

## MARKET CRASH, CARACTERÍSTICAS DAS EMPRESAS E RETORNO – UMA ANÁLISE LOGÍSTICA E DISCRIMINANTE

### MARKET CRASH, FIRM CHARACTERISTICS AND MARKET RETURN – A LOGISTIC AND DISCRIMINANT ANALYSIS

Ricardo Goulart Serra  
[ricardo.serra@usp.br](mailto:ricardo.serra@usp.br)  
USP

Eduardo Augusto Contani  
[contani@usp.br](mailto:contani@usp.br)  
USP

Roy Martelanc  
[rmartela@usp.br](mailto:rmartela@usp.br)  
USP

José Roberto Savoia  
[jrsavoia@usp.br](mailto:jrsavoia@usp.br)  
USP

Maria Aparecida Gouvêa  
[magouvea@usp.br](mailto:magouvea@usp.br)  
USP

#### Resumo

O objetivo do presente artigo é verificar se existem características comuns das empresas que possam discriminar seus retornos em um dia de *market crash*, em dois grupos: o daquelas que perderam mais do que o mercado e o daquelas que perderam menos do que o mercado. O artigo analisou as empresas listadas na NYSE no dia 15 de outubro de 2008, dia de maior queda do S&P 500 dos últimos 23 anos (-9,0%). No total, 461 empresas foram analisadas. Por meio da regressão logística e da análise discriminante, identificou-se que o retorno defasado de 2 meses e o desvio padrão de 1 ano ajudam a discriminar entre os dois grupos. As empresas com maior retorno defasado e com menor desvio padrão apresentaram menor chance de perderem mais do que o S&P 500. As precisões de acerto dos modelos logístico e discriminante alcançaram 71,6% e 72,0% respectivamente, superiores aos critérios da probabilidade máxima, probabilidade proporcional e Press's Q. De forma análoga, o teste da variável beta para discriminação resultou num nível de precisão de 61,4%, inferior aos resultados dos modelos. Este artigo contribui para o estudo das características que influenciam o retorno das empresas em um momento de crise, situação que carece de maiores estudos e aprofundamento.

**Palavras-chave:** *Market Crash*, Crise 2008, Mercado de Capitais, Regressão Logística, Análise Discriminante.

#### Abstract

*We aim to verify whether common firm characteristics can discriminate their returns on a market crash day in two groups: those who have lost more than the market and those who lost less than the market. We analyzed NYSE listed companies on October 15<sup>th</sup>, 2008, the largest drop day in the S&P 500 index over the last 23 years (-9.0%). We analyzed 461 companies. Through logistic regression and discriminant analysis, we identified that 2-months lagged return and 1-year standard deviation can discriminate the two groups. Firms with higher lagged return and less standard deviation were less likely to lose more than the S&P 500. Accuracy level of logistic and discriminant models reached 71.6% and*

72.0% respectively, above the criteria of maximum likelihood, probability proportional and Press's  $Q$ . Similarly, the variable beta test for discrimination resulted in an accuracy level of 61.4%, below those of the models. This article contributes to the study of characteristics that influence the return of companies in a crisis period, field that requires more studies.

**Keywords:** Market Crash, 2008 Crisis, Capital Markets, Logistic Regression, Discriminant Analysis.

Artigo recebido em: 13.11.2013; Aceito em: 07.07.2014

## 1. INTRODUÇÃO

O estudo das características das empresas que explicam o retorno de suas ações em tempos normais tem sido objeto de análise de diversos autores, dentre eles: Ball (1978), Banz (1981), Amihud e Mendelson (1986), Bhandari (1988), Fama e French (1992), Kothari, Shanken e Sloan (1995), Berk, Green e Naik (1999), Pástor e Veronesi (2003), Jiang e Lee (2007), Phalippou (2007) e Chen e Zhang (2010). Em tempos de crise, os estudos normalmente focam outros temas, por exemplo: (i) a análise de volatilidade antes, durante e depois da ocorrência, (ii) o co-movimento e (iii) contágio de países (Arshanapalli e Doukas, 1993; e Bunda, Hamann e Lall, 2009; Cheung, 2000; Hon, Strauss e Yong, 2004; Michayluk e Neuhauser, 2006).

No entanto, são poucos os artigos que relacionam as características das empresas e seus retornos em tempos de crise: “o papel das características da ação, da empresa e da indústria na explicação do impacto de *crashes* no retorno individual da ação não recebeu atenção suficiente” (Wang, et al., 2009, p. 1574).

No intuito de colaborar com o preenchimento desta lacuna, este artigo busca investigar se existem características das empresas que auxiliam na discriminação entre as que perderam mais do que o mercado daquelas que perderam menos, em uma determinada situação de *market crash*. A data selecionada foi 15 de outubro de 2008, em que ocorreu a maior queda do índice S&P 500 (-9,0%) desde outubro de 1987. A amostra foi composta de empresas listadas na NYSE (*New York Stock Exchange*).

Wang *et al.* (2009) utilizam em seu trabalho a regressão linear múltipla. No presente estudo, é utilizada a regressão logística. Adicionalmente, incluímos a análise discriminante, a fim de corroborar os resultados. As duas técnicas, em geral, apresentam similaridades dos resultados, quando se utiliza uma variável dependente com dois grupos. A seção 2 apresenta uma revisão dos conceitos que nortearam o desenvolvimento deste estudo. A seção 3 contém os aspectos metodológicos e as seções 4 e 5 exibem os principais resultados obtidos e as conclusões, respectivamente.

## 2. REVISÃO CONCEITUAL

Esta segunda seção é composta de uma revisão conceitual da literatura a respeito de características das empresas que explicam a variabilidade do seu retorno, *market crashes* e crises.

## 2.1 Características das empresas, volatilidade e retorno

De acordo com a formulação do CAPM (*Capital Asset Pricing Model*), o beta explica completamente e de maneira linear o retorno das ações, portanto, bastaria o beta para prever o retorno esperado de um ativo, dado o retorno esperado do mercado.

No final da década de 70 e durante a década de 80, diversos autores visaram explicar o retorno das ações não apenas com o beta, mas se valendo de outras características das empresas, de forma isolada, ou combinadas com o beta ou, ainda, combinadas entre si: índice lucro / preço (Ball, 1978), índice valor patrimonial / valor de mercado (Rosenberg, Reid e Lanstein, 1985), tamanho (Banz, 1981), liquidez (Amihud e Mendelson, 1986) e alavancagem (Bhandari, 1988). Estes estudos indicaram empiricamente que quanto maior o índice lucro / preço, o índice valor patrimonial / valor de mercado ou a alavancagem maior o retorno esperado da ação e quanto maior o tamanho ou a liquidez menor o retorno esperado da ação. As explicações para estes efeitos identificados empiricamente são dadas de forma hipotética, não sendo desenvolvida nenhuma teoria a respeito, tal como o desenvolvimento teórico que culminou no CAPM.

Fama e French (1992), compilando os estudos anteriores – exceto Amihud e Mendelson (1986), propõem um modelo com duas variáveis explicativas: tamanho da empresa (*size*, medido pelo valor de mercado da empresa) e índice valor patrimonial / valor de mercado (ou *book value / market value*, BV/MV). Os autores encontram evidências, no período de 1963 a 1990, de que empresas menores e com alto índice BV/MV tendem consistentemente a maiores retornos se comparadas às empresas maiores e com baixo índice MV/BV. Observa-se que Fama e French (1992) concluem que, na presença de BV/MV e tamanho, o beta não ajuda a explicar o retorno das ações, assim como lucro/preço e alavancagem também são redundantes.

Kothari, Shanken e Sloan (1995, p. 198) analisam a relação do beta e do tamanho com o retorno das ações, encontrando resultados estatisticamente significantes maiores para o beta do que para o tamanho da empresa na maioria das análises. Os autores criticam duramente os achados de Fama e French (1992). Antes, Black (1993, p. 17) já havia contestado Fama e French indicando que o anúncio da morte do beta parecia prematuro. Diversos autores seguiram testando o beta em conjunto com outras variáveis na explicação do retorno das ações, com amostras em diversos países e períodos de tempo, alguns deles expostos na Tabela 1.

Tabela 1: Estudos empíricos visando explicar o retorno das ações com características das empresas

Autores	Variáveis
Chan, Jegadeesh e Lakonishok (1996)	Efeito momento, retorno futuro, <i>size</i> e BV/MV
Berk, Green e Naik (1999)	BV/MV, <i>size</i> e efeito momento
Amihud (2002)	Iliquidez
Pástor e Veronesi (2003)	MV/BV (variável dependente), idade da empresa, retorno futuro (com diferentes <i>lags</i> ), rentabilidade (ROE) em vários períodos, tamanho da empresa, alavancagem e volatilidade da rentabilidade (VOLP)
Fama e French (2006a)	BV/MV, rentabilidade e investimento
Fama e French (2006b)	BV/MV, lucro / preço e <i>size</i>
Jiang e Lee (2007)	BV/MV e <i>dividend yield</i>

Autores	Variáveis
Phalippou (2007)	BV/MV, <i>size</i> e propriedade
Li, et al. (2008)	Efeito momento, retorno futuro, <i>size</i> , BV/MV e risco da carteira (com diferentes <i>lags</i> )
Mussa, Santos e Famá (2007)	BM/MV, <i>size</i> e momento (usados em fatores, além do fator de mercado)
Fama e French (2008)	BV/MV, <i>size</i> e emissões líquidas
Mussa, et al.(2008)	BV/MV, tamanho e momento
Almeida e Eid (2010)	BV/MV, <i>size</i> e emissões líquidas
Chen e Zhang ( <i>in press</i> )	Nível de investimento e nível de ROA (usados em fatores, além do fator de mercado)
Flister, Bressan e Amaral (2011)	BM/MV, <i>size</i> e momento

**Nota.** Elaborada pelos autores.

Grande parte dos estudos que analisam o relacionamento entre as características das empresas e o retorno de suas ações é feita em períodos normais e não em períodos de crises.

## 2.2 Crises no mercado de ações

Levy (2008, pp. 138-139, 153) identifica os *market crashes* como um complexo, porém instantâneo fenômeno provocado pela associação de eventos econômicos externos a fatores psicológicos e sociais.

Poucos estudos analisam a relação entre características das empresas e retorno em períodos de crise ou em *market crashes*. Wang *et al.* (2009) estudaram oito dias de *market crash* relacionando retorno com diversas características das empresas: beta, *size*, liquidez, BV/MV, retornos defasados (*lagged returns*) entre outras. Michayluk e Neuhauser (2006) examinam o comportamento de ações em 27 de outubro de 1997, relacionando retorno com beta, *size*, preço e risco. Miyajima e Yafeh (2007) estudaram eventos associados à crise bancária no Japão de 1995 a 2000, relacionando desempenho das empresas com tamanho,  $q$  de Tobin, alavancagem, intensidade de pesquisa e desenvolvimento, *rating* das dívidas e medidas de intensidade de laços bancários.

A maioria dos estudos em crises foca outros temas, por exemplo: causas, contágio, volatilidade, co-integração e co-movimento, conforme se pode observar alguns exemplos na Tabela 2.

Tabela 2: Estudos empíricos em períodos de crises

Autores	Tema
Arshanapalli e Doukas (1993)	Cointegração
Cheung (2000)	Transmissão de volatilidade
Hon, Strauss e Yong (2004)	Correlação
Bunda, Hamann e Lall (2009)	Contágio
Catanach e Ragatz (2010)	Causas

**Nota.** Elaborada pelos autores.

### 3. METODOLOGIA

Esta seção descreve a metodologia adotada no presente trabalho a partir do objetivo deste artigo, que é investigar se existem características comuns das empresas que auxiliam na discriminação entre as empresas que perderam mais do que o mercado daquelas que perderam menos do que o mercado em uma determinada situação de *market crash*. A data selecionada foi 15 de outubro de 2008, o dia com a maior queda do índice S&P 500 no ano de 2008 (-9,0%). A amostra foi composta por empresas listadas na NYSE. Para tanto, utilizaram-se duas técnicas: regressão logística e análise discriminante.

#### 3.1 Base de dados

As informações relativas aos demonstrativos financeiros e aos valores de mercado (da firma e do patrimônio líquido) foram obtidas no sistema de informações Economática® em 01 de abril de 2010. As informações relativas às cotações, ajustadas pelos proventos, e o volume de negócios foram obtidas na mesma fonte em 05 de maio de 2010.

A base de dados era composta de 971 empresas, das quais 767 não financeiras. Deste total, foram selecionadas 461 empresas que apresentavam todas as informações relevantes para a análise proposta. Não existem empresas com mais de uma classe de ações na amostra, ou seja, a amostra é composta por 461 ações de empresas diferentes. A Tabela 3 indica as características das empresas utilizadas no presente estudo.

Tabela 3: Características utilizadas no estudo

Característica	Descrição	Data Base
<i>Size</i>	Logaritmo natural do patrimônio líquido da empresa	30 de junho de 2008
BV/MV	<i>Book Value / Market Value</i> , onde <i>market value</i> é o valor de mercado do patrimônio líquido da empresa e o <i>book value</i> é o valor contábil do patrimônio líquido da empresa	<i>Market value</i> de 14 de outubro de 2008 e patrimônio líquido de 30 de junho de 2008
E/P	<i>Earnings / Price</i> onde <i>earnings</i> é o lucro líquido contábil da empresa e <i>price</i> é o valor de mercado do patrimônio líquido da empresa (mesmo que <i>Market Value</i> )	<i>Price</i> de 14 de outubro de 2008 e lucro líquido de 30 de junho de 2008
ROE	Lucro Líquido / Patrimônio Líquido	30 de junho de 2008
ALAV	<i>Firm Value / Market Value</i> , onde <i>firm value</i> é o valor de mercado da firma	14 de outubro de 2008
BETA	Beta calculado versus o índice S&P 500, com histórico de retornos de 5 anos e intervalo de retorno mensal	30 de junho de 2008
DVPD	Desvio padrão dos retornos no 1 ano anterior ao <i>market crash</i> (janela [-252,-22])	15 de outubro de 2008
BETASW	Beta calculado versus o S&P, com histórico de retornos de 1 ano (janela [-252,-22]) e intervalo de retorno diário, ajustado segundo Scholes-Williams (1977)	15 de outubro de 2008
LR_	Retorno defasado, sendo LR_1s o retorno na semana anterior ao <i>market crash</i> (janela [-7,-2]), LR_1m o retorno no mês anterior ao <i>market crash</i> (janela [-24,-2]), LR_2m o retorno nos 2 meses anteriores ao <i>market crash</i> (janela [-	15 de outubro de 2008

Característica	Descrição	Data Base
	46,-2]) e LR_3m o retorno nos 3 meses anteriores ao <i>market crash</i> (janela [-68,-2])	
ILIQ	$\frac{\sum_{t=1}^T \frac{ r_t }{v_t}}{T} \times 1000$ Calculado como (Amihud, 2002), onde $ r_t $ é o módulo do retorno no dia $t$ e $v_t$ é o volume em \$ no dia $t$ , calculado para o período de tempo de 1 ano anterior ao <i>market crash</i> (janela [-252,-22])	15 de outubro de 2008

**Nota.** Elaborada pelos autores.

### 3.2 Partição da amostra

As 461 empresas da amostra foram particionadas em três grupos. O primeiro grupo foi formado com as 150 empresas com as menores perdas, o segundo grupo foi formado com as 150 empresas com as maiores perdas e o terceiro grupo com as demais 161 empresas com perdas intermediárias. Na regressão logística e na análise discriminante foram utilizados os grupos extremos (primeiro e segundo grupos), excluindo-se o grupo do meio da análise (terceiro grupo), segundo a abordagem de extremos polares.

Posteriormente, as 150 empresas do primeiro grupo (menores perdas) foram aleatoriamente divididas em 2 subgrupos com 75 empresas cada: o grupo de análise ( $GA^-$ ) e o grupo de teste ( $GT^-$ ). Da mesma forma, as 150 empresas do segundo grupo (maiores perdas) foram aleatoriamente divididas em 2 subgrupos com 75 empresas cada: análise ( $GA^+$ ) e teste ( $GT^+$ ). Assim, a amostra de análise é composta por  $GA^-$  e  $GA^+$  e conta com 150 empresas e a amostra de teste é composta por  $GT^-$  e  $GT^+$  e conta com 150 empresas. Este procedimento de divisão da amostra em grupo de análise e grupo de teste é recomendado por Hair et al. (2009, p. 236) como uma maneira preferida de validação do modelo logístico ou da função discriminante, denominado de abordagem de partição da amostra ou validação cruzada. Segundo os autores, o procedimento mais comum é a divisão 50-50 da amostra total, metade para o grupo de análise e metade para o grupo de teste.

A abordagem de extremos polares e a abordagem de partição da amostra reduzem a amostra total de 461 empresas para uma amostra de análise de 150. Isto significa 11,5 observações por variável, inferior ao número ideal (20 observações), mas superior ao número mínimo (5 observações) recomendado pela literatura (ver Tabela 4). Esta abordagem, embora prejudicial a relação número de observações por variável, revela vantagens metodológicas para a validação e a consistência dos modelos.

### 3.3 Técnicas: regressão logística e análise discriminante

Foram utilizadas duas técnicas multivariadas: regressão logística e análise discriminante. Na regressão logística, a variável dependente são os grupos (no caso: grupo com maior perda e grupo com menor perda) e as variáveis explicativas são métricas (no caso: características das empresas) ou não métricas (*dummies*). A equação da regressão logística resultará na probabilidade da observação estar em um dos grupos, ou seja, a

probabilidade de um determinado evento ocorrer (equação 1, na qual Z é a combinação linear das variáveis explicativas, obtido por meio do método da verossimilhança).

$$\text{prob} = \frac{e^Z}{1 + e^Z} \quad \text{Equação 1}$$

Na análise discriminante, a variável dependente também são os grupos e as variáveis explicativas são métricas. As variáveis explicativas são combinadas linearmente (formando uma variável estatística) para melhor discriminar entre os grupos (variável dependente). A Tabela 4 detalha as duas técnicas utilizadas neste artigo.

Tabela 4: **Resumo das técnicas: regressão logística e análise discriminante.**

Descrição	Regressão Logística	Análise Discriminante
Objetivo da técnica	Encontrar um modelo de previsão de uma variável dependente binária a partir de variáveis métricas ou dicotômicas.	Identificar variáveis que melhor diferenciam dois ou mais grupos de observações mutuamente exclusivos.
Variáveis	Dependente: não métrica. Explicativas: métricas ou não métricas ( <i>dummies</i> ).	Dependente: não métrica. Explicativas: métricas.
Premissas	A regressão logística é menos afetada do que a análise discriminante quando suposições básicas como a normalidade de variáveis não são satisfeitas.	(i) normalidade multivariada das variáveis explicativas e (ii) variâncias e covariâncias iguais nos grupos.
Tamanho de amostra	Mínimo recomendado: cinco observações por variável explicativa. Sugestão: vinte observações por variável explicativa <sup>a</sup> .	
Qualidade do modelo	Uso da matriz de classificação, cuja finalidade é identificar quantas observações o modelo logístico ou a função discriminante são capazes de classificar corretamente.	
Precisão da classificação do modelo	Para verificar se a precisão da classificação <i>a posteriori</i> é significativamente superior ao acaso, pode-se utilizar os critérios (i) da probabilidade máxima, (ii) da probabilidade proporcional, (iii) do Press's Q e (iv) da regra prática de 25% excedente.	

**Nota.** Elaborada pelos autores.

<sup>a</sup> Segundo Hair *et al.* (2009, pp. 235 e 236).

### 3.4 Precisão da classificação dos modelos

O modelo gerado pela regressão logística ou pela análise discriminante será útil caso ele discrimine entre os grupos significativamente mais do que se poderia conseguir ao acaso. Este artigo trabalha com dois grupos – maiores perdas e menores perdas. Para dois grupos de igual tamanho, o acaso seria capaz de acertar a classificação de 50% das observações. Para dois grupos de tamanhos diferentes, o acaso seria a frequência do maior grupo (critério da probabilidade máxima), critério que gera o mais alto padrão de comparação. A Tabela 5 sintetiza os critérios de precisão de classificação e os resultados esperados.

Tabela 5: Critérios de precisão de classificação e resultados esperados

Critério	Descrição	Resultado Esperado
Critério da probabilidade proporcional	Utilizado para grupos de tamanhos diferentes. Considerando 2 grupos, é dado pela fórmula a seguir, na qual $p$ é a proporção de elementos do grupo 1 e $(1-p)$ é a proporção de elementos do grupo 2: $C_{pro} = p^2 + (1 - p)^2$	A precisão do modelo deve ser superior à probabilidade proporcional calculada.
Critério do acaso pela probabilidade máxima	Quando se tratar de grupos de tamanhos iguais é calculado como $1/n$ , em que $n$ é o número de grupos e quando se tratar de grupos de tamanhos diferentes é igual à probabilidade do grupo maior <sup>a</sup> .	A precisão do modelo deve ser superior à probabilidade máxima calculada.
Critério Press's Q	É uma estatística que avalia o poder discriminatório da matriz de classificação comparando com o modelo ao acaso. Sua hipótese $H_0$ é que a matriz de classificação é igual à classificação ao acaso. O Press's Q é dado pela fórmula a seguir, na qual $N$ é o tamanho da amostra, $n$ é o número de classificações corretas, $k$ é o número de grupos e $gl$ é número de graus de liberdade. $PQ = \frac{[N - (n \cdot k)]^2}{N \times (k - gl)}$	Deve-se rejeitar a hipótese $H_0$ para a obtenção de um modelo válido, ou seja, o valor obtido deve superar o valor crítico dado por esta estatística, quando a matriz de classificação poderá ser estatisticamente melhor do que o acaso.
Critério da regra prática de 25% excedente	Como regra prática, Hair <i>et al.</i> (2009, p. 247) indicam que “a verdadeira precisão preditiva (razão de sucesso) deve exceder qualquer valor de critério em pelo menos 25%”.	A precisão do modelo deve ser superior a 1,25 x a probabilidade máxima.

**Nota.** Elaborada pelos autores.

<sup>a</sup> Segundo Hair *et al.* (2009, p. 245).

## 4. RESULTADOS

A seguir são apresentados os resultados obtidos. A análise dos dados e a apresentação dos resultados estão divididas em quatro subseções: Análise Exploratória, Regressão Logística, Análise Discriminante e Modelo do Beta.

### 4.1 Análise exploratória

Considerando a amostra de análise, as seguintes variáveis apresentam distribuição normal (univariada), testada por meio do teste de Kolmogorov-Smirnov e nível de significância de 0,05: *Size*, *ALAV*, *BETA*, *LR\_1s*, *LR\_1m*, *LR\_2m*, *LR\_3m*, *DVPD* e *BETASW*. Rejeita-se a hipótese da normalidade das distribuições das variáveis: *BV/MV*, *E/P*, *ROE* e *ILIQ*. Mesmo após a aplicação da transformação logarítmica (LN), observa-se que, do conjunto das variáveis que rejeitou a hipótese da normalidade, apenas a variável transformada *BV/MV* apresentou distribuição normal.

A análise discriminante tem como uma de suas premissas a normalidade multivariada, teste que não foi possível ser realizado por meio do SPSS, ferramenta estatística utilizada neste artigo. No entanto, “a suposição mais importante é a igualdade das matrizes de

covariância, o que afeta tanto estimação quanto classificação” (Hair *et al.*, 2009, p. 237), que foi testada por meio da estatística Box’s M.

Como o p-valor do teste Box’s M é 0,802 (não rejeitando a hipótese da igualdade das variâncias e covariâncias), conclui-se pela pertinência de conduzir a análise discriminante. De acordo com Hair *et al.* (2009, p. 321), o teste Box’s M também é sensível a desvios de normalidade. A regressão logística, por sua vez, é menos afetada quando as suposições básicas não são satisfeitas.

Os resultados das análises com as variáveis não transformadas e com as variáveis transformadas (LN) foram similares.

O teste de igualdade de médias entre os grupos a 5% de significância estatística indicou que as seguintes variáveis têm médias diferentes nos grupos de menores perdas e de maiores perdas: BETA, LR\_1s, LR\_1m, LR\_2m, LR\_3m, DVDP e BETASW. Portanto, estas variáveis são boas candidatas a discriminarem os grupos, enquanto as demais aparentemente não o são. Em se tratando de uma análise multivariada optou-se por disponibilizar todas as variáveis para a análise *stepwise*.

#### 4.2 Regressão logística

A equação do logit (Z) resultante do modelo pode ser escrita como:

$$Z = - 7,136 - 9,928 * LR\_2m + 142,329 * DVDP$$

O modelo foi montado considerando a probabilidade de estar no grupo das maiores perdas, sendo assim, quanto maior o logit (Z), maior a chance de a empresa estar no grupo de maiores perdas. Desta forma, quanto menor o retorno dos últimos 2 meses antes do *crash* (LR\_2m) e quanto maior o desvio padrão dos retornos no último ano antes do *crash* (DVDP), maior a chance de estar no grupo de maiores perdas. Entre as duas variáveis, segundo a estatística de Wald, o retorno dos últimos 2 meses é a que mais contribui para o modelo (42,358, 24,466 e 19,426 são as estatísticas Wald para o coeficiente linear, o coeficiente angular de LR\_2m e o coeficiente angular de DVDP, respectivamente).

O R<sup>2</sup> de Cox e Snell é 0,498, o R<sup>2</sup> de Nagelkerke é 0,664 e o pseudo R<sup>2</sup> é 0,497. O p-valor da estatística de Hosmer e Lemeshow é 0,288, não rejeitando a hipótese de que os valores reais e esperados sejam iguais, indicando que o modelo logístico tem um ajuste aceitável. O modelo foi capaz de classificar, *a posteriori*, a amostra de análise com 87,3% de acerto (painel a da Tabela 6). A classificação da amostra de teste atingiu 84,0% de acerto (painel b da Tabela 6). Como ambas as amostras são compostas de grupos de tamanhos iguais, observa-se que o modelo é bem superior a uma classificação aleatória (50% de chance de acerto) mesmo considerando a regra prática de 25% excedente (1,25 x 50% = 62,5%).

Tabela 6: Classificação *a posteriori* – amostra de análise (150 empresas) e amostra de teste (150 empresas).

Painel (a): Classificação da amostra de análise (GA<sup>-</sup> e GA<sup>+</sup>): 150 empresas

		Previsto		<b>Total</b>
		Menores Perdas	Maiores Perdas	
Real	Menores Perdas	65 (86,7%)	10 (13,3%)	<b>75 (100%)</b>
	Maiores Perdas	9 (12,0%)	66 (88,0%)	<b>75 (100%)</b>
	<b>Total</b>	<b>74</b>	<b>76</b>	
Acerto		131 (87,3%)		

Painel (b): Classificação da amostra de teste (GT<sup>-</sup> e GT<sup>+</sup>): 150 empresas

		Previsto		<b>Total</b>
		Menores Perdas	Maiores Perdas	
Real	Menores Perdas	60 (80,0%)	15 (20,0%)	<b>75 (100%)</b>
	Maiores Perdas	9 (12,0%)	66 (88,0%)	<b>75 (100%)</b>
	<b>Total</b>	<b>69</b>	<b>81</b>	
Acerto		126 (84,0%)		

**Nota.** Elaborada pelos autores baseado nas saídas do software SPSS e também em análises próprias.

A Tabela 6 indica, além do alto índice de acerto da classificação, uma simetria entre os acertos para as menores perdas e para as maiores perdas. Na amostra de análise o modelo classificou corretamente 86,7% das empresas com as menores perdas e 88,0% das empresas com as maiores perdas (80,0% e 88,0%, respectivamente, para a amostra de teste).

Além da verificação da classificação na amostra de análise e na amostra teste, também se verificou a classificação do modelo na amostra inteira (considerando as 461 empresas), visando verificar se o modelo é capaz de discriminar entre as empresas que perderam mais que o S&P 500 (indicado como Perda > M) e aquelas que perderam menos (indicado como Perda < M).

Tabela 7: Classificação *a posteriori* – amostra toda (461 empresas)

		Previsto		<b>Total</b>
		Perda < M	Perda > M	
Real	Perda < M	159 (79,9%)	40 (20,1%)	<b>199 (100%)</b>
	Perda > M	91 (34,7%)	171 (65,3%)	<b>262 (100%)</b>
	<b>Total</b>	<b>250</b>	<b>211</b>	
Acerto		330 (71,6%)		

**Nota.** Elaborada pelos autores baseado nas saídas do software SPSS e também em análises próprias.

A Tabela 7 indica que no conjunto, o modelo teve um alto nível de acerto (71,6%). Considerando que o maior grupo tem 262 empresas em 461 (56,8%).

Verifica-se também que o modelo classificou melhor as empresas que tiveram perdas menores do que o S&P 500 (79,9% de acerto) comparativamente a classificação das empresas que tiveram perdas maiores do que o S&P 500 (65,3%). O modelo foi feito para discriminar entre as maiores e menores perdas para uma amostra que tem mediana de retorno negativo de 9,8%. O S&P 500, em 15 de outubro de 2008, perdeu 9,0%. Este fato justifica o modelo de ter sido mais eficiente para discriminar entre as que perderam menos do que o S&P 500.

Observa-se, na tabela 8, que a precisão da classificação pelos critérios da probabilidade máxima, probabilidade proporcional e Press's Q indicam que o modelo é superior a uma classificação ao acaso mesmo considerando a regra prática de 25% excedente ( $1,25 \times 56,8\% = 71,0\%$ ). Sendo assim, a utilização do modelo logístico é pertinente.

Tabela 8: Precisão da classificação – amostra toda (461 empresas)

Descrição	Valor	Decisão
Acerto do modelo	71,6%	-
Critério da Probabilidade Máxima	56,8%	Precisão do modelo é superior
Regra prática de 25% excedente	71,0%	Precisão do modelo é superior
Critério da Probabilidade Proporcional	50,9%	Precisão do modelo é superior
Press's Q <sup>a</sup>	85,9	H <sub>0</sub> rejeitado (precisão do modelo é superior ao acaso)

**Nota.** Elaborada pelos autores.

<sup>a</sup> Valor crítico de 3,84. Parâmetros utilizados: N=461 (tamanho da amostra), n=332 (número de classificações corretas), k=2 (número de grupos).

#### 4.3 Análise discriminante

A função discriminante não padronizada resultante do modelo pode ser escrita como:

$$\text{Escore discriminante} = - 3,935 - 5,377 * LR\_2m + 78,562 * DVDP$$

O escore discriminante de corte é zero, por se tratar de dois grupos de igual tamanho (75 observações em cada grupo). Observando-se os centróides das equações, verifica-se que as empresas com maiores perdas têm centróide positivo (1,053) e têm escore discriminante também positivo. Desta forma, as empresas com menores retornos nos últimos 2 meses antes do *crash* e com maiores desvio padrão dos retornos no último ano antes do *crash* apresentaram as maiores perdas. As duas variáveis obtidas no modelo, seus sinais e o ranking (segundo a estatística Wilk's Lambda) são iguais aos obtidos no modelo logístico (0,587 e 0,637 são as estatísticas Wilk's Lambda para LR\_2m e DVDP, respectivamente). A significância do Wilk's Lambda do modelo final (0,000) indica que as duas variáveis coletivamente discriminam bem os grupos.

A precisão de classificação *a posteriori* da amostra de análise (87,3%), da amostra de teste (84,7%) e da amostra total (72,0%) se assemelham à precisão do modelo logístico e, portanto, suas tabelas não são apresentadas.

A classificação *a posteriori* do modelo também é bastante satisfatória quando submetida aos critérios da probabilidade máxima (incluindo o critério de pelo menos 25% a mais do que o obtido ao acaso), probabilidade proporcional e Press's Q (89,4, superior ao valor

crítico de 3,84). Portanto, o modelo discriminante apresenta uma precisão de classificação superior ao acaso, conferindo-lhe pertinência na sua utilização.

#### 4.4 Modelo do beta

A teoria de finanças que norteia a maioria dos acadêmicos e profissionais indica que empresas que apresentam beta menor do que 1 têm menos risco do que o mercado e deveriam ter, portanto, perdas menores do que a perda do mercado. O inverso também seria verdadeiro, indicando que as empresas com betas maiores do que 1 têm mais risco do que o mercado e deveriam ter, portanto, perdas maiores do que a perda do mercado. Como o beta não compõe o conjunto de variáveis do modelo logístico ou do modelo discriminante, buscou-se verificar qual seria a precisão da classificação utilizando-se a teoria convencional.

A Tabela 10 indica a classificação das empresas utilizando-se um modelo de discriminação baseado apenas no beta: se maior do que 1 classifica a ação em perda esperada maior do que a perda do mercado e quando menor do que 1 classifica a ação em perda esperada menor do que a perda do mercado. Verifica-se que o beta (BETA) acerta 61,4% das classificações, desempenho inferior ao que os modelos apurados por meio da regressão logística e da análise discriminante produziram (71,6% e 72,0%, respectivamente), utilizando as duas variáveis: (i) retorno defasado de 2 meses (LR\_2m) e (ii) desvio padrão de 1 ano (DVPD).

Tabela 10: Classificação *a posteriori* – amostra toda (461 empresas) – Beta

		Previsto		Total
		Perda < M	Perda > M	
Real	Perda < M	117 (58,8%)	82 (41,2%)	<b>199 (100%)</b>
	Perda > M	96 (36,6%)	166 (63,4%)	<b>262 (100%)</b>
Total		<b>213</b>	<b>248</b>	
Acerto		283 (61,4%)		

**Nota.** Elaborada pelos autores baseado nas saídas do software SPSS e também em análises próprias.

Como a variável DVPD é calculada a partir de um ano de retornos diários e, tendo em vista que a variável BETA foi calculada com cinco anos de retornos mensais, foi realizado outro cálculo nas variáveis candidatas a discriminar os grupos: o beta calculado com um ano de retornos diários. Sabendo-se que retornos diários podem enviesar para baixo os betas de ações com baixa liquidez, aplicou-se o ajuste segundo Scholes e Williams (1977). O beta calculado conforme o ajuste de Scholes e Williams (BETASW) também não compõe o conjunto de variáveis do modelo logístico ou do modelo discriminante. Utilizado sozinho para discriminação dos grupos, assim como feito para o BETA, o BETASW acerta 62,7% das classificações.

A precisão de acerto do modelo do beta é superior aos critérios de probabilidade máxima (56,8%), critério de probabilidade proporcional (50,9%) e rejeita a  $H_0$  do critério Press's Q (23,9, acima do valor crítico de 3,84), mostrando-se satisfatório por estes critérios. No entanto, não é satisfatório quando se utiliza a regra prática de pelo menos 25% a mais do

que o obtido ao acaso ( $1,25 \times 56,8\% = 71,0\%$ ). Ressalta-se que esta regra prática deve ser vista com cautela.

## 5. CONCLUSÕES

O artigo propõe-se a verificar se existem características comuns das empresas que auxiliam na discriminação das ações em dois grupos: aquelas que perderam mais do que o mercado e aquelas que perderam menos do que o mercado, em uma determinada situação de *market crash*. Neste sentido, o artigo contribui, por meio da aplicação de duas técnicas multivariadas de objetivos similares, para a compreensão das características das empresas que ajudam a explicar a variabilidade dos seus retornos em um momento de crise.

No *market crash* analisado em 15 de outubro de 2008, observou-se que, para a amostra de 461 empresas do NYSE, o retorno de dois meses antes do *crash* e o desvio padrão de retornos diários de um ano anteriormente ao evento são relevantes para a discriminação entre os dois grupos.

Este resultado foi obtido aplicando-se duas técnicas multivariadas: regressão logística e análise discriminante. Os modelos resultantes de ambas as técnicas classificaria as empresas corretamente, significativamente acima do acaso, conferindo-lhes pertinência nas suas utilizações. A regressão logística acertaria 71,6% dos casos e a análise discriminante acertaria 72,0%, resultados que são superiores aos critérios da probabilidade máxima, probabilidade proporcional e Press's Q.

As empresas com menor desvio padrão de retornos no último ano antes do *crash* teriam menos chance de integrarem o grupo com perdas maiores do que o S&P 500, resultado consistente com a teoria, que diz que maiores riscos apresentam maiores oscilações. As empresas com maiores retornos nos dois meses anteriores ao *crash* teriam, igualmente, menos chance de integrarem o grupo com perdas maiores.

Observa-se que o beta não integra o conjunto de variáveis que compõe os modelos logístico e discriminante. Pelo CAPM o beta deveria ser, sozinho, capaz de explicar o retorno das ações. A sua ausência no conjunto de variáveis dos modelos indica que, para a data analisada, ele é redundante na presença de retorno defasado de 2 meses (LR\_2m) e (ii) desvio padrão de 1 ano (DVDP).

Além disto, utilizando o beta como única variável para separar as ações entre os grupos (beta maior do que 1 classifica a ação no grupo de perdas maiores do que o mercado e beta menor do que 1 classifica a ação no grupo de perdas menores do que o mercado), verifica-se que os resultados de acerto são inferiores aos obtidos pelos modelos logístico e discriminante. Aplicando-se um beta de 5 anos e retornos mensais, o acerto foi de 61,4% dos casos. Aplicando-se um beta de 1 ano e retornos diários (mesmos parâmetros da variável desvio padrão), ajustado pela metodologia de Scholes e Williams (1977), o acerto foi de 62,7% dos casos. Em ambos os casos, a classificação alcançada pelo beta é inferior a regra prática dos 25% excedente a probabilidade máxima.

As limitações do estudo são: (i) estuda-se apenas o comportamento em um dia, (ii) este dia está inserido em um período de grande turbulência do mercado e (iii) trata-se de um único evento estudado (uma única data). Futuros estudos poderiam estender a análise para um período de tempo maior do que um dia, além de estudar outros momentos de *market crash*.

## REFERÊNCIAS

ALMEIDA, J.R., EID, W. Estimando o retorno das ações com decomposição do índice book-to-market: evidências na bovespa. **Revista Brasileira de Finanças**, v.8, n.4, 2010, pp. 417-441.

AMIHUD, Y. Illiquidity and Stock Returns: Cross-Section and Time-Series Effects. **Journal of Financial Markets**, v.5, n.1, 2002, pp. 31-56.

AMIHUD, Y., MENDELSON, H. Liquidity and stock returns. **Financial Analysts Journal**, v.42, n.3, 1986, pp. 43-48.

ARSHANAPALLI, B., DOUKAS, J. International stock market linkages: Evidence from the pre- and post-October 1987 periods. **Journal of Banking and Finance**, v.17, 1993, pp. 193-208.

BALL, R. Anomalies in relationships between securities' yield and yield-surrogates. **Journal of Financial Economics**, v.6, 1978, pp. 103-126.

BANZ, R. W. The relationship between return and market value of common stocks. **Journal of Financial Economics**, v.9, 1981, pp. 3-18.

BERK, J.B., GREEN, R.C., NAIK, V. Optimal Investment, Growth Options, and Security Returns. **Journal of Finance**, v. 54, n.5, 1999, pp. 1553-1607.

BHANDARI, L.C. Debt/equity ratio and expected common stock returns: empirical evidence. **Journal of Finance**, v.43, 1988, pp. 507-528.

BLACK, F. Beta and return. **Journal of Portfolio Management**, v.20, 1993, pp. 8-18.

BUNDA, I., HAMANN, J., LALL, S. Correlations in emerging market bonds: the role of local and global factors. **Emerging Markets Review**, v. 10, 2009, pp. 67-96.

CATANACH, A.H., RAGATZ, J.A. Market Crisis: Black Swan, Perfect Storm or Tipping Point? **Bank Accounting & Finance**, v.23, n.3, April-May 2010, pp. 20-26.

CHAN, L.C., JEGADEESH, N., LAKONISHOK, J. Momentum Strategies. **The Journal of Finance**, v.51, n.5, 1996, pp. 1681-1713.

CHEN, L., & ZHANG, L. A better three-factor model that explain more anomalies. **The Journal of Finance**, no Prelo, 2010.

CHEUNG, D.W. The Impulse of Stock Market Volatility and the Market Crash of October 1987. **Journal of Business Finance & Accounting**, v.27, June-July, 2000, pp. 761-776.

- FAMA, E.F., FRENCH, K.R. The Cross-Section of Expected Stock Returns. *The Journal of Finance*, v.47, n.2, 1992, pp. 427-465.
- FAMA, E.F., FRENCH, K.R. Profitability, investment and average returns. *Journal of Financial Economics*, v.82, 2006a, pp. 491-518.
- FAMA, E.F., FRENCH, K.R. The value premium and the CAPM. *The Journal of Finance*, v.61, n. 5, 2006b, pp. 2163-2185.
- FAMA, E.F., FRENCH, K.R. Average returns, b/m, and share issues. *The Journal of Finance*, v.63, n.6, 2008, pp. 2971-2995.
- FLISTER, F.V., BRESSAN, A.A., AMARAL, H.F. CAPM condicional no mercado brasileiro: um estudo dos efeitos momento, tamanho e book-to-market entre 1995 e 2008. *Revista Brasileira de Finanças*, v.9, n.1, 2011, pp. 105-129.
- HAIR, J.F. et al.. **Análise multivariada de dados** (6ª. ed.). Porto Alegre: Bookman, 2009.
- HON, M.T., STRAUSS, J., YONG, S. Contagion in Financial Markets after September 11: myth or reality? *The Journal of Financial Research*, v.27, n.1, 2004, pp. 95-114.
- JIANG, X., LEE, B. Stock Returns, dividend yield and book-to-market ratio. *Journal of Banking & Finance*, v. 31, 2007, pp. 455-475.
- KOTHARI, S.P., SHANKEN, J., SLOAN R.G. (1995). Another Look at the Cross-section of Expected Stock Returns. *The Journal of Finance*, v.50, n.1, 1995, pp. 185-224.
- LEVY, M. Stock market crashes as social phase transitions. *Journal of Economic Dynamics & Control*, v.32, 2008, pp. 137-155.
- LI, X., et al. Momentum profits and time-varying unsystematic risk. *Journal of Banking & Finance*, v.32, 2008, pp. 541-558.
- MICHAYLUK, D., NEUHAUSER, K.L. Investor overreaction during market declines: evidence from the 1997 Asian Financial crisis. *The Journal of Financial Research*, v.29, n.2, 2006, pp. 217-234.
- MIYAJIMA, H., YAFEH, Y. Japan's banking crisis: an event study perspective. *Journal of Banking & Finance*, v.31, n.9, 2007, pp. 2866-2885.
- MUSSA, A., SANTOS, J.O., FAMÁ, R. A adição do fator de risco momento ao modelo de precificação de ativos dos três fatores de Fama & French aplicado ao mercado acionário brasileiro. *Anais do VII Congresso USP de Controladoria e Contabilidade*, São Paulo, SP, Brasil, 2007.

MUSSA, A. et al.. A influência das condições do mercado acionário e da política monetária no comportamento dos indicadores de risco tamanho, índice book-to-market e momento, no mercado acionário brasileiro. *Anais do VII Encontro da Sociedade Brasileira de Finanças*, Rio de Janeiro, RJ, Brasil, 2008.

PÁSTOR, L., VERONESI, P. (2003). Stock valuation and learning about profitability. *The Journal of Finance*, v.58, n.5, 2003, pp. 1749–1789.

PHALIPPOU, L. Can risk-based theories explain the value premium? *Review of Finance*, v.11, 2007, pp. 143-166.

ROSENBERG, B., REID, K., LANSTEIN, R. Persuasive evidence of market inefficiency. *Journal of Portfolio Management*, v.9, 1985, pp. 18-28.

SCHOLES, M., WILLIAMS, J. Estimating betas from nonsynchronous data. *Journal of Financial Economics*, v.5, n.3, 1977, pp. 309-327.

WANG, et al.. Stock market crashes, firm characteristics, and stock return. *Journal of Banking & Finance*, v.33, 2009, pp. 1563-1574.