

3

Metodologia

A amostra considerada nesta pesquisa abrange 58 ações negociadas na BOVESPA durante o período de julho de 2005 a dezembro de 2010. Esta amostra representa todas as ações negociadas neste período, obedecendo ao critério adotado em Fu (2009), onde se exige que cada ação tenha um mínimo de 15 dias de negociação em cada mês da amostra. Por critério de conveniência, foram consideradas apenas as ações que estavam presentes em todos os meses do período amostral. Além disso, seguindo outros estudos que usaram o modelo de Fama e French (FAMA e FRENCH, 1993), como Malaga e Securato (2004) e Rogers e Securato (2009), foram excluídas da amostra as ações de empresas financeiras, por elas apresentarem normalmente um alto nível de endividamento, normal para este setor, que afeta o índice *book-to-market*. Adicionalmente, seguindo a metodologia desses estudos, foram também excluídas as ações de empresas que apresentavam Patrimônio Líquido negativo em 31 de dezembro em pelo menos um dos anos de 2004 a 2009.

3.1.

Estimação da Volatilidade Idiossincrática Através do Modelo de Fama & French (VI)

Para a primeira parte deste estudo, foram calculadas as volatilidades idiossincráticas de cada ação em cada mês, utilizando-se o desvio-padrão dos resíduos das regressões mensais de cada ação nos três fatores de Fama e French, como foi feito em Fu (2009) para o mercado acionário americano. Segundo o modelo de Fama e French (1993), três fatores explicam os retornos dos ativos: o excesso de retorno do mercado (o retorno da carteira de mercado menos o retorno do ativo livre de risco), o retorno de uma carteira SMB (*small minus big*, ou seja, a diferença entre o retorno de um portfólio com ações de empresas pequenas e um com ações de

empresas grandes), e o retorno de uma carteira HML (*high minus low*, ou seja, a diferença entre o retorno de um portfólio com ações de empresas com alto índice *book-to-market* e um com ações de empresas que apresentam baixo índice *book-to-market*). Este modelo é uma extensão do modelo CAPM, e baseia-se nos estudos de finanças comportamentais, segundo os quais o desempenho de ações de empresas pequenas tende a superar o de ações de empresas grandes, ao mesmo tempo em que o retorno de ações de empresas que apresentam um alto índice *book-to-market* tende a ser maior que o de ações de empresas que apresentam um baixo índice *book-to-market*.

Para a aplicação deste modelo, portanto, seguiu-se a metodologia de Fu (2009), onde utilizou-se os dados diários destas ações em todo o período da amostra. O modelo utilizado é o Modelo de Três Fatores de Fama e French, correspondente à Equação (3.1) abaixo:

$$R_{it} - r_t = \alpha_{it} + b_{it}(R_{mt} - r_t) + s_{it}SMB_t + h_{it}HML_t + \varepsilon_{it} \quad (3.1)$$

onde τ indica o dia e t indica o mês, $\tau \in t$, R_{it} representa o retorno de cada ação em cada dia, r_t representa a taxa livre de risco diária, R_{mt} é o retorno diário da carteira de mercado, e SMB_t e HML_t representam os retornos diários das carteiras SMB e HML. b_{it} , s_{it} e h_{it} são os coeficientes relativos a cada fator.

A taxa de retorno livre de risco utilizada foi o CDI. Todas as variáveis de retorno foram calculadas de forma contínua, isto é, usou-se o logaritmo natural da razão preço no dia τ / preço no dia $\tau - 1$. Além disso, para todas as ações da amostra foram utilizados os preços ajustados por proventos. Todos os dados foram obtidos no sistema Economática®, com exceção do CDI, que foi obtido no site do Banco Central do Brasil (www.bcb.gov.br). Para os valores de Patrimônio Líquido, os dados foram obtidos nos demonstrativos contábeis consolidados das empresas, no entanto, quando este não estava disponível, utilizou-se os demonstrativos contábeis não-consolidados.

A carteira de mercado utilizada foi a obtida pela ponderação de cada uma das 58 ações amostrais segundo o seu valor de mercado, como em Malaga e Securato (2004). O excesso de retorno de mercado (primeiro fator do modelo) foi calculado diariamente utilizando-se a diferença entre o retorno da carteira de mercado e o CDI. Para a obtenção dos prêmios pelo fator de risco SMB e HML, seguiu-se a metodologia de Fama e French (1993), que foi replicada em Malaga e Securato (2004):

- As ações amostrais foram ordenadas, em junho de cada ano do período, segundo o seu valor de mercado. Estas ações foram separadas em dois grupos, divididos pela mediana do valor de mercado, classificados como B (*big*) e S (*small*).

- Em cada um desses grupos, nesse mesmo mês, as ações foram ordenadas segundo o seu índice *book-to-market*. Este índice foi calculado através da razão entre o patrimônio líquido da empresa e o valor de mercado da ação em 31 de dezembro do ano anterior. Cada grupo foi então dividido entre três outros grupos, de acordo com o 30º e o 70º percentil, formando os grupos H (*high*), M (*medium*) e L (*low*).

- Desta forma, foram compostas seis carteiras em junho de cada ano, são elas:

- BH (*big and high*)
- BM (*big and medium*)
- BL (*big and low*)
- SH (*small and high*)
- SM (*small and medium*)
- SL (*small and low*)

- O retorno de cada uma destas seis carteiras foi calculado diariamente, através da ponderação das ações das carteiras segundo o seu valor de mercado.

- Para obtenção da variável SMB, calculou-se diariamente a diferença entre a média dos retornos das três carteiras *small* e a média dos retornos das três carteiras *big*. Já

para o cálculo da variável HML, usou-se a diferença entre a média dos retornos das duas carteiras *high* e a média dos retornos das duas carteiras *low*.

Por fim, como variável dependente, foi considerado o excesso de retorno diário de cada ação amostral, isto é, o seu retorno menos a taxa livre de risco.

Foram feitas regressões em séries temporais para cada ação em cada mês da amostra (66 meses), conforme a Equação (3.1), obtendo-se o desvio-padrão dos resíduos em cada uma delas. A volatilidade idiossincrática mensal foi encontrada multiplicando-se o desvio-padrão dos resíduos pela raiz quadrada do número de dias de negociação de cada ação em cada mês. Vale destacar que apenas nesta parte da pesquisa foram utilizados dados com frequência diária para o cálculo da volatilidade mensal.

3.2.

Propriedades das Séries Temporais de Volatilidade Idiossincrática (VI)

A Tabela 3.1 abaixo mostra as estatísticas das volatilidades idiossincráticas individuais (média, desvio-padrão, assimetria, curtose e auto-correlações). Primeiro, foram computadas as estatísticas da série temporal de volatilidade idiossincrática de cada ação amostral, e, em seguida, foi calculada a média para todas as 58 ações. Foram consideradas duas variáveis: a volatilidade idiossincrática (VI), e sua variação em tempo contínuo ($\ln(VI_t/VI_{t-1})$).

A média da volatilidade idiossincrática entre as ações foi de 7,54%, com desvio-padrão médio de 2,92%.

Tabela 3.1

Estatísticas das séries de volatilidade idiossincrática individuais (% mensal)

Estatísticas das séries de volatilidade idiosincrática individuais (% mensal)												
	N	Média	D. P.	Assim.	Curtose							
VI	58	7,54	2,92	1,49	6,96							
ln (VI _t /VI _{t-1})	58	-0,42	37,83	0,10	3,08							
	Auto-correlação - lags											
	1	2	3	4	5	6	7	8	11	12	13	
VI	0,40	0,29	0,22	0,17	0,18	0,11	0,07	0,05	-0,03	-0,05	-0,08	
ln (VI _t /VI _{t-1})	-0,63	0,13	0,02	-0,05	0,07	-0,02	-0,02	0,00	0,00	0,03	-0,07	

A parte de baixo da Tabela 3.1 mostra as auto-correlações das variáveis. O objetivo é verificar se a série temporal de volatilidade idiossincrática pode ser considerada um passeio aleatório, como foi constatado em Ang et al. (2006). Caso isto seja verdade, significa que é válido utilizar o valor da volatilidade idiossincrática em um dado mês para estimar o seu valor no mês seguinte. Nesse caso, a auto-correlação da variável em nível deve ser igual a um na primeira defasagem. Fora isso, ela deve ser igual a zero em todos os *lags* da primeira diferença ($\ln(VI_t/VI_{t-1})$).

Para a volatilidade em nível (VI), a auto-correlação na primeira defasagem é de 0,40, e vai diminuindo na medida em que são considerados mais *lags*. Já para a volatilidade na primeira diferença ($\ln(VI_t/VI_{t-1})$), a auto-correlação na primeira defasagem é de -0,63, seguida por 0,13, enquanto as demais se aproximam de zero. Este resultado se assemelha aos encontrados por Fu (2009) no mercado acionário americano.

As auto-correlações encontradas mostram, portanto, que a variável VI não segue um passeio aleatório. Para testar com mais robustez se a série segue um passeio aleatório deve-se fazer o teste de raiz unitária (teste de Dickey-Fuller), conforme descrito na Tabela 3.2 abaixo.

Tabela 3.2

Teste de passeio aleatório para as séries de volatilidade idiossincrática (VI)

	N	Média	Mediana	Q1	Q3	P.A. rej. %*
Modelo: $VI_{i,t+1} - VI_{i,t} = \alpha_i + \beta_i VI_{i,t} + \varepsilon_{i,t}$ $i = 1, 2, \dots, N$, $t = 1, 2, \dots, T_i$						
β	58	-0,59	-0,58	-0,69	-0,49	100%
$t(\beta)$	58	-5,16	-5,07	-5,77	-4,49	
Modelo: $\ln VI_{i,t+1} - \ln VI_{i,t} = \alpha_i + \beta_i \ln VI_{i,t} + \varepsilon_{i,t}$ $i = 1, 2, \dots, N$, $t = 1, 2, \dots, T_i$						
β	58	-0,59	-0,57	-0,67	-0,48	98%
$t(\beta)$	58	-5,11	-5,05	-5,67	-4,43	

Valores críticos do teste de Dickey-Fuller (1996)

<u>amostra</u>	<u>valor crítico da estatística-t</u>
25	-3,75
50	-3,59
100	-3,50

* porcentagem das ações para as quais se rejeita a hipótese de passeio aleatório

A Tabela 3.2 verifica, através de teste de raiz unitária, se as séries de volatilidade idiossincrática das ações amostrais seguem um passeio aleatório. São estimados dois modelos, o Modelo 1 e o Modelo 2, explicitados abaixo:

$$\text{Modelo 1: } VI_{i,t+1} - VI_{i,t} = \alpha_i + \beta_i VI_{i,t} + \varepsilon_{i,t}, \quad i = 1, 2, \dots, N, \quad t = 1, 2, \dots, T_i$$

$$\text{Modelo 2: } \ln VI_{i,t+1} - \ln VI_{i,t} = \alpha_i + \beta_i \ln VI_{i,t} + \varepsilon_{i,t}, \quad i = 1, 2, \dots, N, \quad t = 1, 2, \dots, T_i$$

O Modelo 1 é descrito pela variável VI em diferença de nível ($VI_{i,t+1} - VI_{i,t}$), enquanto o Modelo 2 é descrito por sua diferença em tempo contínuo ($\ln VI_{i,t+1} - \ln VI_{i,t}$). Para o caso da série temporal em questão ser um passeio aleatório, o parâmetro β da regressão não deve ser significativamente diferente de zero. Desta forma, para cada regressão são estimadas as estatísticas-t referentes a cada β , e as mesmas são comparadas com os valores críticos de Dickey-Fuller (FULLER, 1996).

A Tabela 3.2 mostra a média, mediana, e os quartis 1 e 3 referentes aos β 's encontrados para cada uma das 58 ações amostrais. São apresentadas também as estatísticas-t encontradas. A coluna “P.A. rej. %” mostra a porcentagem de ações para as quais foi rejeitada a hipótese de passeio aleatório em cada modelo, a um nível de significância de 1%: 100% no primeiro modelo e 98% no segundo. Desta forma, pode-se dizer que não é adequado representar esta variável através de um passeio aleatório.

3.3. Os Modelos ARCH

Os modelos ARCH (*autoregressive conditionally heteroscedastic*) são muito usados em Finanças para estimação da volatilidade, por apresentarem duas características importantes (BROOKS, 2008). Primeiramente, estes modelos não assumem que a variância dos erros é constante, isto é, são modelos heterocedásticos. Não é apropriado considerar a variância dos erros constante em séries temporais financeiras, por isso é preciso utilizar um modelo que não tenha esta premissa, e que mostre a evolução da variância dos erros.

A segunda característica está relacionada com um fenômeno chamado *volatility clustering*, que representa a tendência, presente em séries financeiras, de grandes variações de preços de ativos (positivas ou negativas) serem seguidas por grandes variações, e pequenas variações de preços de ativos (positivas ou negativas) serem seguidas de pequenas variações (BROOKS, 2008). Em outras palavras, a volatilidade tende a apresentar algum grau de auto-correlação.

No modelo ARCH, estas duas características estão presentes devido à forma como a variância condicional é modelada. Um exemplo de modelagem de série financeira usando um modelo ARCH(q) pode ser visto na Equação (3.2) abaixo. A variância dos erros da regressão hipotética é tida como dependente dos q *lags* dos erros quadrados. Com isso, ambas as características descritas acima são respeitadas, e é construída uma série de variância condicional, que representa a volatilidade da série financeira.

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_n x_{nt} + u_t, \quad u_t \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_q u_{t-q}^2 \quad (3.2)$$

Apesar de todas as vantagens, o modelo ARCH possui certas limitações. O valor de q pode ser difícil de ser determinado, e pode ser muito grande, tornando o modelo de variância condicional pouco parcimonioso (com muitos *lags*). Além disso, não há como impedir a existência de parâmetros negativos no modelo da variância condicional (α 's), o que poderia tornar a variância negativa.

Com isso, vale destacar uma extensão natural do modelo ARCH, que é o modelo GARCH (*generalised* ARCH), desenvolvido por Bollerslev (1986). Este modelo apresenta as mesmas propriedades do primeiro, mas é mais parcimonioso, e por isso corrige em parte as limitações descritas acima. No modelo GARCH, a variância condicional pode depender de seus próprios *lags*, além dos *lags* dos erros, e desta forma o modelo permite que informações sobre erros quadrados passados influenciem a variância corrente sem a necessidade de inclusão de múltiplos parâmetros. Um exemplo de modelo GARCH(p,q) pode ser visto na Equação (3.3) abaixo.

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_n x_{nt} + u_t, \quad u_t \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_q u_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \dots + \beta_p \sigma_{t-p}^2 \quad (3.3)$$

Entretanto, apesar de ser amplamente utilizado em séries financeiras, algumas críticas persistem ao modelo exemplificado acima. As condições de não-negatividade dos parâmetros, no caso da variância condicional, ainda podem não ser respeitadas, o que poderia gerar variâncias negativas.

Além disso o modelo não corrige o chamado *leverage effect*, muito comum em séries financeiras. O *leverage effect* diz respeito ao fato de que choques negativos têm um efeito maior sobre a volatilidade do que choques positivos, isto é, geram um aumento maior da volatilidade. De acordo com o modelo GARCH, choques positivos e negativos exerceriam a mesma influência sobre a variância.

Desta forma, vale destacar, por fim, uma extensão ao modelo GARCH: o modelo EGARCH (*exponential GARCH*), que foi proposto por Nelson (1991). Nesse caso, uma série de vantagens podem ser destacadas, como a impossibilidade de se gerar variâncias negativas, e a permissão da existência de assimetrias no modelo (*leverage effect*). O modelo EGARCH(p,q), em termos gerais, é descrito abaixo:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_n x_{nt} + u_t, \quad u_t \sim N(0, \sigma_t^2)$$

$$\ln(\sigma_t^2) = \omega + \sum_{l=1}^p b_l \ln(\sigma_{t-l}^2) + \sum_{k=1}^q c_k \left\{ \gamma \left(\frac{u_{t-k}}{\sqrt{\sigma_{t-k}^2}} \right) + \alpha \left[\frac{|u_{t-k}|}{\sqrt{\sigma_{t-k}^2}} - \sqrt{\frac{2}{\pi}} \right] \right\} \quad (3.4)$$

Na Equação (3.4), é o logaritmo da variância que é especificado, portanto, mesmo que os parâmetros do modelo sejam negativos, σ_t^2 será positiva.

Por apresentar propriedades que se adéquam perfeitamente à modelagem de volatilidade em séries financeira, este estudo utilizou o modelo EGARCH como uma forma alternativa de estimação da volatilidade idiossincrática. Como fornece uma série de variância condicional para cada modelo estimado, o modelo EGARCH é útil como meio de se estimar a volatilidade idiossincrática esperada (FU, 2009).

3.4.

Estimação da Volatilidade Idiossincrática Através do Modelo EGARCH (E(VI))

O segundo método de cálculo da volatilidade idiossincrática das ações da amostra foi o modelo EGARCH descrito acima. O modelo de Três Fatores de Fama e French (1993) é novamente utilizado, mas desta vez todos os dados são mensais. É estimada uma regressão para cada ação, abrangendo todo o período amostral. A regressão para esta etapa é a Equação (3.5) explicitada abaixo.

$$\begin{aligned}
R_{it} - r_t &= \alpha_i + b_i (R_{mt} - r_t) + s_i \text{SMB}_t + h_i \text{HML}_t + u_{it}, & u_t &\sim N(0, \sigma_{it}^2) \\
\ln(\sigma_{it}^2) &= \omega + \sum_{l=1}^p b_{i,l} \ln(\sigma_{i,t-l}^2) + \sum_{k=1}^q c_{i,k} \left\{ \gamma \left(\frac{u_{i,t-k}}{\sqrt{\sigma_{i,t-k}^2}} \right) + \alpha \left[\frac{|u_{i,t-k}|}{\sqrt{\sigma_{i,t-k}^2}} - \sqrt{\frac{2}{\pi}} \right] \right\}
\end{aligned}
\tag{3.5}$$

A variável dependente ($R_{it} - r_t$) é o excesso de retorno mensal de cada ação, ou seja, seu retorno mensal descontado do CDI. A variável $R_{mt} - r_t$ representa o excesso de retorno de mercado. O retorno de mercado mensal é calculado através da ponderação dos retornos mensais individuais segundo o valor de mercado de cada ação. As variáveis SMB e HML seguem a mesma metodologia descrita na seção 3.1. Todos os retornos foram calculados de forma contínua, isto é, por meio do logaritmo natural.

As regressões referentes a cada ação foram feitas utilizando-se o método EGARCH. Seguindo-se o método de Fu (2009), nove modelos EGARCH(p,q) foram estimados para cada ação: EGARCH(1,1), EGARCH(1,2), EGARCH(1,3), EGARCH(2,1), EGARCH(2,2), EGARCH(2,3), EGARCH(3,1), EGARCH(3,2) e EGARCH(3,3). Entre os modelos que convergiam, foi selecionado o que apresentava o menor critério de Akaike (AIC) para cada ação.

Na amostra de ações deste estudo, o modelo mais comum foi o EGARCH(2,2), que foi o selecionado para 13 das 58 das ações, seguido pelo EGARCH(3,1), presente em 10 casos.

Por fim, foi obtida a série de variância condicional individual de acordo com o modelo escolhido. Dos valores encontrados foi retirada a raiz quadrada, para torná-los equivalentes aos valores de volatilidade idiossincrática calculados segundo o primeiro método (desvio-padrão dos resíduos das regressões mensais segundo o modelo de Fama e French). Desta maneira, as novas séries temporais encontradas correspondem à volatilidade idiossincrática esperada de cada ação. Esta nova variável foi denominada E(VI).

As estatísticas descritivas de $E(VI)$, assim como da respectiva variável em tempo contínuo ($\ln(E(VI)_t/E(VI)_{t-1})$) estão apresentadas na Tabela 3.3 abaixo. Da mesma forma que na Tabela 3.1, foram computadas a média, o desvio-padrão, a assimetria e a curtose de cada variável para cada uma das 58 ações amostrais, e então foi calculada a média de cada uma dessas estatísticas. A média da volatilidade idiossincrática condicional entre as ações foi de 7,11%, com desvio-padrão médio de 3,53%.

Tabela 3.3

Estatísticas das séries de volatilidade idiossincrática condicional individuais (% mensal)

	N	Média	D. P.	Assim.	Curtose
$E(VI)$	58	7,11	3,53	1,08	5,05
$\ln(E(VI)_t/E(VI)_{t-1})$	58	-0,11	63,98	-0,13	3,00

3.5.

Demais Variáveis a Serem Incluídas nas Regressões *Cross-Section*

Para investigar a relação entre o retorno de uma ação e seu risco idiossincrático, foram feitas diversas regressões *cross-section*, utilizando-se também outras variáveis específicas às ações, além das duas variáveis de volatilidade idiossincrática calculadas anteriormente.

Em primeiro lugar, foi estimado o beta mensal de cada ação da amostra para todo o período analisado na pesquisa. Esta etapa foi feita através de regressões lineares simples mensais entre o retorno da ação e o retorno da carteira de mercado, formada pelas ações amostrais ponderadas segundo seu valor de mercado. Foram utilizados os retornos mensais dos últimos 60 meses, 36 meses, ou 24 meses, o que estivesse disponível. Esta variável foi chamada de BETA.

Adicionalmente, foi incluída uma segunda variável como beta das ações. A variável BETAIBOV corresponde ao beta calculado pelo sistema Economática®. A diferença é que, neste caso, a carteira de mercado considerada é o índice Bovespa: o sistema Economática® calcula o beta de cada ação através de regressões lineares simples entre o seu retorno e o retorno do Ibovespa. Também nesse caso, foram utilizados os retornos mensais dos últimos 60 meses, 36 meses, ou 24 meses, o que estivesse disponível.

Vale destacar que ambas as variáveis de beta estimadas seguem uma metodologia bastante simples. Nos artigos de Fu (2009) e Fama e French (1992), entretanto, o beta é o *post-ranking beta*, estimado através de regressões de carteiras igualmente ponderadas formadas com base na alocação de ações segundo o seu índice beta (nesse caso, o beta de mercado calculado através dos retornos dos últimos 60 meses) e seu tamanho. As ações são alocadas a dez carteiras diferentes segundo o seu beta de mercado, e a dez carteiras diferentes segundo o seu valor de mercado, formando 100 carteiras no total. O retorno destas carteiras, enfim, é regredido no retorno de mercado, a fim de se estimar o *post-ranking beta*, e este beta é então associado a cada ação presente na carteira correspondente.

Contudo, cada uma destas 100 carteiras deve possuir uma quantidade de ações suficiente para torná-la diversificada. Por isso, não é possível replicar esta metodologia no presente estudo, pois não há ações suficientes negociadas no mercado brasileiro para compor esta quantidade de carteiras. Desta forma, optou-se por um método mais parcimonioso de estimação do beta, através de regressões simples do retorno individual de cada ação no retorno da carteira de mercado.

Seguindo a metodologia de Fu (2009), outras quatro importantes variáveis são incluídas nas regressões *cross-section*, por serem historicamente utilizadas como variáveis explicativas na modelagem de retornos financeiros. São elas: o valor de mercado (tamanho), o índice *book-to-market*, uma variável para *momentum* e uma variável para liquidez. O valor de mercado e o índice *book-to-market* foram utilizados para análise dos retornos em Fama e French (1992), e estão presentes em diversos estudos sobre este tema. Já o efeito *momentum* e a liquidez são outras duas variáveis que passaram a estar presentes em estudos acadêmicos sobre modelagem de retornos

após Fama e French (1992), por terem sido consideradas significantes na explicação dos retornos *cross-section* das ações.

A variável VM representa o valor de mercado mensal das ações. Como pode ser visto na Tabela 3.4, esta variável é muito assimétrica, e portanto foi utilizado o logaritmo natural do valor de mercado, representado por $\ln(VM)$.

O índice *book-to-market*, representado por B/M, foi calculado por meio da metodologia de Fama e French (1992). Em dezembro de cada ano, o Patrimônio Líquido relativo a cada ação foi dividido pelo valor de mercado da mesma. O índice encontrado é referente a julho do ano seguinte e aos onze meses posteriores a ele. Esta diferença de seis meses garante que as os retornos reflitam as informações disponíveis, visto que as empresas costumam divulgar os resultados fiscais do final do ano cerca de três a quatro meses após o início do ano seguinte. As séries obtidas apresentaram também assimetria (Tabela 3.4), e portanto foi utilizado o seu logaritmo natural ($\ln(B/M)$).

O efeito *momentum*, muito citado em estudos de finanças comportamentais, é comumente observado em mercados acionários em todo o mundo. Ele descreve o fato de que ações ganhadoras (isto é, que tiveram desempenho positivo em determinado horizonte de tempo) tendem a continuar ganhadoras, e ações perdedoras (que tiveram desempenho negativo) tendem a continuar perdedoras. Ou seja, retornos passados tendem a prever retornos futuros (JEGADEESH e TITMAN, 1993).

O estudo de Fu (2009), que abrange ações negociadas em bolsas americanas, considera como variável de *momentum* o retorno acumulado da ação entre os meses t-7 e t-2, ou seja, o período de formação é de cinco meses. O mês t-1 é retirado do cálculo para evitar qualquer influência sobre o mês t causada por *thin trading* ou efeitos do *spread* de compra e venda⁴. Desta forma, para representar o efeito *momentum*, a variável $RET(-2;-7)$ é adicionada ao presente modelo. Ela é calculada através do logaritmo do retorno entre o mês t-7 e o mês t-2.

⁴ *Thin trading* quer dizer volume de transações abaixo do normal. Este fenômeno pode influenciar o preço dos ativos pois ordens grandes têm mais dificuldade de serem executadas. Da mesma forma, um grande *spread* entre os preços de compra e de venda pode causar alterações anormais nos preços dos ativos

Contudo, Bonomo e Lacerda (2007) encontram um efeito *momentum* significativo para o mercado acionário brasileiro utilizando-se um período de formação de três meses, um pouco mais curto do que o utilizado no modelo de Fu (2009). Desta forma, por ser mais adequada ao caso brasileiro, é incluída no presente modelo uma segunda variável para representar o efeito *momentum*, que considera o retorno acumulado do mês t-5 ao mês t-2. A nova variável é chamada de RET(-2;-5), e também é calculada através do logaritmo natural do retorno.

Como variável de liquidez, é usado um índice de giro, calculado através da razão entre a média de volume mensal negociado e a média do valor de mercado para cada ação nos últimos seis meses. Esta variável foi denominada VOL. A alta assimetria das séries obtidas, como pode ser visto na Tabela 3.4, também levou à adoção do seu logaritmo natural ($\ln(VOL)$).

Para completar, a variável RET, como descrito anteriormente, representa o logaritmo natural do retorno mensal de cada ação (retorno contínuo), EXRET, por sua vez, é o excesso de retorno, isto é, o retorno mensal descontado da taxa livre de risco (CDI), também calculado de forma contínua. Todas as variáveis estão apresentadas na Tabela 3.4 abaixo, com suas respectivas estatísticas descritivas (média, desvio-padrão, mediana, o primeiro e terceiro quartis e a assimetria). N representa o número de observações em cada variável (ação-mês). Todas as estatísticas foram calculadas utilizando-se as amostras combinadas das ações (*pooled sample*).

Tabela 3.4Estatísticas descritivas das variáveis (*pooled sample*)

Variáveis	Média	Desv-Pad	Mediana	Q1	Q3	Assimetria	N
RET	0,015	0,106	0,016	-0,044	0,077	-0,482	3828
EXRET	0,005	0,105	0,007	-0,052	0,066	-0,494	3828
VI	0,075	0,035	0,069	0,053	0,090	1,961	3828
E(VI)	0,071	0,043	0,062	0,043	0,088	2,526	3828
BETA	0,810	0,342	0,795	0,555	1,059	0,200	3746
BETAIBOV	0,819	0,335	0,826	0,548	1,034	0,236	3758
VM	26.940*	62.232*	9.188*	2.825*	17.916*	3,927	3828
ln(VM)	15,854	1,595	16,033	14,854	16,701	-0,127	3828
B/M	0,679	0,608	0,504	0,348	0,747	2,608	3828
ln(B/M)	-0,687	0,792	-0,686	-1,055	-0,291	-0,374	3828
lnRET(-2,-7)	0,060	0,274	0,076	-0,075	0,227	-0,710	3828
lnRET(-2;-5)	0,037	0,203	0,046	-0,063	0,159	-0,681	3828
VOL	0,039	0,033	0,032	0,016	0,052	1,782	3828
ln(VOL)	-3,651	1,070	-3,449	-4,127	-2,954	-1,150	3828

* dados em R\$ milhões