retorno Anual

*19) Usa dentre as variáveis analisadas no passo 17, as variáveis de maior correlação com retorno anual (sendo apenas um indicador contábil de Lucratividade, Liquidez, Giro do Ativo e Alavancagem; as variáveis macroeconômicas; e o índice de mercado de capitais do país),
 * para rodar o logit multinomial (mlogit), 1º tercil como base e 15 iterações para forçar o logit a convergir para uma determinada função de verossimilhança, caso a convergência não seja normalmente alcançada

* utiliza como variável dependente tret1a_ano_a_ano, criada a seguir:

```
gen tret1a_ano_a_ano=.
```

```
replace tret1a_ano_a_ano=tret1a1996 if year_==1996
```

```
replace tret1a_ano_a_ano=tret1a1997 if year_==1997
```

```
replace tret1a_ano_a_ano=tret1a1998 if year_==1998
```

```
replace tret1a_ano_a_ano=tret1a1999 if year_==1999
```

```
replace tret1a_ano_a_ano=tret1a2000 if year_==2000
```

```
replace tret1a_ano_a_ano=tret1a2001 if year_==2001
```

```
replace tret1a_ano_a_ano=tret1a2002 if year_==2002
```

```
replace tret1a_ano_a_ano=tret1a2003 if year_==2003
```

```
replace tret1a_ano_a_ano=tret1a2004 if year_==2004
```

```
replace tret1a_ano_a_ano=tret1a2005 if year_==2005
```

```
replace tret1a_ano_a_ano=tret1a2006 if year_==2006
```

```
replace tret1a_ano_a_ano=tret1a2007 if year_==2007
```

```
replace tret1a_ano_a_ano=tret1a2008 if year_==2008
```

```
replace tret1a_ano_a_ano=tret1a2009 if year_==2009
```

```

replace tret1a_ano_a_ano=tret1a2010 if year_==2010
replace tret1a_ano_a_ano=tret1a2011 if year_==2011
replace tret1a_ano_a_ano=tret1a2012 if year_==2012
replace tret1a_ano_a_ano=tret1a2013 if year_==2013
replace tret1a_ano_a_ano=tret1a2014 if year_==2014
replace tret1a_ano_a_ano=tret1a2015 if year_==2015
replace tret1a_ano_a_ano=tret1a2016 if year_==2016
replace tret1a_ano_a_ano=tret1a2017 if year_==2017
mlogit tret1a_ano_a_ano alav_1ano ga_1ano liq_1ano lucr_1ano PIB inflação juros
ind_mercado if inrange(year_,1996,2012), base(1) iterate(15)
*20) Prediz a probabilidade de ser categoria 1 (baixo retorno) para os retornos anuais nos anos
2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do modelo rodado no passo 18,
criando-se as variáveis pla_1_2013, pla_1_2014, pla_1_2015, pla_1_2016 e pla_1_2017
predict pla_1_2013 if year_==2013, outcome(1) p
predict pla_1_2014 if year_==2014, outcome(1) p
predict pla_1_2015 if year_==2015, outcome(1) p
predict pla_1_2016 if year_==2016, outcome(1) p
predict pla_1_2017 if year_==2017, outcome(1) p
*21) Detalha a distribuição das variáveis criadas para mostrar a probabilidade de ser categoria
1 (baixo retorno)
sum pla_1_2013, detail
sum pla_1_2014, detail
sum pla_1_2015, detail
sum pla_1_2016, detail
sum pla_1_2017, detail
*22) Prediz a probabilidade de ser categoria 2 (médio retorno) para os retornos anuais nos
anos 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do modelo rodado no passo 18,
criando-se as variáveis pla_2_2013, pla_2_2014, pla_2_2015, pla_2_2016 e pla_2_2017
predict pla_2_2013 if year_==2013, outcome(2) p
predict pla_2_2014 if year_==2014, outcome(2) p
predict pla_2_2015 if year_==2015, outcome(2) p
predict pla_2_2016 if year_==2016, outcome(2) p
predict pla_2_2017 if year_==2017, outcome(2) p
*23) Detalha a distribuição das variáveis criadas para mostrar a probabilidade de ser categoria
2 (médio retorno)
sum pla_2_2013, detail
sum pla_2_2014, detail
sum pla_2_2015, detail
sum pla_2_2016, detail
sum pla_2_2017, detail
*24) Prediz a probabilidade de ser categoria 3 (alto retorno) para os retornos anuais nos anos
2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do modelo rodado no passo 18,
criando-se as variáveis pla_3_2013, pla_3_2014, pla_3_2015, pla_3_2016 e pla_3_2017
predict pla_3_2013 if year_==2013, outcome(3) p
predict pla_3_2014 if year_==2014, outcome(3) p
predict pla_3_2015 if year_==2015, outcome(3) p
predict pla_3_2016 if year_==2016, outcome(3) p
predict pla_3_2017 if year_==2017, outcome(3) p
*25) Detalha a distribuição das variáveis criadas para mostrar a probabilidade de ser categoria
3 (alto retorno)

```

sum pla_3_2013, detail

sum pla_3_2014, detail

sum pla_3_2015, detail

sum pla_3_2016, detail

sum pla_3_2017, detail

*26) Cria variáveis de categorias preditas para os anos de 2013, 2014, 2015, 2016 e 2017 para cada empresa que teve dados calculados pelo modelo logit multinomial (mlogit) no passo 18, a variável assumirá o valor da categoria com maior probabilidade de ocorrer (probabilidade dominante)

gen pla_cat_2013=.

replace pla_cat_2013=1 if pla_1_2013>pla_2_2013 & pla_1_2013>pla_3_2013

replace pla_cat_2013=2 if pla_2_2013>pla_1_2013 & pla_2_2013>pla_3_2013

replace pla_cat_2013=3 if pla_3_2013>pla_1_2013 & pla_3_2013>pla_2_2013

gen pla_cat_2014=.

replace pla_cat_2014=1 if pla_1_2014>pla_2_2014 & pla_1_2014>pla_3_2014

replace pla_cat_2014=2 if pla_2_2014>pla_1_2014 & pla_2_2014>pla_3_2014

replace pla_cat_2014=3 if pla_3_2014>pla_1_2014 & pla_3_2014>pla_2_2014

gen pla_cat_2015=.

replace pla_cat_2015=1 if pla_1_2015>pla_2_2015 & pla_1_2015>pla_3_2015

replace pla_cat_2015=2 if pla_2_2015>pla_1_2015 & pla_2_2015>pla_3_2015

replace pla_cat_2015=3 if pla_3_2015>pla_1_2015 & pla_3_2015>pla_2_2015

gen pla_cat_2016=.

replace pla_cat_2016=1 if pla_1_2016>pla_2_2016 & pla_1_2016>pla_3_2016

replace pla_cat_2016=2 if pla_2_2016>pla_1_2016 & pla_2_2016>pla_3_2016

replace pla_cat_2016=3 if pla_3_2016>pla_1_2016 & pla_3_2016>pla_2_2016

gen pla_cat_2017=.

replace pla_cat_2017=1 if pla_1_2017>pla_2_2017 & pla_1_2017>pla_3_2017

replace pla_cat_2017=2 if pla_2_2017>pla_1_2017 & pla_2_2017>pla_3_2017

replace pla_cat_2017=3 if pla_3_2017>pla_1_2017 & pla_3_2017>pla_2_2017

*27) Verifica a distribuição das variáveis (valores previstos): pla_cat_2013, pla_cat_2014, pla_cat_2015, pla_cat_2016 e pla_cat_2017

tab pla_cat_2013

tab pla_cat_2014

tab pla_cat_2015

tab pla_cat_2016

tab pla_cat_2017

*28) Compara os resultados previstos e realizados. Observar, na matriz gerada, a quantidade da diagonal principal em relação à quantidade total:

tab tretla2013 pla_cat_2013

tab tretla2014 pla_cat_2014

tab tretla2015 pla_cat_2015

tab tretla2016 pla_cat_2016

tab tretla2017 pla_cat_2017

*29) Cria a variável prev_corr1a para as previsões dos anos 2013, 2014, 2015, 2016 e 2017 para mostrar a categoria prevista corretamente(valor previsto pla_cat_20YY é igual ao valor realizado tretla20YY)

gen prev_corr1a2013=pla_cat_2013 if year==2013 & tretla2013==pla_cat_2013

gen prev_corr1a2014=pla_cat_2014 if year==2014 & tretla2014==pla_cat_2014

gen prev_corr1a2015=pla_cat_2015 if year==2015 & tretla2015==pla_cat_2015

gen prev_corr1a2016=pla_cat_2016 if year==2016 & tretla2016==pla_cat_2016

```
gen prev_corr1a2017=p1a_cat_2017 if year==2017 & tret1a2017==p1a_cat_2017
```

*30) Verifica a distribuição das variáveis (prev_corr1a): anos 2013, 2014, 2015, 2016 e 2017

```
tab prev_corr1a2013
```

```
tab prev_corr1a2014
```

```
tab prev_corr1a2015
```

```
tab prev_corr1a2016
```

```
tab prev_corr1a2017
```

*31) Cria a variável pr_acerto1a2013 e reporta nessa variável a probabilidade informada no predict, conforme passos 19, 21 e 23 anteriores

```
gen pr_acerto1a2013=.
```

```
replace pr_acerto1a2013=p1a_1_2013 if p1a_1_2013>p1a_2_2013 &
p1a_1_2013>p1a_3_2013
```

```
replace pr_acerto1a2013=p1a_2_2013 if p1a_2_2013>p1a_1_2013 &
p1a_2_2013>p1a_3_2013
```

```
replace pr_acerto1a2013=p1a_3_2013 if p1a_3_2013>p1a_1_2013 &
p1a_3_2013>p1a_2_2013
```

*32) Cria a variável pr_acerto1a2014 e reporta nessa variável a probabilidade informada no predict, conforme passos 19, 21 e 23 anteriores

```
gen pr_acerto1a2014=.
```

```
replace pr_acerto1a2014=p1a_1_2014 if p1a_1_2014>p1a_2_2014 &
p1a_1_2014>p1a_3_2014
```

```
replace pr_acerto1a2014=p1a_2_2014 if p1a_2_2014>p1a_1_2014 &
p1a_2_2014>p1a_3_2014
```

```
replace pr_acerto1a2014=p1a_3_2014 if p1a_3_2014>p1a_1_2014 &
p1a_3_2014>p1a_2_2014
```

*33) Cria a variável pr_acerto1a2015 e reporta nessa variável a probabilidade informada no predict, conforme passos 19, 21 e 23 anteriores

```
gen pr_acerto1a2015=.
```

```
replace pr_acerto1a2015=p1a_1_2015 if p1a_1_2015>p1a_2_2015 &
p1a_1_2015>p1a_3_2015
```

```
replace pr_acerto1a2015=p1a_2_2015 if p1a_2_2015>p1a_1_2015 &
p1a_2_2015>p1a_3_2015
```

```
replace pr_acerto1a2015=p1a_3_2015 if p1a_3_2015>p1a_1_2015 &
p1a_3_2015>p1a_2_2015
```

*34) Cria a variável pr_acerto1a2016 e reporta nessa variável a probabilidade informada no predict, conforme passos 19, 21 e 23 anteriores

```
gen pr_acerto1a2016=.
```

```
replace pr_acerto1a2016=p1a_1_2016 if p1a_1_2016>p1a_2_2016 &
p1a_1_2016>p1a_3_2016
```

```
replace pr_acerto1a2016=p1a_2_2016 if p1a_2_2016>p1a_1_2016 &
p1a_2_2016>p1a_3_2016
```

```
replace pr_acerto1a2016=p1a_3_2016 if p1a_3_2016>p1a_1_2016 &
p1a_3_2016>p1a_2_2016
```

*35) Cria a variável pr_acerto1a2017 e reporta nessa variável a probabilidade informada no predict, conforme passos 19, 21 e 23 anteriores

```
gen pr_acerto1a2017=.
```

```
replace pr_acerto1a2017=p1a_1_2017 if p1a_1_2017>p1a_2_2017 &
p1a_1_2017>p1a_3_2017
```

```
replace pr_acerto1a2017=p1a_2_2017 if p1a_2_2017>p1a_1_2017 &
p1a_2_2017>p1a_3_2017
```

```
replace pr_acerto1a2017=pla_3_2017 if pla_3_2017>pla_1_2017 &
pla_3_2017>pla_2_2017
```

*36) Detalha a distribuição das probabilidades dos acertos

```
sum pr_acerto1a2013, detail
```

```
sum pr_acerto1a2014, detail
```

```
sum pr_acerto1a2015, detail
```

```
sum pr_acerto1a2016, detail
```

```
sum pr_acerto1a2017, detail
```

```
*****
```

** Retorno Quinquenal

```
*****
```

*37) Usa dentre as variáveis analisadas no passo 17, as variáveis quinquenais recíprocas às de maior correlação com retorno anual (utilizadas no passo 18 e sendo apenas um indicador contábil de Lucratividade, Liquidez, Giro do Ativo e Alavancagem; as variáveis macroeconômicas; e o índice de mercado de capitais do país),

* para rodar o logit multinomial (mlogit), 1º tercil como base e 15 iterações para forçar o logit a convergir para uma determinada função de verossimilhança, caso a convergência não seja normalmente alcançada

* utiliza como variável dependente tret5a_ano_a_ano, criada a seguir:

```
gen tret5a_ano_a_ano=.
```

```
replace tret5a_ano_a_ano=tret5a2000 if year_==2000
```

```
replace tret5a_ano_a_ano=tret5a2001 if year_==2001
```

```
replace tret5a_ano_a_ano=tret5a2002 if year_==2002
```

```
replace tret5a_ano_a_ano=tret5a2003 if year_==2003
```

```
replace tret5a_ano_a_ano=tret5a2004 if year_==2004
```

```
replace tret5a_ano_a_ano=tret5a2005 if year_==2005
```

```
replace tret5a_ano_a_ano=tret5a2006 if year_==2006
```

```
replace tret5a_ano_a_ano=tret5a2007 if year_==2007
```

```
replace tret5a_ano_a_ano=tret5a2008 if year_==2008
```

```
replace tret5a_ano_a_ano=tret5a2009 if year_==2009
```

```
replace tret5a_ano_a_ano=tret5a2010 if year_==2010
```

```
replace tret5a_ano_a_ano=tret5a2011 if year_==2011
```

```
replace tret5a_ano_a_ano=tret5a2012 if year_==2012
```

```
replace tret5a_ano_a_ano=tret5a2013 if year_==2013
```

```
replace tret5a_ano_a_ano=tret5a2014 if year_==2014
```

```
replace tret5a_ano_a_ano=tret5a2015 if year_==2015
```

```
replace tret5a_ano_a_ano=tret5a2016 if year_==2016
```

```
replace tret5a_ano_a_ano=tret5a2017 if year_==2017
```

```
mlogit tret5a_ano_a_ano alav_5anos ga_5anos liq_5anos lucr_5anos PIB5a inflação5a
```

```
juros5a ind_mercado5a if inrange(year_,2000,2012), base(1) iterate(15)
```

*38) Prediz a probabilidade de ser categoria 1 (baixo retorno) para os retornos quinquenais nos anos 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do modelo rodado no passo 36, criando-se as variáveis p5a_1_YYYY para cada um dos 5 anos

```
predict p5a_1_2013 if year_==2013, outcome(1) p
```

```
predict p5a_1_2014 if year_==2014, outcome(1) p
```

```
predict p5a_1_2015 if year_==2015, outcome(1) p
```

```
predict p5a_1_2016 if year_==2016, outcome(1) p
```

```
predict p5a_1_2017 if year_==2017, outcome(1) p
```

*39) Detalha a distribuição das variáveis criadas para mostrar a probabilidade de ser categoria 1 (baixo retorno)

sum p5a_1_2013, detail

sum p5a_1_2014, detail

sum p5a_1_2015, detail

sum p5a_1_2016, detail

sum p5a_1_2017, detail

*40) Prediz a probabilidade de ser categoria 2 (médio retorno) para os retornos anuais nos anos 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do modelo rodado no passo 36, criando-se as variáveis p5a_2_YYYY para cada um dos 5 anos

predict p5a_2_2013 if year_==2013, outcome(2) p

predict p5a_2_2014 if year_==2014, outcome(2) p

predict p5a_2_2015 if year_==2015, outcome(2) p

predict p5a_2_2016 if year_==2016, outcome(2) p

predict p5a_2_2017 if year_==2017, outcome(2) p

*41) Detalha a distribuição das variáveis criadas para mostrar a probabilidade de ser categoria 2 (médio retorno)

sum p5a_2_2013, detail

sum p5a_2_2014, detail

sum p5a_2_2015, detail

sum p5a_2_2016, detail

sum p5a_2_2017, detail

*42) Prediz a probabilidade de ser categoria 3 (alto retorno) para os retornos anuais nos anos 2013, 2014, 2015, 2016 e 2017, utilizando os parâmetros do modelo rodado no passo 36, criando-se as variáveis p5a_3_YYYY para cada um dos 5 anos

predict p5a_3_2013 if year_==2013, outcome(3) p

predict p5a_3_2014 if year_==2014, outcome(3) p

predict p5a_3_2015 if year_==2015, outcome(3) p

predict p5a_3_2016 if year_==2016, outcome(3) p

predict p5a_3_2017 if year_==2017, outcome(3) p

*43) Detalha a distribuição das variáveis criadas para mostrar a probabilidade de ser categoria 3 (alto retorno)

sum p5a_3_2013, detail

sum p5a_3_2014, detail

sum p5a_3_2015, detail

sum p5a_3_2016, detail

sum p5a_3_2017, detail

*44) Cria variáveis de categorias preditas para os anos de 2013, 2014, 2015, 2016 e 2017 para cada empresa que teve dados calculados pelo modelo logit multinomial (mlogit) no passo 36, a variável assumirá o valor da categoria com maior probabilidade de ocorrer (probabilidade dominante)

gen p5a_cat_2013=.

replace p5a_cat_2013=1 if p5a_1_2013>p5a_2_2013 & p5a_1_2013>p5a_3_2013

replace p5a_cat_2013=2 if p5a_2_2013>p5a_1_2013 & p5a_2_2013>p5a_3_2013

replace p5a_cat_2013=3 if p5a_3_2013>p5a_1_2013 & p5a_3_2013>p5a_2_2013

gen p5a_cat_2014=.

replace p5a_cat_2014=1 if p5a_1_2014>p5a_2_2014 & p5a_1_2014>p5a_3_2014

replace p5a_cat_2014=2 if p5a_2_2014>p5a_1_2014 & p5a_2_2014>p5a_3_2014

replace p5a_cat_2014=3 if p5a_3_2014>p5a_1_2014 & p5a_3_2014>p5a_2_2014

gen p5a_cat_2015=.

replace p5a_cat_2015=1 if p5a_1_2015>p5a_2_2015 & p5a_1_2015>p5a_3_2015

replace p5a_cat_2015=2 if p5a_2_2015>p5a_1_2015 & p5a_2_2015>p5a_3_2015

replace p5a_cat_2015=3 if p5a_3_2015>p5a_1_2015 & p5a_3_2015>p5a_2_2015
gen p5a_cat_2016=.

replace p5a_cat_2016=1 if p5a_1_2016>p5a_2_2016 & p5a_1_2016>p5a_3_2016

replace p5a_cat_2016=2 if p5a_2_2016>p5a_1_2016 & p5a_2_2016>p5a_3_2016

replace p5a_cat_2016=3 if p5a_3_2016>p5a_1_2016 & p5a_3_2016>p5a_2_2016

gen p5a_cat_2017=.

replace p5a_cat_2017=1 if p5a_1_2017>p5a_2_2017 & p5a_1_2017>p5a_3_2017

replace p5a_cat_2017=2 if p5a_2_2017>p5a_1_2017 & p5a_2_2017>p5a_3_2017

replace p5a_cat_2017=3 if p5a_3_2017>p5a_1_2017 & p5a_3_2017>p5a_2_2017

*45) Verifica a distribuição das variáveis (valores previstos): p5a_cat_2013, p5a_cat_2014, p5a_cat_2015, p5a_cat_2016 e p5a_cat_2017

tab p5a_cat_2013

tab p5a_cat_2014

tab p5a_cat_2015

tab p5a_cat_2016

tab p5a_cat_2017

*46) Compara os resultados previstos e realizados. Observar, na matriz gerada, a quantidade da diagonal principal em relação à quantidade total:

tab tret5a2013 p5a_cat_2013

tab tret5a2014 p5a_cat_2014

tab tret5a2015 p5a_cat_2015

tab tret5a2016 p5a_cat_2016

tab tret5a2017 p5a_cat_2017

*47) Cria a variável prev_corr5a para as previsões dos anos 2013, 2014, 2015, 2016 e 2017 para mostrar a categoria prevista corretamente (valor previsto p5a_cat_20YY é igual ao valor realizado tret5a20YY)

gen prev_corr5a2013=p5a_cat_2013 if year==2013 & tret5a2013==p5a_cat_2013

gen prev_corr5a2014=p5a_cat_2014 if year==2014 & tret5a2014==p5a_cat_2014

gen prev_corr5a2015=p5a_cat_2015 if year==2015 & tret5a2015==p5a_cat_2015

gen prev_corr5a2016=p5a_cat_2016 if year==2016 & tret5a2016==p5a_cat_2016

gen prev_corr5a2017=p5a_cat_2017 if year==2017 & tret5a2017==p5a_cat_2017

*48) Verifica a distribuição das variáveis (prev_corr5a): anos 2013, 2014, 2015, 2016 e 2017

tab prev_corr5a2013

tab prev_corr5a2014

tab prev_corr5a2015

tab prev_corr5a2016

tab prev_corr5a2017

*49) Cria a variável pr_acerto5a2013 e reporta nessa variável a probabilidade informada no predict, conforme passos 37, 39 e 41 anteriores

gen pr_acerto5a2013=.

replace pr_acerto5a2013=p5a_1_2013 if p5a_1_2013>p5a_2_2013 &
p5a_1_2013>p5a_3_2013

replace pr_acerto5a2013=p5a_2_2013 if p5a_2_2013>p5a_1_2013 &
p5a_2_2013>p5a_3_2013

replace pr_acerto5a2013=p5a_3_2013 if p5a_3_2013>p5a_1_2013 &
p5a_3_2013>p5a_2_2013

*50) Cria a variável pr_acerto5a2014 e reporta nessa variável a probabilidade informada no predict, conforme passos 37, 39 e 41 anteriores

gen pr_acerto5a2014=.

```
replace pr_acerto5a2014=p5a_1_2014 if p5a_1_2014>p5a_2_2014 &
p5a_1_2014>p5a_3_2014
```

```
replace pr_acerto5a2014=p5a_2_2014 if p5a_2_2014>p5a_1_2014 &
p5a_2_2014>p5a_3_2014
```

```
replace pr_acerto5a2014=p5a_3_2014 if p5a_3_2014>p5a_1_2014 &
p5a_3_2014>p5a_2_2014
```

*51) Cria a variável pr_acerto5a2015 e reporta nessa variável a probabilidade informada no predict, conforme passos 37, 39 e 41 anteriores

```
gen pr_acerto5a2015=.
```

```
replace pr_acerto5a2015=p5a_1_2015 if p5a_1_2015>p5a_2_2015 &
p5a_1_2015>p5a_3_2015
```

```
replace pr_acerto5a2015=p5a_2_2015 if p5a_2_2015>p5a_1_2015 &
p5a_2_2015>p5a_3_2015
```

```
replace pr_acerto5a2015=p5a_3_2015 if p5a_3_2015>p5a_1_2015 &
p5a_3_2015>p5a_2_2015
```

*52) Cria a variável pr_acerto5a2016 e reporta nessa variável a probabilidade informada no predict, conforme passos 37, 39 e 41 anteriores

```
gen pr_acerto5a2016=.
```

```
replace pr_acerto5a2016=p5a_1_2016 if p5a_1_2016>p5a_2_2016 &
p5a_1_2016>p5a_3_2016
```

```
replace pr_acerto5a2016=p5a_2_2016 if p5a_2_2016>p5a_1_2016 &
p5a_2_2016>p5a_3_2016
```

```
replace pr_acerto5a2016=p5a_3_2016 if p5a_3_2016>p5a_1_2016 &
p5a_3_2016>p5a_2_2016
```

*53) Cria a variável pr_acerto5a2017 e reporta nessa variável a probabilidade informada no predict, conforme passos 37, 39 e 41 anteriores

```
gen pr_acerto5a2017=.
```

```
replace pr_acerto5a2017=p5a_1_2017 if p5a_1_2017>p5a_2_2017 &
p5a_1_2017>p5a_3_2017
```

```
replace pr_acerto5a2017=p5a_2_2017 if p5a_2_2017>p5a_1_2017 &
p5a_2_2017>p5a_3_2017
```

```
replace pr_acerto5a2017=p5a_3_2017 if p5a_3_2017>p5a_1_2017 &
p5a_3_2017>p5a_2_2017
```

*54) Detalha a distribuição das probabilidades dos acertos

```
sum pr_acerto5a2013, detail
```

```
sum pr_acerto5a2014, detail
```

```
sum pr_acerto5a2015, detail
```

```
sum pr_acerto5a2016, detail
```

```
sum pr_acerto5a2017, detail
```

*55) Fecha o arquivo log

```
log close
```