

ESTRATÉGIA DE INVESTIMENTO LONG-SHORT COM CARTEIRAS CRIADAS A PARTIR DE MODELOS LONGITUDINAIS BINÁRIOS E MULTINOMIAIS

JOSÉ ERASMO SILVA

UNIVERSIDADE METODISTA DE PIRACICABA (UNIMEP)

MARIA JOSÉ DE CAMARGO MACHADO

UNIVERSIDADE METODISTA DE PIRACICABA (UNIMEP)

MARIA IMACULADA DE LIMA MONTEBELLO

UNIVERSIDADE METODISTA DE PIRACICABA (UNIMEP)

JORGE LUIZ DOS SANTOS SILVA

ESTRATÉGIA DE INVESTIMENTO *LONG-SHORT* COM CARTEIRAS CRIADAS A PARTIR DE MODELOS LONGITUDINAIS BINÁRIOS E MULTINOMIAIS

1 INTRODUÇÃO

Estratégias de investimento sempre foram muito discutidas e o ponto central da discussão é a possibilidade de minimizar o risco e maximizar o retorno dos investimentos. Desde o artigo seminal de Markowitz (1952), que transformou drasticamente a forma de investir, migrando do conceito de investir todo o capital onde acreditava-se haver a maior probabilidade de retornos, para o conceito de investir de forma diversificada, muitos outros estudos surgiram com o objetivo de validar a proposta em diferentes ambientes econômicos. Apesar de inúmeras ferramentas terem surgido desde Markowitz (1952) a ideia da diversificação para minimizar riscos e proteger a riqueza do investidor continua atual.

Harvey et al (2019) estudaram diversos métodos de proteção no mercado americano, desde os tradicionais, como por exemplo, opções de vendas de longo prazo do índice S&P 500, passando por investimento em ouro, até chegar aos modelos alternativos como investimento em ações utilizando o método *long-short*. Os autores concluíram que os métodos tradicionais de proteção comprometem o resultado das carteiras pois corroem uma parcela considerável dos retornos e que os modelos *long-short* podem ser utilizados como uma carteira autoimune.

Com o objetivo de validar o poder de proteção e retorno das carteiras *long-short* novos modelos tem surgido e antigos modelos tem sido reformulados com o objetivo de selecionar as melhores ações para compor as carteiras (BEAVER; MCNICHOLS; PRICE, 2016), (MIKUTOWSKI; KAMBOURIS; ZAREMBA, 2019) e (ZAREMBA et al., 2019).

Devido a melhora na liquidez do mercado, assim como o aperfeiçoamento dos investidores a partir de publicações acadêmicas, a lucratividade de algumas estratégias de investimento tende a diminuir e até mesmo desaparecer (AVRAMOV et al., 2017). Nesse sentido, estratégias antigas devem se adaptar as novas realidades do mercado bem como novas estratégias devem ser desenvolvidas.

Desta forma questiona-se: é possível construir carteiras de investimento *long-short* a partir de modelos longitudinais binários e multinomiais e obter retornos maiores e menos arriscados que a carteira de mercado?

O objetivo principal do presente artigo é criar duas carteiras *long-short* a partir de dois modelos longitudinais, um binário e um multinomial, e verificar se são capazes de promover retornos maiores e com menor risco que a carteira teórica do Ibovespa.

Não foi encontrado em artigos anteriores a utilização de modelos longitudinais binários e multinomiais na construção de carteiras *long-short* e, sendo assim, o trabalho pretende fechar esta lacuna, contribuindo para o meio acadêmico e para o mercado.

Esses modelos, apesar de bastante úteis e de fácil aplicação, ainda são pouco utilizados em muitas áreas do conhecimento humano. Embora o desenvolvimento de softwares e o incremento da capacidade de processamento dos computadores tenham propiciado a sua aplicação de forma mais direta, muitos pesquisadores ainda desconhecem as suas utilidades e, sobretudo, as condições para que seu uso seja correto (FÁVERO; BELFIORE, 2017).

2 REFERENCIAL TEÓRICO

2.1 Carteiras de investimento

Segundo Markowitz (1952), deve-se considerar todos os investimentos como uma só carteira. Para ele o investidor deve possuir investimentos que, combinados entre si, formam uma carteira que apresente o maior retorno possível com o menor risco possível. Ainda segundo o autor existe uma relação direta entre o risco e o retorno e quanto maior o risco que o investidor estiver disposto a assumir maior poderá ser o retorno da sua carteira. Ele destaca

que a importância da formação de carteiras se dá principalmente pelo poder que esta tem de dissipar o risco tendo em vista que o risco de um ativo insulado é maior que o risco do mesmo ativo quando inserido em uma carteira.

Fama (2002) argumenta que o mercado é eficiente e que preço dos ativos refletem completamente todas as informações disponíveis e essas informações são processadas por agentes maximizadores de utilidade e minimizadores de risco. De acordo com Ehrhardt e Brigham (2012) a Hipótese de Mercado Eficiente (HME) afirma que as ações estão sempre em equilíbrio e que é impossível um investidor vencer o mercado e consistentemente obter uma taxa de retorno mais alta do que é justificado pelo risco das ações. Damodaran (2006) defende que o mercado não é eficiente e as estratégias de investimento visam explorar as ineficiências do mercado para obter vantagens sobre elas.

Partindo do princípio que o mercado na maioria das vezes não é eficiente os investidores buscam identificar as companhias que estão supervalorizadas ou subvalorizadas (ARTUSO; CHAVES NETO, 2012). O preço das ações de uma empresa depende do histórico de lucros e da expectativa dos lucros futuros. Um investidor inteligente está preparado para ler, entender e implementar uma estratégia de investimentos que incorpora as informações divulgadas pelas empresas (GUERARD; XU; WANG, 2019).

De acordo com Jacobs, Levy e Starer (1999) os investidores buscam na maioria das vezes ações subvalorizadas e que possam ser adquiridas para realização de ganhos futuros. Ainda segundo os autores existe um lado pouco explorado que é o de obter lucros com as ações supervalorizadas através da chamada venda a descoberto. Venda a descoberto é quando um investidor vende as ações emprestadas (ou alugadas visto que há uma taxa) com a expectativa de recompra-las mais barato futuramente (SIMON, 2019). Neste contexto o investidor obtém lucro com a queda das ações e ainda pode utilizar a receita obtida com a venda a descoberto para comprar ações subvalorizadas. Esta estratégia de investimento é conhecida como *long-short* e tem como vantagem aproveitar as duas direções de preços das ações.

Assumir uma posição vendida a descoberto introduz riscos, custo e atritos adicionais que não existem na posição comprada. O máximo que um investidor pode perder em uma posição comprada é 100%, enquanto que o retorno máximo é ilimitado. Com uma posição vendida acontece o contrário, pois o máximo que se pode ganhar é 100%, mas o máximo que se pode perder é ilimitado, pois no lado vendido a descoberto quanto mais a ação sobe mais o investidor perde. Além disso, os locatários podem solicitar a devolução das ações sem aviso prévio forçando os investidores a encerrarem a posição vendida em momentos não favoráveis (BEAVER; MCNICHOLS; PRICE, 2016)

De acordo com Jião, Massa e Zhang (2016) a estratégia *long-short* é também chamada de estratégia de zero investimento (*zero-investment strategy*) tendo em vista que o lado *short* vai servir de recursos para comprar o lado *long*. Para que se possa vender a descoberto é necessário que se dê garantias que podem ser desde saldo em conta como até mesmo investimentos em CDBs de liquidez diária (ZAREMBA et al., 2019).

Segundo Jacobs, Levy e Starer (1999) há três maneiras de implementar estratégias *long-short*, que são *Market-neutral*, *equitized* e estratégia *hedge*. A estratégia *Market-neutral*, mantém o lado *long* e o lado *short* em equilíbrio o tempo todo. Dessa forma elimina-se a exposição ao risco de mercado fazendo com que a carteira não seja afetada independente da direção do mercado. Nessa estratégia os lucros são obtidos a partir da diferença de desempenho entre os lados *long* e *short*. A estratégia *equitized* complementa a estratégia anterior adicionando um índice de mercado de ações na carteira de igual valor ao capital investido. Os lucros ou prejuízos são obtidos das variações das pontas *long* e *short* e também das variações no índice de ações. Finalmente, a estratégia *hedge* é similar as estratégias anteriores, porém é adicionado a carteira um índice futuro de ações que varia de acordo com

as perspectivas do mercado. Assim como a estratégia anterior, os lucros ou prejuízos são alcançados nos dois lados do *long/short* bem como nos retornos do índice.

As estratégias *long-short* são vistas como excelentes proteções em tempos de crises ou recessões. Em um estudo realizado por Harvey et al (2019) foram testadas importantes estratégias no mercado americano com o objetivo de verificar qual é mais eficiente para mitigar perdas durante tempos de crise ou de recessão. É notório que as estratégias de *hedge* são difíceis e caras. Sendo assim, além de testar as principais estratégias utilizadas com esse objetivo os autores testaram outras estratégias de investimento que podem exercer o mesmo papel de proteção. A primeira estratégia testada foi a utilização de opções de venda de longo prazo do índice S&P500. Esse é o mais confiável e defensivo método utilizado, porém pode custar até 7% dos retornos da carteira, o que a torna muitas vezes inviável. Outra ferramenta de proteção, conhecida como “Porto Seguro” (*safe-haven*) é a alocação de títulos do tesouro americano (*US Treasury bonds*). Além destas, em termos de confiabilidade e custo existem duas estratégias que podem ser inseridas entre as duas anteriores que são investimento em ouro e proteção de crédito (*credit protection*). Com o objetivo de buscar estratégias de menor custo os autores evoluíram os testes para mais duas estratégias diferentes. Primeira, estratégia de *momentum* de séries temporais que apresentou bom desempenho tanto em momentos de crise quanto de recessão. Segundo, estratégia *long-short* em ações, utilizando diferentes métricas de qualidade para classificar as empresas transversalmente. Os melhores resultados foram apresentados pelas estratégias *long-short*, que mostraram bom desempenho no período estudado (1985-2018), inclusive em momentos de crise e recessão.

Beaver e McNichols e Price (2016) estudaram as estratégias *long-short* de maneira diferente do que geralmente propõe os estudos acadêmicos. Os autores assumiram que o custo das carteiras não é zero e atribuíram um custo que foi descontado dos retornos das carteiras. Os resultados mostraram que o principal benefício das estratégias *long-short* é promover a diversificação em conjunto com outras estratégias. Isoladamente a estratégia *long-short* não foi capaz de bater o retorno do mercado.

Mikutowski, Kambouris e Zarembam (2019) estudaram 124 companhias listadas nos Emirados Árabes Unidos. Foram elaboradas carteiras *long-short* utilizando-se quatro diferentes variáveis preditivas de retorno representando o valor da empresa: preço sobre lucro (P/L), valor contábil sobre valor de mercado (B/M), *dividend yield* (DY) e EBITDA sobre valor de mercado (E/EV). Cada um desses indicadores foi classificado e dividido em tercis. As melhores classificações formaram o lado *long* da carteira, as piores classificações formaram o lado *short* da carteira. Os indicadores foram revistos e as carteiras reformuladas mensalmente pelo período de janeiro de 2004 a março de 2019. O indicador B/M trouxe o melhor resultado com média mensal de 1,66%.

A próxima seção apresenta a metodologia que será utilizada para atingir o objetivo do presente artigo.

METODOLOGIA

As informações para compor a amostra foram coletadas na base de dados Economática. Foram utilizadas as informações trimestrais das companhias listadas na B3 de 2010 a 2018. A utilização desse período justifica-se por compor diferentes cenários políticos e econômicos e com relativa estabilidade da moeda. Para evitar o viés de sobrevivência foram coletadas tanto informações de empresas ativas como de empresas inativas. Foram desconsiderados os custos de transações que, segundo Harvey et al., (2019), no contexto das estratégias *long-short* representam em torno de 1 a 2% ao ano no mercado americano. Foi desconsiderado também o imposto de renda (IR) sobre o ganho de capital.

Utilizou-se as informações consolidadas das empresas e, sendo assim, para que não houvesse mais de uma ação de uma mesma empresa em determinada carteira foram

escolhidas as ações mais líquidas das empresas. Além disso foi exigido ainda que a ação apresentasse no mínimo uma negociação dentro do trimestre.

Os *outliers* foram excluídos da amostra utilizando a proposta de Fávero e Belfiori (2017). Segundo os autores a detecção e exclusão de *outliers* é extremamente importante quando da aplicação de praticamente todas as técnicas de análise multivariadas de dados. No caso do presente artigo utilizou-se o algoritmo *Blocked Adaptive Computationally Efficient Outlier Nominators* com percentil de 0.15. Foram detectadas e eliminadas 65 observações classificadas como *outliers* multivariados.

Foram ainda eliminadas as variáveis que apresentaram problemas de multicolinearidade. Os modelos logísticos de regressão não disponibilizam um teste específico para esse fim, mas pode-se simular uma regressão linear utilizando as variáveis que serão utilizadas na regressão logística e em seguida obter a tolerância e VIF. (FIELD, 2009). Após os filtros restaram na amostra 251 empresas que apresentaram todos os dados necessários para as regressões e que geraram 5308 observações entre 2010 e 2018.

A opção de se trabalhar com informações trimestrais (ITR – Informações trimestrais) e carteiras reformuladas trimestralmente se deu pelo fato da influência que a divulgação dessas informações provoca no mercado. Há um consenso no mercado de que o investimento em ações é para longo prazo, no entanto, nada impede que elas sejam reavaliadas trimestralmente e que informações que possam mostrar uma oportunidade ou comprometer o retorno futuro demande uma ação imediata do investidor.

O formulário de informações trimestrais (ITR) deve ser divulgado no prazo máximo de 45 dias após o término do trimestre. Para o trimestre finalizado em 31/06, por exemplo, as informações devem ser divulgadas até 15/08 (COMISSÃO DE VALORES MOBILIÁRIOS (CVM), 1993). Dado este prazo os retornos, que foram ajustados aos proventos (dividendos e outros eventos corporativos), foram computados a partir deste prazo até um dia antes da divulgação do próximo ITR. A Tabela 1 exemplifica como foram considerados os retornos utilizados no estudo.

Tabela 1 – Exemplo da apuração dos retornos utilizados nas regressões

data demonstração	data máxima publicação	período de apuração do retorno
31/03/2017	15/05/2017	entre 16/05/2017 e 14/08/2017
30/06/2017	14/08/2017	entre 15/08/2017 e 14/11/2017
30/09/2017	14/11/2017	entre 15/11/2017 e 31/03/2018
31/12/2017	31/03/2018	entre 01/04/2018 e 15/05/2018

Fonte: elaborada pelo autor

Com todas as informações organizadas em planilhas eletrônicas o próximo passo do trabalho foi realizar os testes estatísticos para analisar primeiramente quais variáveis independentes explicam os retornos e posteriormente verificar qual a probabilidade destes retornos serem positivos, negativos ou estarem enquadradas em determinados intervalos de retornos.

Foram utilizados os modelos de regressão logística binária e multinomial. O modelo binário é utilizado quando a variável dependente é uma *dummy* que apresenta os valores zero ou um. O modelo multinomial é utilizado quando a variável dependente apresenta duas ou mais categorias, por exemplo ruim, regular, bom e ótimo.

No que se refere a variável *dummy* utilizada na regressão logística binária e que foi chamada de DUMMYRET, sempre que o retorno de uma ação em determinado período foi positivo esta variável recebeu o valor um, sempre que o retorno em determinado período for zero ou negativo ela recebeu o valor zero. Exemplo: digamos que determinada empresa apurou retorno de 10% entre 16/05/2019 e 14/08/2019, neste caso a *dummy* recebeu valor um. No mesmo sentido, digamos que no mesmo período outra empresa apresenta retorno de -5%,

a *dummy* desta empresa para este período recebeu valor zero. A partir de regressão logística binária deseja-se saber qual a probabilidade de DUMMYRET ser um no próximo trimestre.

Neste contexto, a carteira derivada do modelo binário foi criada da seguinte maneira: as ações com probabilidade de retornos positivos maior que 60% no próximo trimestre fizeram parte do lado *long*, as ações com probabilidade menor que 40% fizeram parte do lado *short*. O peso de cada ação na carteira é o mesmo para todas elas.

A respeito da variável QUART que foi utilizada na regressão logística multinomial, diferente do modelo logístico binário, o modelo logístico multinomial trabalhou com quartis, ou seja, dividiu-se em quartis por intervalos de retorno e através das variáveis independentes procurou-se demonstrar como cada variável influencia na probabilidade de determinada ação pertencer a um dos quartis. A Tabela 2 mostra como estão divididos os quartis.

Tabela 2 - Distribuição dos retornos por quartil

Variável	Obs.	Média	Desvio Padrão	Min.	Max.
QUART=1	1.327	-20,94	11,11	-87,44	-9,57
QUART=2	1.327	-4,00	2,95	-9,58	0,90
QUART=3	1.327	6,15	3,20	0,91	12,14
QUART=4	1.327	27,82	21,14	12,15	282,72

Valores em percentual

Fonte: Elaborado pelo autor com base na saída do sistema Stata

As ações com retorno entre -87,44 e -9,57 foram classificadas no primeiro quartil, as ações entre -9,58 e 0,90 foram classificadas no segundo quartil, as ações entre 0,91 e 12,14 foram classificadas no terceiro quartil e por fim as ações entre 12,15 e 282,72 foram classificadas no quarto quartil. Foi utilizado para fins de formação da carteira *long-short* o quarto quartil (ações com maior probabilidade de subir) e o primeiro quartil (ações com maior probabilidade de cair).

A regressão logística binária é uma forma especializada de regressão que é formada para prever e explicar uma variável categórica binária (dois grupos). A variável estatística representa uma relação multivariada com coeficientes como os da regressão indicando o impacto relativo de cada variável preditora (HAIR et al., 2009).

Nesse tipo de regressão há o interesse em avaliar a probabilidade p de ocorrência de um determinado evento com base no comportamento de variáveis explicativas. Sendo assim, sabendo-se que a chance de ocorrência de um evento é dada por $chance = (p/1-p)$, o modelo de regressão logística pode ser definido de acordo com o apresentado a seguir (FÁVERO; BELFIORI, 2017).

$$\ln(chance) = Z = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (1)$$

Que, ao desenvolver-se, chega-se a:

$$p = \frac{1}{1 + e^{-Z}} = \frac{1}{1 + e^{-(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}} \quad (2)$$

Em que:

Z: conhecido por *logit*;

p : probabilidade estimada de ocorrência do evento de interesse;

x_i : são as variáveis explicativas, com $i=1,2,\dots,k$; e

α e β_i : são os parâmetros do modelo.

Quando a variável dependente que representa o fenômeno em estudo oferece mais de duas possibilidades de resposta (categorias), deve-se fazer uso da regressão logística multinomial para estimar as probabilidades de ocorrência de cada alternativa (FÁVERO; BELFIORE, 2017).

No caso do presente estudo, conforme Tabela 2, a variável dependente apresenta quatro categorias possíveis de resposta (1, 2, 3 ou 4). A categoria de referência escolhida arbitrariamente foi a categoria 2. Desta forma, tem-se outras três possibilidades de evento em relação a esta categoria, que são as categorias 1, 3 e 4. Serão definidos então três vetores de variáveis explicativas com os respectivos parâmetros estimados, ou seja, três logitos:

$$Z_{i1} = \alpha_1 + \beta_{11} \cdot X_{1i} + \beta_{21} \cdot X_{2i} + \dots + \beta_{k1} \cdot X_{ki} \quad (3)$$

$$Z_{i3} = \alpha_1 + \beta_{13} \cdot X_{1i} + \beta_{23} \cdot X_{2i} + \dots + \beta_{k3} \cdot X_{ki} \quad (4)$$

$$Z_{i4} = \alpha_1 + \beta_{14} \cdot X_{1i} + \beta_{24} \cdot X_{2i} + \dots + \beta_{k1} \cdot X_{ki} \quad (5)$$

Com base nas expressões 3, 4 e 5 é possível estimar a probabilidade de ocorrência da categoria de referência 2 e as probabilidades de ocorrência dos três eventos distintos, representados pelas categorias 1, 3 e 4. Conforme Fávero e Belfiori (2017), as expressões podem ser escritas da seguinte forma:

Probabilidade de ocorrência da categoria 2 (referência):

$$p_{i2} = \frac{1}{1 + e^{Z_{i1}} + e^{Z_{i3}} + e^{Z_{i4}}} \quad (6)$$

Probabilidade de ocorrência da categoria 1:

$$p_{i1} = \frac{e^{Z_{i1}}}{1 + e^{Z_{i1}} + e^{Z_{i3}} + e^{Z_{i4}}} \quad (7)$$

Probabilidade de ocorrência da categoria 3:

$$p_{i3} = \frac{e^{Z_{i3}}}{1 + e^{Z_{i1}} + e^{Z_{i3}} + e^{Z_{i4}}} \quad (8)$$

Probabilidade de ocorrência da categoria 4:

$$p_{i4} = \frac{e^{Z_{i4}}}{1 + e^{Z_{i1}} + e^{Z_{i3}} + e^{Z_{i4}}} \quad (9)$$

A soma das probabilidades dos eventos será sempre 1.

Na forma complexa as expressões 6, 7, 8 e 9 podem ser escritas, respectivamente, da seguinte forma:

$$p_{io} = \frac{1}{1 + e^{(\alpha_1 + \beta_{11} \cdot X_{1i} + \beta_{21} \cdot X_{2i} + \dots + \beta_{k1} \cdot X_{ki})} + e^{(\alpha_2 + \beta_{12} \cdot X_{1i} + \beta_{22} \cdot X_{2i} + \dots + \beta_{k2} \cdot X_{ki})}} \quad (10)$$

$$p_{i1} = \frac{e^{(\alpha_1 + \beta_{11} \cdot X_{1i} + \beta_{21} \cdot X_{2i} + \dots + \beta_{k1} \cdot X_{ki})}}{1 + e^{(\alpha_1 + \beta_{11} \cdot X_{1i} + \beta_{21} \cdot X_{2i} + \dots + \beta_{k1} \cdot X_{ki})} + e^{(\alpha_2 + \beta_{12} \cdot X_{1i} + \beta_{22} \cdot X_{2i} + \dots + \beta_{k2} \cdot X_{ki})}} \quad (11)$$

$$p_{i2} = \frac{e^{(\alpha_2 + \beta_{12} \cdot X_{1i} + \beta_{22} \cdot X_{2i} + \dots + \beta_{k2} \cdot X_{ki})}}{1 + e^{(\alpha_1 + \beta_{11} \cdot X_{1i} + \beta_{21} \cdot X_{2i} + \dots + \beta_{k1} \cdot X_{ki})} + e^{(\alpha_2 + \beta_{12} \cdot X_{1i} + \beta_{22} \cdot X_{2i} + \dots + \beta_{k2} \cdot X_{ki})}} \quad (12)$$

$$p_{i3} = \frac{e^{(\alpha_3 + \beta_{13} \cdot X_{1i} + \beta_{23} \cdot X_{2i} + \dots + \beta_{k3} \cdot X_{ki})}}{1 + e^{(\alpha_1 + \beta_{11} \cdot X_{1i} + \beta_{21} \cdot X_{2i} + \dots + \beta_{k1} \cdot X_{ki})} + e^{(\alpha_2 + \beta_{12} \cdot X_{1i} + \beta_{22} \cdot X_{2i} + \dots + \beta_{k2} \cdot X_{ki})}} \quad (13)$$

Através das expressões apresentadas é possível calcular manualmente, através dos coeficientes estatisticamente significantes apresentados nas regressões, qual a probabilidade de cada um dos eventos representados nas variáveis dependentes ocorrer.

No que se refere às variáveis independentes, elas foram selecionadas a partir de estudos anteriores sobre o tema análise de investimentos, falência, carteiras de investimento e carteiras de investimento utilizando o modelo *long-short* (ALTMAN, 1968), (OHLSON, 1980), (FÁVERO; ALMEIDA, 2011), (SERRA; FÁVERO; MARTELANC, 2012), (ARTUSO; CHAVES NETO, 2012), (ROTELA JUNIOR; PAMPLONA; SALOMON, 2014), (SINHA; CHANDWANI; SINHA, 2015), (GUIMARÃES JÚNIOR; CARMONA; GUIMARÃES, 2015), (ALEXANDRINO et al., 2017), (MEDEIROS BARBOSA LEITE; FARIAS GUIMARÃES JÚNIOR, 2017), (MOREIRA; PENEDO, 2018), (ALBERTO et al., 2018), (PORTELA; SANTOS, 2018), (LYLE; YOHAN, 2019), (GUERARD; XU; WANG, 2019). Foram selecionadas 28 variáveis independentes, sendo elas: valor de mercado (LNVM), *dividend yield* (DY), preço sobre lucro (P/L), preço sobre vendas (P/V), preço sobre capital de giro líquido (P/CGL), retorno sobre o patrimônio líquido (ROE), retorno sobre o ativo (ROA), margem líquida (ML), variação do lucro (MCL), liquidez seca (LS), garantia do capital próprio ao capital de terceiros (GCPCT), composição do endividamento (CE), liquidez geral modificada (LGM), ativo (LNAT), valor patrimonial da ação (VPA), variação do VPA (CVPA), alavancagem (ALAV), receita (LNRECEITA), lucro líquido (LL), EBITDA, variação do EBITDA (CEBITDA), retorno do primeiro trimestre anterior (LNTRT1), retorno do segundo trimestre anterior (LNTRT2), retorno do terceiro trimestre anterior (LNTRT3), retorno do quarto trimestre anterior (LNTRT4), *dummy* que representa que os dados são do primeiro trimestre (_ITRI_1), *dummy* que representa que os dados são do segundo trimestre (_ITRI_2), *dummy* que representa que os dados são do terceiro trimestre (_ITRI_3), *dummy* que representa que os dados são do quarto trimestre (_ITRI_4), SELIC, DOLAR e IBOV. O objetivo das variáveis independentes é explicar o comportamento das variáveis dependentes.

RESULTADOS

Após rodar as regressões o modelo *stepwise* apresentou o melhor ajuste e também o melhor poder preditivo. Este método ajusta o modelo inserindo somente as variáveis com significância estatística definida pelo usuário. No caso foi definido o limite de 5% (sig.<0,05) para que as variáveis fizessem parte do modelo.

Tabela 3 - Regressão logística binária

Logistic regression	Number of obs	=	5,308
Wald chi2(10)	=	222.07	
Prob > chi2	=	0.0000	
Log pseudolikelihood	=	-3558.4337	
Pseudo R2	=	0.0314	

(Std. Err. adjusted for 251 clusters in id)

DUMMYRET	Coef.	Odds Ratio	Robust Std. Err.	z	P>z	[95% Conf.	Interval]
_ITRI_3	.3575618	1.429.839	.0668123	5.35	0.000	.2266121	.4885115
CP	-.3398913	.7118477	.1281143	-2.65	0.008	-.5909906	-.0887919
IBOV	-.0195192	.98067	.0039811	-4.90	0.000	-.0273221	-.0117164
LL	-4.20e-08	1	1.61e-08	-2.61	0.009	-7.36e-08	-1.05e-08
LNTRT1	.5618316	1.753.882	.1526233	3.68	0.000	.2626955	.8609677
LNTRT3	.8780714	2.406.255	.138627	6.33	0.000	.6063675	1.149.775
LNTRT4	.3166453	1.372.516	.1239295	2.56	0.011	.073748	.5595426
ROA	.0669644	1.069.257	.0110874	6.04	0.000	.0452334	.0886954
SELIC	.156396	1.169.289	.0240761	6.50	0.000	.1092076	.2035843
VPA	-.0006957	.9993046	.0002686	-2.59	0.010	-.0012221	-.0001692
_cons	-.2445594	.7830494	.0859873	-2.84	0.004	-.4130914	-.0760275

Fonte: saída do sistema Stata

Através da coluna *Odds Ratio* (taxa de probabilidade) é possível verificar qual o impacto das variáveis independentes no aumento ou diminuição da probabilidade de retornos positivos. Nota-se que a variável que traz maior impacto para a probabilidade de retorno positivo é a variável LNTRT3 (LN do retorno do terceiro trimestre anterior), ou seja, aumentos no logaritmo neperiano do retorno das ações três trimestres atrás aumentam em 2,40 vezes a probabilidade de retornos positivos no próximo trimestre, mantidas as demais condições constantes (*ceteris paribus*).

A variável que mais impacta negativamente é a variável CP (composição do endividamento). A composição do endividamento é calculada dividindo-se o passivo circulante pela soma do passivo circulante com o passivo não circulante (PC/(PC+PNC)). Quanto maior este indicador, maiores são os compromissos de curto prazo da empresa. Observa-se através da coluna *Odds Ratio* que o aumento de uma unidade da variável CP, representando que a empresa precisará de mais recursos no curto prazo, diminui em 29% (*Odds Ratio* 0,71) a probabilidade de retornos positivos no próximo trimestre, *ceteris paribus*.

Através dos coeficientes mostrados na Tabela 3 é possível calcular a probabilidade manualmente utilizando-se a expressão 2 apresentada na metodologia. No entanto, como o Stata calcula automaticamente serão utilizadas as probabilidades já calculadas para fins de seleção das carteiras. As ações com probabilidade de retorno positivo acima de 60% no próximo trimestre fazem parte do lado *long* da carteira e as ações com probabilidade de retornos positivos abaixo de 40% fazem parte do lado *short* da carteira.

Antes de apresentar os resultados da carteira é mostrado através da Tabela 4 o nível de classificação do modelo, ou seja, o que foi classificado como provável retorno positivo corretamente, o que foi classificado como improvável retorno positivo corretamente e o que foi classificado incorretamente.

Tabela 4 - Nível de classificação do modelo logístico binário

Modelo logístico de classificação para DUMMYRET				
Classificação	D	D~	Total	
+	1960	1343	3303	
-	814	1191	2005	
Total	2774	2534	5308	
Classificado como + se probabilidade de retorno $\geq 0,50$				
Sensibilidade			0,71	
Especificidade			0,47	
Eficiência Geral			0,59	

Fonte: elaborado pelo autor com base na saída do sistema Stata

De acordo com Fávero e Belfiori (2017) mais importante do que o pseudo R2 nos modelos logísticos é o nível de acerto das previsões. Observa-se que o nível de eficiência

geral do modelo foi de 59% que significa que de maneira geral, entre as previsões negativas e positivas, o modelo acertou 59% delas. A linha sensibilidade mostra que o modelo foi preciso em 71% das previsões positivas e a linha especificidade mostra que o modelo foi preciso em 47% dos retornos negativos. Nota-se que o modelo é mais preciso para prever retornos positivos futuros do que os retornos negativos.

Acredita-se que dois fatores podem melhorar o nível de acerto do sistema: (1) inserção de novas variáveis não utilizadas no modelo atual e (2) trabalhar com datas mais próximas da divulgação dos resultados das empresas. No presente artigo trabalhou-se com as datas máximas de publicação, no entanto, a maioria das empresas publica seus relatórios antes da data limite.

Através da Tabela 5 são apresentados os retornos médios trimestrais da carteira *long-short* e Ibovespa. Foi simulado um investimento com saldo inicial de 50 mil para o lado *long* da carteira, 50 mil para o lado *short* e 100 mil para o Ibovespa.

Tabela 5 - Retorno trimestral das carteiras formadas a partir da regressão logística binária

	<i>logit long cutoff 0,60</i>	<i>logit short cutoff 0,40</i>	Total <i>long-short</i>	Ibovespa
SD	13,22	17,01	12,30	8,17
RET	1074,69	100,03	587,36	36,52
MÉDIA	8,49	3,74	6,28	1,20
SHARP	64,22	21,97	51,05	14,68
AÇÕES	35	15		
SALDO	587,34	100,01	687,36	136,52

DP = Desvio Padrão. RET = Retorno em percentual. MÉDIA = média trimestral do retorno. SHARP = retorno ajustado ao risco. Saldo = Valores do investimento em milhares de Reais. Saldo Total = Soma dos saldos da carteira *long-short*.

Fonte: Elaborada pelo autor

O lado *long* da carteira apresentou desvio padrão de 13,22%, retorno total no período de 1074%, média trimestral de 8,49% e retorno ajustado ao risco (Sharp) de 64,22%. Simulou-se um investimento de 50 mil reais no primeiro trimestre de 2010 que ao final do período, no quarto trimestre de 2018, acumulou o montante de 587 mil reais.

O mesmo procedimento foi feito para o lado *short* da carteira onde obteve-se desvio padrão de 17,01%, retorno total no período de 100%, retorno médio de 3,74% e retorno ajustado ao risco de 21,97%. O resultado substancialmente menor no lado *short* se deu pelo baixo poder de previsão dos resultados negativos apresentados pelo modelo na Tabela 4. Simulou-se também para o lado *short* um investimento de 50 mil reais no primeiro trimestre de 2010 que ao final do período, no quarto trimestre de 2018, acumulou o montante de 100 mil reais. Sendo assim, o retorno total da carteira *long-short* é a soma dos dois lados, ou seja, 687 mil reais. Referente à quantidade de ações alocadas para cada lado da carteira, a carteira ficou com média de 35 ações no lado *long* e média de 15 ações no lado *short*.

Sobre o resultado da carteira *long-short* de forma integrada, nota-se que o risco foi minimizado apresentando desvio padrão de 12,30, no entanto, o retorno ajustado ao risco não foi melhor que o lado *long* de forma isolada, apresentando 51,05%. Vale lembrar que uma das vantagens da carteira *long-short* é que o lado *short* pode servir de recursos para comprar o lado *long*.

Simulou-se ainda um investimento de 100 mil reais no Ibovespa. A carteira teórica apresentou desvio padrão de 8,17%, retorno total de 36,52%, média trimestral de 1,20% e retorno ajustado ao risco de 14,68%. O retorno total da carteira teórica no período foi de 136 mil reais. Observa-se que o risco do Ibovespa, quando mensurado através do desvio padrão é menor que a carteira *long-short*, no entanto, como os retornos também são substancialmente menores fez com que o retorno ajustado ao risco também fosse menos interessante que a carteira *long-short*. Com o objetivo de verificar se a diferença entre as médias da carteira

long-short e as médias da carteira teórica Ibovespa são estatisticamente significantes utilizou-se o teste Mann-Whitney. Os resultados são apresentados na Tabela 6.

Tabela 6 – Teste Mann-Whitney de diferença de médias

Carteira	obs	rank sum	expected
(1) long-short binário	36	1543	1314
(2) Ibovespa	36	1085	1314
combined	72	2628	2628
unadjusted variance	7884.00		
adjustment for ties	0.00		
adjusted variance	7884.00		
Ho: Ret(cod==1) = Ret(cod==2)			
z = 2.579			
Prob > z = 0.0099			

Fonte: saída do sistema Stata

As médias da carteira *long-short* são estatisticamente superiores do que as médias do índice. Através da coluna *rank sum* apresentado na Tabela 6 nota-se que o posto de média da carteira *long-short* foi 50% maior que o posto de média da carteira Ibovespa. Desta forma, com Prob > |z| = 0.0099, rejeita-se a hipótese nula de que a diferença entre os retornos médios das carteiras é zero.

A Tabela 7 apresenta o modelo logístico multinomial, que conforme mencionado anteriormente trabalhou com quartis, ou seja, dividiu-se em quartis as possibilidades de retorno e através das variáveis independentes procurou-se demonstrar como cada variável influencia na probabilidade de determinada ação pertencer a um dos quartis.

Tabela 7 - Regressão logística multinomial

QUART	Var.	Coef.	Odds Ratio	Robust Std. Err.	z	P>z	[95% Conf. Interval]
1							
	ALAV	.0458129	1.046878	.0133606	3.43	0.001	.0196265 .0719992
	CVPA	.8160504	2.26155	.3029938	2.69	0.007	.2221935 1.409.907
	DOLAR	-.0120639	.9880086	.0061605	-1.96	0.050	-.0241382 .0000104
	IBOV	-.0023037	.9976989	.0057143	-0.40	0.687	-.0135036 .0088961
	LL	1.07e-08	1	4.61e-08	0.23	0.816	-7.96e-08 1.01e-07
	LNRECEITA	.0375165	1.038229	.0437661	0.86	0.391	-.0482634 .1232964
	LNTRT1	-.1310584	.8771666	.1983591	-0.66	0.509	-.519835 .2577183
	LNTRT2	-.7020453	.4955707	.1984034	-3.54	0.000	-1.090909 -.3131818
	LNTRT3	-.862613	.4220578	.2104326	-4.10	0.000	-1.275053 -.4501728
	LNTRT4	-.1844128	.8315925	.1777036	-1.04	0.299	-.5327054 .1638798
	LNVM	-.1538845	.8573711	.0388884	-3.96	0.000	-.2301043 -.0776646
	LS	-.0624006	.9395064	.029462	-2.12	0.034	-.1201451 -.0046562
	PSCGL	-.0002135	.9997865	.0001537	-1.39	0.165	-.0005148 .0000878
	ROA	-.0640037	.9380015	.0182562	-3.51	0.000	-.0997852 -.0282222
	SELIC	-.2302543	.7943316	.0405833	-5.67	0.000	-.3097961 -.1507125
	VPA	.0002631	1.000.263	.0001153	2.28	0.023	.0000371 .000489
	_cons	2.308807	1.006.241	.45107	5.12	0.000	1.424.726 3.192.888
2		(base		outcome)			
3							
	ALAV	.0096362	1.009683	.0135342	0.71	0.476	-.0168903 .0361627
	CVPA	.8806503	2.412468	.2923773	3.01	0.003	.3076013 1.453.699
	DOLAR	-.0130794	.9870058	.006331	-2.07	0.039	-.025488 -.0006709

IBOV	-.0269135	.9734455	.0059517	-4.52	0.000	-.0385786	-.0152484
LL	-5.25e-08	.9999999	3.56e-08	-1.48	0.140	-1.22e-07	1.72e-08
LNRECEITA	-.0596282	.9421147	.0407849	-1.46	0.144	-.1395652	.0203087
LNTRT1	.8513139	2.342723	.2126534	4.00	0.000	.434521	1.268.107
LNTRT2	-.4065568	.6659393	.1990342	-2.04	0.041	-.7966568	-.0164569
LNTRT3	.1916171	1.211207	.2157974	0.89	0.375	-.231338	.6145721
LNTRT4	.1444279	1.155378	.1819602	0.79	0.427	-.2122075	.5010632
LNVM	.0740134	1.076821	.0411211	1.80	0.072	-.0065825	.1546093
LS	-.0939621	.9103173	.0293313	-3.20	0.001	-.1514505	-.0364737
PSCGL	.0002622	1.000262	.000121	2.17	0.030	.0000251	.0004993
ROA	.033667	1.03424	.0188506	1.79	0.074	-.0032796	.0706135
SELIC	.1004232	1.105639	.0417406	2.41	0.016	.0186132	.1822332
VPA	-.0006305	.9993697	.0003308	-1.91	0.057	-.0012788	.0000179
_cons	-.433385	.6483108	.4020187	-1.08	0.281	-1.221327	.3545572
4							
ALAV	.0298832	1.030334	.0163126	1.83	0.067	-.0020889	.0618553
CVPA	.5044637	1.656097	.2914257	1.73	0.083	-.0667201	1.075.647
DOLAR	.0113244	1.011389	.0055761	2.03	0.042	.0003955	.0222533
IBOV	-.0096529	.9903935	.0056108	-1.72	0.085	-.0206498	.001344
LL	-4.25e-08	1	1.70e-08	-2.50	0.012	-7.57e-08	-9.18e-09
LNRECEITA	.1654067	1.179873	.0478943	3.45	0.001	.0715356	.2592779
LNTRT1	.0859321	1.089732	.2141879	0.40	0.688	-.3338684	.5057326
LNTRT2	-.6636752	.5149553	.2101035	-3.16	0.002	-1.075471	-.2518798
LNTRT3	.5745363	1.776307	.2026687	2.83	0.005	.1773129	.9717597
LNTRT4	.3939243	1.482788	.2161568	1.82	0.068	-.0297352	.8175838
LNVM	-.2407143	.7860662	.0431395	-5.58	0.000	-.3252662	-.1561624
LS	-.0022653	.9977373	.029885	-0.08	0.940	-.0608388	.0563082
PSCGL	-.0000363	.9999637	.0003293	-0.11	0.912	-.0006817	.000609
ROA	.0208003	1.021018	.0183147	1.14	0.256	-.015096	.0566965
SELIC	-.0386188	.9621174	.0396191	-0.97	0.330	-.1162708	.0390332
VPA	-.0003836	.9996165	.000247	-1.55	0.120	-.0008678	.0001006
_cons	1.328823	3.776594	.4790411	2.77	0.006	.3899194	2.267.726

Variáveis sem significância estatística não apresentadas.

Fonte: elaborado pelo autor com base na saída do Stata

Assim como a modelagem anterior, para a presente modelagem foi utilizado o método *stepwise*, que ajusta o modelo inserindo somente as variáveis com significância estatística definida pelo usuário. No caso foi definido o limite de 5% (sig.<0,05) para que as variáveis fizessem parte do modelo.

Apesar de existirem várias variáveis que influenciam a probabilidade de a ação pertencer a um ou outro quartil, será comentada apenas a atuação das mais influentes. Para primeiro quartil (observações com retorno entre -87,44 e -9,57) a variável que mais influenciou positivamente foi a CVPA (crescimento do Valor Patrimonial por Ação) que apresentou *Odds Ratio* de 2,26. Isso significa que aumentos nesta variável aumentam a probabilidade de uma ação se enquadrar no primeiro quartil. Por outro lado, a variável LNTRT3 influenciou negativamente mostrando que, mantidos os demais fatores constantes, aumentos nesta variável diminuem em 58% (*Odds Ratio* = 0,4220578) a probabilidade da ação se enquadrar no primeiro quartil.

O segundo quartil (observações com retorno entre -9,58 e 0,90) foi atribuído como base e desta forma não são apresentados os coeficientes para ele. O Stata adota como categoria de referência a que apresenta maior frequência de observações, no entanto como a quantidade de observações é a mesma para todos os quartis foi atribuído arbitrariamente o segundo quartil como base.

Quanto ao terceiro quartil (observações com retorno entre 0,91 e 12,14) duas variáveis influenciaram positivamente de maneira mais robusta que foram as variáveis CVPA (*Odds Ratio* = 2.412468) e LNTRT1 (*Odds Ratio* = 2.342723), ou seja, o aumento de uma unidade

na variável CVPA aumenta em 2,41 vezes a probabilidade de a ação pertencer ao terceiro quartil, *ceteris paribus*. O aumento de uma unidade na variável LNTRT1 aumenta em 2,34 vezes a probabilidade de a ação pertencer ao terceiro quartil, *ceteris paribus*.

O último quartil (observações com retorno entre 12,15 e 282,72) apresentou como variável mais relevante positivamente a LNTRT3 (*Odds Ratio* = 1.776307). Aumentos nesta variável, mantidas as demais constantes, representam aumento de 77% na probabilidade da ação se enquadrar no quarto quartil. Por outro lado, a variável LNTRT2 mostra influência negativa (*Odds Ratio* = 0.5149553), ou seja, aumentos nesta variável diminuem a probabilidade da ação se enquadrar no quarto quartil.

Através dos coeficientes mostrados na Tabela 7 é possível calcular a probabilidade manualmente utilizando-se as expressões 10, 11, 12 e 13 apresentadas na metodologia. No entanto, como o Stata calcula automaticamente serão utilizadas as probabilidades já calculadas para fins de seleção das carteiras. As ações com maior probabilidade de pertencer ao quartil 1 formaram o lado *short* da carteira, enquanto que as ações com maior probabilidade de pertencer ao quartil 4 formaram o lado *long* da carteira.

Assim como destacado na regressão logística binária mais importante do que o pseudo R2 é o nível de classificação do modelo. A Tabela 8 apresenta a comparação do que foi classificado corretamente para cada um dos quartis.

Tabela 8 - Nível de classificação do modelo logístico multinomial

Modelo logístico de classificação para QUART						
Classificação Real	Classificação Prevista				Total	
	1	2	3	4		
1	553	228	270	276	1327	
2	308	295	458	266	1327	
3	274	203	584	266	1327	
4	343	215	357	412	1327	
Total	1478	941	1669	1220	5308	
Classificação Correta por Quartil (%)						
1	37,42					
2	31,35					
3	34,99					
4	33,77					

Fonte: elaborado pelo autor com base na saída do sistema Stata

Nota-se que o nível de acerto é baixo, em especial no que mais interessa para o presente trabalho, que é a classificação do quartil 1 e 4, obteve-se nível de acerto de 37,42% e 33,77%, respectivamente. É um nível de acerto abaixo do ideal pois olhando de outra perspectiva pode se afirmar que mais de 60% das classificações estão erradas.

A Tabela 9 apresenta o retorno trimestral das carteiras formadas a partir do modelo logístico multinomial. Nota-se a princípio que tanto o lado *long* quanto o lado *short* da carteira formada a partir deste modelo trouxeram menor desvio padrão, que representa menor risco a partir dessa perspectiva.

Tabela 9 - Retorno trimestral das carteiras formadas a partir da regressão logística multinomial

	mlogit long quart = 4	mlogit short quart = 1	Total long-short	Ibovespa
SD	11,37	13,95	6,83	8,175
RET	175,56	-17,50	79,03	36,52
MÉDIA	3,48	-0,53	1,87	1,205
SHARP	30,57	-3,83	27,32	14,73
AÇÕES	34	41		
SALDO	137,78	41,25	179,03	136,52

SD = Desvio Padrão. RET = Retorno em percentual. MÉDIA = média trimestral do retorno. SHARP = retorno ajustado ao risco. Saldo = Valores do investimento em milhares de Reais. Saldo Total = Soma dos saldos da carteira *long-short*. Pesos iguais para todas as ações.

Fonte: Elaborada pelo autor

Observa-se que a quantidade de ações alocadas no lado *long* da carteira baseada no modelo multinomial foi equivalente às alocadas no modelo binário, ou seja, média de 34 ações. No entanto, o lado *short* ficou com quantidade bastante superior, sendo que a média de 15 do modelo binário subiu para média de 41 no modelo multinomial. Quanto maior a quantidade de ações na carteira maiores são os custos de transação.

Referente aos retornos, o modelo trouxe menores retornos absolutos e também menores retornos ajustados ao risco quando comparado ao modelo binário. Entretanto, o modelo apresentou melhor desempenho que o Ibovespa, tanto no retorno absoluto quanto no retorno ajustado ao risco. A carteira apresentou retorno total de 79,03% no período com média de 1,87% ao trimestre. Analisando os lados de maneira separada nesse caso seria mais interessante manter somente o lado *long* tendo em vista que apresentou maior retorno (175%) e menor risco (11,37%) que o lado *short* que apresentou perda de 17,50% e maior risco (13,95%). Quanto ao índice de Sharp a carteira apresenta 27,32% de retorno ajustado ao risco, sendo 30,57% no lado *long* e -3,83% no lado *short*. O Ibovespa apresentou 14,73% de Sharp no período.

Da mesma forma que no modelo anterior, também foi simulado um investimento de 50 mil no lado *long*, 50 mil no lado *short* e 100 mil no Ibovespa que apresentaram saldo total acumulado de 179,03 mil para a carteira *long-short* e 136,52 mil para o Ibovespa.

A Tabela 10 apresenta o teste Mann-Whitney que faz a comparação entre as médias da carteira *long-short* com a carteira teórica do Ibovespa.

Tabela 10 – Teste Mann-Whitney de diferença de médias

Carteira	obs	rank sum	expected
(1) long-short multinomial	36	1342	1314
(2) Ibovespa	36	1286	1314
combined	72	2628	2628

unadjusted variance 7884.00
 adjustment for ties 0.00
 adjusted variance 7884.00
 Ho: Ret(cod==1) = Ret(cod==2)
 z = 0.315
 Prob > |z| = 0.7525

Fonte: saída do sistema Stata

A carteira *long-short* apresentou maior retorno e menor risco que o Ibovespa, no entanto essa diferença não é estatisticamente significativa, conforme mostra o teste de diferença de médias. Os postos de médias ficaram muito próximos sendo 1342 para a carteira *long-short* e 1286 para a carteira teórica. Desta forma, com Prob > |z| = 0.7525, não foi possível rejeitar a hipótese nula de que há diferença entre as carteiras.

Além dos testes realizados anteriormente, ambas as regressões passaram por testes de qualidade dos ajustes dos modelos. Testar a qualidade do ajuste é um importante passo dos modelos estatísticos. Para verificação da qualidade de ajuste em regressões logísticas binárias um teste frequentemente utilizado é o teste de Hosmer-Lameshow (*estat gof*). No entanto para as regressões logísticas multinomiais ainda não existe um teste nativo no sistema Stata, então optou-se por utilizar o teste *mlogitgof*, que é um *user-written* desenvolvido por Fagerland e Hosmer (2012). Os testes fazem a comparação entre o resultado de fato e o resultado estimado. Espera-se que não seja possível rejeitar a hipótese nula, ou seja, espera-se que não exista diferença estatisticamente significativa entre o resultado real e o resultado estimado.

Dada a significância destes testes, sendo sig. 0.3910 para o modelo binário e sig. 0.223 para o modelo multinomial, não foi possível rejeitar a hipótese nula do teste. Desta forma é possível afirmar que o teste não tem problemas de ajuste, ou seja, não há diferenças estatisticamente diferentes entre o resultado estimado e o resultado observado nas regressões.

Os resultados confirmam a maioria dos achados anteriores que mostraram que as carteiras *long-short* possibilitam a oportunidade de se obter retornos maiores do que as carteiras de mercado, mas não maiores que a carteira *long-only*. Assim, entende-se que o modelo ainda carece de ajustes. Estão de acordo ainda com os argumentos de Beaver e McNichols e Price (2016) que afirmam que as estratégias *long-short* apresentam maior risco, tendo em vista que o modelo binomial mostrou desvio padrão maior que o Ibovespa, contrariando os achados de Harvey et al., (2019).

Conclui-se, portanto, que é possível construir carteiras de investimento *long-short* a partir dos modelos longitudinais binários e multinomiais, sendo que o modelo binário apresentou maior retorno, porém maior risco que o mercado e o modelo multinomial se mostrou mais seguro e mais rentável, mas sem validação da diferença nas médias.

Para os trabalhos futuros sugere-se a consideração dos custos de transações, compras, vendas e especialmente os custos de aluguel no caso das ações *short*. É importante considerar ainda o Imposto de Renta sobre os ganhos de capital durante a revisão das carteiras. Não incluir esses custos foi uma limitação para o trabalho.

CONSIDERAÇÕES FINAIS

Através de dois modelos longitudinais de regressão foram construídas duas carteiras de investimento que foram renovadas trimestralmente. A carteira construída a partir do modelo binário apresentou desvio padrão de 12,30%, retorno total de 587,36%, média trimestral de 6,28% e índice Sharpe de 51,05%. A carteira construída a partir do modelo multinomial apresentou 6,83% de desvio padrão, 79,90% de retorno total, média trimestral de 1,87% e índice Sharpe de 27,32%. Para o mesmo período a carteira teórica Ibovespa apresentou 8,17% de desvio padrão, 36,52 de retorno total, 1,20% de média trimestral e 14,68% de índice Sharpe. Nota-se que o modelo binário trouxe resultado substancialmente maior que o Ibovespa no período, inclusive com médias trimestrais estatisticamente diferentes ao nível de 1%. Para o modelo multinomial, apesar de apresentar maior retorno e menor risco que o Ibovespa essa diferença não foi estatisticamente significativa.

Este artigo dá uma parcela de contribuição para a estabilidade do mercado financeiro, fornecendo um novo modelo para investimento em ações e, desta forma, fazendo com o que haja mais alternativas para diversificar as formas de investir. Conforme apontado por Damodaran (2006) e Avramov et al., (2017), quando muitos investidores se concentram nos mesmos modelos de investimento a tendência é que com o tempo os retornos se anulem e também que momentos de crise ou recessão sejam potencializados. O trabalho contribui ainda com o meio acadêmico, incrementando os estudos a respeito de técnicas de investimento.

REFERÊNCIAS

ALBERTO, J. G. C. et al. Análise do desempenho de uma carteira construída seguindo as ideias de investimento de Greenblatt. p. 157–169, 2018.

ALEXANDRINO, F. Q. DE L. et al. Utilização da análise envoltória de dados (DEA) na composição de carteira de investimento diversificada e eficiente. **Revista Produção Online**, v. 17, n. 2, p. 507, 2017.

ALTMAN, E. I. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. **The Journal of Finance**, v. 23, n. 4, p. 589–609, 1968.

ARTUSO, A. R.; CHAVES NETO, A. Identificação de Fatores Relevantes e Construção de Portfólio Diversificado no Mercado Acionário Brasileiro. **Produção Online**, v. 12, n. 4, p. 1002–1030, 2012.

AVRAMOV, D. et al. Scaling Up Market Anomalies. **The Journal of Investing**, v. 26, n. 3, p. 89–105, 2017.

BEAVER, W.; MCNICHOLS, M.; PRICE, R. The costs and benefits of long-short investing: A perspective on the market efficiency literature. **Journal of Accounting Literature**, v. 37, p. 1–18, 2016.

COMISSÃO DE VALORES MOBILIÁRIOS (CVM). Texto integral da instrução cvm nº 202. . 1993.

DAMODARAN, Aswath. Filosofias de investimento: estratégias bem-sucedidas e os investidores que a fizeram funcionar. Rio de Janeiro. Qualitymark, 2006.

EHRHARDT, M. C; BRIGHAM, E. F. **Administração Financeira: Teoria e Prática**. São Paulo: Cengage Learning, 2012.

FAGERLAND, M. W.; HOSMER, D. W. A generalized Hosmer-Lemeshow goodness-of-fit test for multinomial logistic regression models. **Stata Journal**, v. 12, n. 3, p. 447–453, 2012.

FAMA, E. F. Market efficiency, long-term returns, and behavioral finance | The comments of Brad Barber, David Hirshleifer, S.P. Kothari, Owen Lamont, Mark Mitchell, Hersh Shefrin, Robert Shiller, Rex Sinquefeld, Richard Thaler, Theo Vermaelen, Robert Vishny, Ivo Welch. **Journal of Financial Economics**, v. 49, n. 3, p. 283–306, 2002.

FÁVERO, L. P. L.; ALMEIDA, J. E. F. DE. Modelo de painel logit para avaliação de retornos positivos em mercados acionários. **XLIII Simpósio Brasileiro de Pesquisa Operacional**, p. 922–934, 2011.

FÁVERO, Luiz Paulo; BELFIORE, Patrícia. Manual de análise de dados: estatística e modelagem multivariada com Excel®, SPSS® e Stata®. Elsevier Brasil, 2017.

FIELD, A. Descobrindo a estatística usando o SPSS - 2.ed. **Descobrindo a Estatística usando o SPSS**, 2009.

GUERARD, J. B.; XU, G.; WANG, Z. **Portfolio and Investment Analysis with SAS: Financial Modeling Techniques for Optimization**. Cary, NC: Inc., SAS Institute, 2019.

GUIMARÃES JÚNIOR, F. R. F.; CARMONA, C. U. DE M.; GUIMARÃES, L. G. DE A. Carteiras Formadas Por Meio De Variáveis Fundamentalistas Apresentam Bom Desempenho De Mercado? **Gestão & Regionalidade**, v. 31, n. 91, 2015.

HAIR, J. et al. Análise multivariada de dados. 2009.

HARVEY, C. R. et al. The Best of Strategies for the Worst of Times: Can Portfolios be Crisis Proofed? **The Journal of Portfolio Management**, 2019.

JACOBS, B. I.; LEVY, K. N.; STARER, D. Long-short portfolio management: An integrated approach. **The Journal of Portfolio Management**, p. 23–32, 1999.

JIAO, Y.; MASSA, M.; ZHANG, H. Short selling meets hedge fund 13F: An anatomy of informed demand. **Journal of Financial Economics**, v. 122, n. 3, p. 544–567, 2016.

ROTELA JUNIOR, P. R.; PAMPLONA, E. D. O.; SALOMON, F. L. R. Otimização de portfólios: análise de eficiência. **Revista de Administração de Empresas | FGV-EAESP**, p. 405–413, 2014.

LYLE, M. R.; YOHN, T. L. Optimized Fundamental Portfolios. **SSRN Electronic Journal**, 2019.

MARKOWITZ, H. Portfolio Selection. **The Journal of Finance**, v. 7, n. 1, p. 77–91, 1952.
MEDEIROS BARBOSA LEITE, U.; FARIAS GUIMARÃES JÚNIOR, F. R. Selecionando portfólio de ativos utilizando variáveis fundamentais no mercado brasileiro. **Revista Eletrônica do Mestrado Profissional em Administração**, v. 9, n. 1, p. 62–72, 2017.

MIKUTOWSKI, M.; KAMBOURIS, G. D.; ZAREMBA, A. A note on value investing in the UAE stock market. **Journal of Research in Emerging Markets**, v. 1, n. 2, p. 33–38, 2019.

MOREIRA, K. D. S.; PENEDO, A. S. T. Seleção de Portfólios: Uma análise comparativa dos cinco fatores de Fama e French e Redes Neurais Artificiais. **Enfoque: Reflexão Contábil**, v. 37, n. 2, p. 141, 2018.

OHLSON, J. A. Financial Ratios and the Probabilistic Prediction of Bankruptcy. **Journal of Accounting Research**, v. 18, n. 1, p. 24, 1980.

PORTELA, D. L. M.; SANTOS, J. F. DOS. RISCO MACROECONÔMICO E O MODELO DE CINCO FATORES NO MERCADO ACIONÁRIO BRASILEIRO. **REAd. Revista Eletrônica de Administração (Porto Alegre)**, v. 24, n. 3, p. 269–293, 2018.

SERRA, R. G.; FÁVERO, L. P. L.; MARTELANC, R. Stock Returns, Efficiency of Beta and the Probability to Grow at an Above-Average Rate Relative to the Market: Evidence from a Logit Model. **American International Journal of Contemporary Research**, v. 2, n. January 2016, p. 7–14, 2012.

SIMON, S. **DEVELOPING A NEW STOCK SCORING MODEL FOR SHARIAH - COMPLIANT INVESTMENT**. [s.l.] (Tese de Doutorado) - University of Bolton, 2019.

SINHA, P.; CHANDWANI, A.; SINHA, T. Algorithm of construction of optimum portfolio of stocks using genetic algorithm. **International Journal of Systems Assurance Engineering and Management**, v. 6, n. 4, p. 447–465, 2015.

ZAREMBA, A. et al. Picking Winners to Pick Your Winners: The Momentum Effect in Commodity Risk Factors. **The North American Journal of Economics and Finance**, p. 37, 2019.