

# Modelagem da volatilidade condicional incorporando o período não regular do pregão ao modelo APARCH

Breno Valente Fontes Araújo<sup>1</sup>

 <https://orcid.org/0000-0002-8855-5830>  
E-mail: brenovfa@gmail.com

Marcos Antônio de Camargos<sup>1,2</sup>

 <https://orcid.org/0000-0002-3456-8249>  
E-mail: marcosac@face.ufmg.br

Frank Magalhães de Pinho<sup>1,2</sup>

 <https://orcid.org/0000-0002-5063-3389>  
E-mail: frank.pinho@ibmec.edu.br

<sup>1</sup> Universidade Federal de Minas Gerais, Faculdade de Ciências Econômicas, Centro de Pós-Graduação e Pesquisas em Administração, Belo Horizonte, MG, Brasil

<sup>2</sup> Ibmec BH, Departamento de Administração, Belo Horizonte, MG, Brasil

Recebido em 08.01.2018 – Desk aceita em 23.01.2018 – 2ª versão aprovada em 21.06.2018 – Ahead of print em 01.11.2018  
Editora Associada: Fernanda Finotti Cordeiro Perobelli

## RESUMO

O estudo busca avaliar como os períodos *after-market* e pré-abertura impactam a estimação da volatilidade condicional de um dia à frente. A volatilidade tem bastante destaque nos estudos de finanças, pois é parâmetro fundamental na precificação de derivativos, gestão de portfólios e de risco. Os resultados são relevantes para que agentes de investimentos possam refinar os modelos de previsão da volatilidade e obter melhores resultados na precificação de derivativos, na gestão de risco, na composição e otimização de carteiras. Utilizou-se o modelo *asymmetric power autoregressive conditional heteroskedasticity* (APARCH), incorporando o período *after-market*, o leilão de pré-abertura e o *overnight* total, para avaliar se eles carregam informações relevantes para a modelagem da volatilidade. Analisaram-se as 20 ações de empresas brasileiras listadas na Bolsa de Valores, Mercadorias e Futuros de São Paulo (BM&FBOVESPA) e pertencentes ao índice Dow Jones Brazil Titans 20 ADR Index (BR Titans 20) com ADRs (*American depositary receipts*) listados nas bolsas de Nova York e na Nasdaq. Os resultados foram avaliados na amostra pelo critério de informação Akaike corrigido (AICc) e pela significância estatística dos coeficientes, e fora da amostra pelos critérios raiz dos erros quadráticos médios (*root mean squared error* – RMSE), erro percentual médio absoluto (*mean absolut percentage error* – MAPE),  $R^2$  da regressão de Mincer e Zarnowitz e teste de Diebold Mariano. A análise não permite afirmar o melhor modelo, pois não há unanimidade entre todas as ações. Entretanto, os períodos não regulares do pregão demonstraram incorporar informações relevantes para a maior parte das ações. Ademais, os modelos que incorporaram o período pré-abertura obtiveram, em geral, resultados superiores aos demais, demonstrando que tal período carrega informações relevantes para a previsão da volatilidade condicional.

**Palavras-chave:** volatilidade condicional e realizada, modelo APARCH, dados intradiários, *after-market*, pré-abertura, gestão de risco.

## Endereço para correspondência

Breno Valente Fontes Araujo

Universidade Federal de Minas Gerais, Faculdade de Ciências Econômicas, Centro de Pós-Graduação e Pesquisas em Administração  
Avenida Presidente Antônio Carlos, 6627, Sala 4012 – CEP: 31270-901  
Pampulha – Belo Horizonte – MG – Brasil

\*Trabalho apresentado no XLI Encontro da Associação Nacional de Pós-Graduação e Pesquisa em Administração, São Paulo, SP, Brasil, outubro de 2017.



## 1. INTRODUÇÃO

Diversos estudos no Brasil e no exterior utilizam dados diários para previsão da volatilidade condicional do dia seguinte. No Brasil, a maior parte desses estudos desenvolveu-se nos últimos 20 anos, apontando para o bom desempenho dos modelos da família ARCH (*autoregressive conditional heteroskedasticity*) para previsão da volatilidade (Cavaleri & Ribeiro, 2011; Ceretta & Costa, 2001; Gaio, Pessanha, Oliveira & de Ázara, 2007; Galdi & Pereira, 2007; Silva, C. A. G., 2009). Entretanto, a maior parte dos estudos que busca modelar a volatilidade condicional de ações ou índices desconsidera a variação que ocorre entre o período de abertura de um dia e o fechamento do dia anterior, também conhecido como período *overnight*. No Brasil, foram encontrados alguns estudos que buscaram avaliar a significância das informações encontradas no *overnight* (Accioly & Mendes, 2015; Souza, 2004), enquanto no exterior foram encontrados diversos estudos que debatem o tema (Barclay & Hendershott, 2004; Chen, Yu & Zivot, 2012; Gallo & Pacini, 1998; Martens, 2002; Taylor, 2007).

A Instrução Normativa n. 358 da Comissão de Valores Mobiliários (CVM, 2002), em seu artigo 5º, determina que a divulgação de ato ou fato relevante deve ocorrer, sempre que possível, antes do início ou após o encerramento dos negócios nas bolsas de valores. Nesse sentido, é provável que informações relevantes sejam divulgadas no período não regular do pregão e se reflitam no *after-market* ou no preço de abertura das ações. Tais informações tendem a influenciar a dinâmica do mercado durante o pregão regular, impactando diretamente na volatilidade das ações.

Nicolau (2012) aponta que a chegada de informação de forma intensa tende a aumentar a volatilidade dos retornos. O autor sugere inserir variáveis explicativas ao modelo de previsão de volatilidade condicional quando essas variáveis incorporarem informações que impactam na volatilidade. Zivot (2009) destaca estudos que identificam variáveis explicativas que, incorporadas aos modelos da família GARCH, melhoram os resultados de previsão, tais como volume transacionado, anúncios de dados macroeconômicos, retorno *overnight*, volatilidade *after-hours*, volatilidade implícita nos preços de opções e volatilidade realizada.

Com o avanço da tecnologia e a disponibilidade de dados em maior frequência, surgiram estudos com novas formas de modelagem e previsão da volatilidade. Trabalhos recentes, tanto no exterior quanto no Brasil, utilizam dados intradiários em alta frequência para cálculo da volatilidade realizada. Tal variável é utilizada como

medida “observável” da volatilidade de um dia, o que proporciona uma análise *out-of-sample* mais eficaz do que outras medidas utilizadas, como o próprio retorno ao quadrado do dia (Andersen & Bollerslev, 1998). Além disso, os dados intradiários permitem analisar o comportamento do preço no período não regular do pregão de forma fracionada, pois existem informações específicas do período *after-market* e do período pré-abertura.

Acredita-se, então, que durante o período *overnight* ocorrem fatos relevantes para a modelagem da volatilidade condicional de uma ação e que essas informações podem estar contidas tanto no *after-market* quanto no período de pré-abertura, refletidas no preço de abertura. Dessa forma, o objetivo deste trabalho é avaliar como os períodos de *after-market*, pré-abertura e *overnight* total impactam sobre a estimação da volatilidade condicional das empresas brasileiras listadas na Bolsa de Valores, Mercadorias e Futuros de São Paulo (BM&FBOVESPA) e pertencentes ao Dow Jones Brazil Titans 20 ADR Index (BR Titans 20).

O presente trabalho se diferencia dos demais principalmente ao fragmentar o retorno do *overnight* em subperíodos, assim como feito por Chen et al. (2012). Porém, devido às peculiaridades do mercado brasileiro, o período *overnight* foi dividido em dois subperíodos, em vez de três: o *after-market* e o pré-abertura. Para isso, utilizou-se o modelo APARCH, incorporando a variação do período *after-market* em relação ao fechamento do pregão e a variação do preço de abertura em relação ao preço de fechamento do *after-market* do dia anterior (variação do período pré-abertura), além da variação do *overnight* total (variação entre preço de abertura e preço de fechamento do pregão regular do dia anterior) como variáveis explicativas do modelo.

A pesquisa visa, então, a contribuir para a literatura de três maneiras: (i) apresentando a análise do período *overnight* ainda pouco estudado no Brasil; (ii) utilizando subperíodos do horário não regular do pregão como variáveis explicativas para modelar a volatilidade condicional; (iii) realizando a análise *out-of-sample*, utilizando, como parâmetro de comparação, a volatilidade realizada, calculada a partir de dados intradiários.

O artigo tem a seguinte estrutura: após essa introdução, a seção 2 apresenta o referencial teórico de suporte à pesquisa empírica. A metodologia e os resultados são apresentados nas seções 3 e 4, respectivamente, e encerra-se com as considerações finais na seção 5, seguidas das referências.

## 2. REFERENCIAL TEÓRICO

### 2.1. Volatilidade Condicional – Modelos da Família ARCH

As séries de retornos de ativos financeiros apresentam, como propriedade característica, a não correlação serial, mas a existência de dependência em sua estrutura. Ou seja, os retornos passados não influenciam o retorno presente, mas a volatilidade é correlacionada aos retornos passados ou às “inovações” (resíduos) em torno da equação da média (Tsay, 2010). Em trabalho seminal, Engle (1982) propôs o modelo ARCH em que a variância condicional pode ser modelada por meio de uma função quadrática. Os estimadores dos parâmetros dos modelos da família ARCH, assim como os demais da família expostos a seguir, podem ser obtidos pelo estimador de máxima verossimilhança condicional (Engle, 1982; Tsay, 2010).

A equação da volatilidade condicional, na maioria dos casos, exige um modelo ARCH de ordem elevada (muitos parâmetros) para ser descrito adequadamente (Tsay, 2010). Tal fato pode gerar problemas de estimação durante a convergência do algoritmo de otimização (Nicolau,

2012). Para contornar essa limitação do modelo ARCH, Bollerslev (1986) propôs um modelo alternativo para a modelagem das “inovações” (resíduos) de uma série de retorno de ativos, mais conhecido como modelo GARCH.

Ao longo dos anos, diversos modelos foram sendo desenvolvidos na busca de incorporar diferentes comportamentos das séries de dados financeiros: IGARCH, EGARCH, GJR-GARCH etc. Neste estudo, optou-se por utilizar o modelo APARCH devido ao fato de este ser capaz de se comportar como os outros sete modelos da família ARCH.

O modelo *asymmetric power autoregressive conditional heteroskedasticity*, o APARCH, apresentado por Ding, Granger e Engle (1993), surgiu do questionamento de que a variância condicional não necessariamente segue uma função quadrática ou linear. Desse modo, o modelo oferece uma forma geral em que a potência da equação da variância condicional também é estimada, assim como os parâmetros tradicionais, como *alpha* e *beta*. Segundo esses autores, o modelo APARCH pode ser representado pela equação 1:

$$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i (|\epsilon_{t-i}| - \gamma_i \epsilon_{t-i})^\delta + \sum_{j=1}^p \beta_j \sigma_{t-j}^\delta \quad [1]$$

em que  $\omega$ ,  $\alpha_i$ ,  $\gamma_i$ ,  $\delta$ , e  $\beta_j$  são parâmetros a serem estimados pelo modelo. Como nos demais modelos,  $\omega$  é o intercepto do modelo, que retrata o nível médio da variância condicional, ou seja, pode ser considerada a variância incondicional.  $\alpha_i$  e  $\beta_j$ , assim como apresentado anteriormente, representam o quanto o choque (inovação) impacta sobre a variância condicional e o quanto a própria variância condicional defasada persiste no período corrente, respectivamente.  $\gamma_i$ , assim como no modelo GJR-GARCH, capta a resposta assimétrica da variância condicional a choques positivos e negativos, também conhecido como efeito alavancagem, ou seja, se choques positivos e negativos impactam diferentemente sobre a variância condicional de um período à frente. Se  $\gamma_i$  for estatisticamente significativo e positivo, indica a existência do efeito alavancagem, ou seja, choques negativos têm impacto maior sobre a variância condicional de um dia à frente. Caso  $\gamma_i$  seja estatisticamente significativo e negativo, indica que choques positivos têm maior impacto sobre a variância condicional. Por fim,  $\delta$  permite estimar outras potências

para a equação da variância condicional por meio de uma transformação Box-Cox do  $\sigma_t$ .

O modelo APARCH pode ser considerado um dos mais promissores da família ARCH, pois é capaz de abranger ao menos sete modelos dessa família, como pode ser visto nos seguintes casos especiais: ARCH ( $\delta = 2$ ,  $\gamma_i = 0$  e  $\beta_j = 0$ ), GARCH ( $\delta = 2$ ,  $\gamma_i = 0$ ), ARCH não linear ( $\delta = 0$  e  $\beta_j = 0$ ), GARCH de Taylor/Schwartz's ( $\delta = 1$ ,  $\gamma_i = 0$ ), TARARCH ( $\delta = 1$  e  $\beta_j = 0$ ), Log-ARCH ( $\delta \rightarrow 0$ ) e GJR-GARCH ( $\delta = 2$ ). Dessa forma, opção pelo modelo APARCH se deu devido ao fato de este estudo não ter como objetivo avaliar os diversos modelos, e sim o impacto das variáveis exógenas.

### 2.2. Volatilidade Realizada

A grande dificuldade em avaliar a validade da estimação da volatilidade se dá pelo fato de ela não ser observável. Faz-se necessária, então, uma técnica para medição da volatilidade diária *a posteriori* para avaliar o poder de previsão dos modelos. Andersen e Bollerslev

(1998) indicam que o retorno ao quadrado do dia, apesar de pouco eficaz, poderia ser utilizado como *proxy* para a volatilidade diária, como parâmetro de comparação para análise *out-of-sample* dos modelos. Estudos como os de Cumby, Figlewski e Hasbrouck (1993), Jorion (1995) e Figlewski (1997) constataram que os modelos GARCH apresentam resultados ruins para estimação da volatilidade, não por deficiência do modelo, mas pelo fato de o retorno ao quadrado de um dia não ser uma boa *proxy* a ser utilizada para a volatilidade diária da ação, visto que o preço de uma ação pode atingir valores altos e baixos em relação ao preço de abertura e depois voltar para o preço de abertura, chegando-se à volatilidade igual a 0.

Para contornar tal problema quanto à determinação da variável a ser utilizada como *proxy* para volatilidade diária, Andersen e Bollerslev (1998) apresentaram o conceito de volatilidade realizada, ou volatilidade percebida. Tal medida pode ser estimada pelo somatório dos retornos ao quadrado de dados intradiários (alta frequência). Os autores mostram que tal medida é muito mais confiável para se utilizar como *proxy* da volatilidade diária, pois se aproxima da volatilidade integrada do dia. No presente trabalho foi utilizada essa medida como *proxy* para a volatilidade diária a ser comparada com a volatilidade condicional estimada pelos modelos.

A volatilidade realizada pode ser descrita da seguinte forma simplificada, como indicado por Andersen, Bollerslev, Diebold e Ebens (2001a), Andersen, Bollerslev, Diebold e Labys (2001b) e Bollerslev e Wright (2001):

$$r_{t,i} = P_{t,i} - P_{t,i-1} \quad 2$$

$$VR_t^2 = \sum_{i=1}^n r_{t,i}^2 \quad 3$$

em que  $P$  é o logaritmo do preço,  $i$  é a fração do pregão regular, nesse caso a cada 15 minutos,  $r_{t,i}$  é o log-retorno do  $i$ -ésimo intervalo de 15 minutos do dia,  $n$  é o número de observações para cada dia e  $VR^2$  é a variância realizada do dia.

Andersen, Bollerslev, Diebold e Labys (2001c) indicaram que quanto maior a frequência dos dados intradiários, mais perto a volatilidade realizada se aproxima da volatilidade integrada, que pode ser considerada efetivamente realizada em um determinado horizonte de tempo. Entretanto, os autores apontaram que a utilização de dados “contínuos” para estimação da volatilidade realizada pode acarretar em grandes vieses, devido à existência de atritos de microestrutura de mercado, como salto do preço (*bid-ask bounce*), baixa frequência de transação, dentre outros.

Andersen et al. (2001c) propõem, então, a amostragem em intervalos de 5 minutos para amenizar tais problemas de microestrutura. Existe um debate na literatura sobre qual a melhor janela para o cálculo da volatilidade realizada. Os trabalhos empíricos na literatura indicam que a frequência ótima para o cálculo encontra-se entre 5 e 25 minutos (Mota & Fernandes, 2004). Oomen (2001) apresenta, como frequência ótima, intervalos de 25 minutos. Já Giot e Laurent (2004) encontraram a frequência ótima de 15 minutos para seu estudo. Optou-se por utilizar a frequência de 15 minutos, devido à disponibilidade dos dados e à utilização da mesma frequência em inúmeros estudos nacionais (Milach, 2010; Moreira & Lemgruber, 2004; Mota & Fernandes, 2004; Reis, 2011; Silva, J. C., 2002).

### 2.3. Funcionamento da BM&FBOVESPA e o Período Overnight

Conforme se observa na Tabela 1, a BM&FBOVESPA apresenta dois horários de negociação de ações: (i) pregão regular, que ocorre entre 10h e 17h, e (ii) *after-market*, que ocorre entre 17h30min e 18h, sendo a única forma de negociação após o fechamento do pregão, representando, devido a isso, um dos meios de variação entre os valores de fechamento e abertura, justificando o estudo do funcionamento e suas particularidades. Além do *after-market*, o período de pré-abertura também contribui diretamente para a variação entre o preço de fechamento do dia anterior e o preço de abertura do dia seguinte, haja vista que, apesar de não ocorrer a transação, as ordens já são casadas, refletindo no preço de abertura.

**Tabela 1**

*Funcionamento da Bolsa de Valores, Mercadorias e Futuros de São Paulo (BM&FBOVESPA)*

Mercado	Cancelamento de ofertas		Pré-abertura		Negociação		Call de fechamento		After-market				
									Cancelamento de ofertas		Negociação		
	Início	Fim	Início	Fim	Início	Fim	Início	Fim	Início	Fim	Início	Fim	
Mercado à vista	09h30min	09h45min	09h45min	10h	10h	16h:55min	16h55min	17h		17h25min	17h30min	17h30min	18h

Fonte: Adaptada de BM&FBOVESPA (2014).

O limite estipulado de variação no *after* de uma ação de 2%, seja para mais ou para menos, em relação ao preço de fechamento, restringe a análise do trabalho, pois podem haver dias em que a variação do *after-market* chegue a 2% e não ultrapasse por impossibilidade, não representando a real situação das ordens de compra e venda. Para contornar tal situação, verificou-se o número de vezes em que a variação do período *after-market* se aproximou de 2%. Constatou-se que para nenhuma empresa isso ocorreu em mais de 1% das observações.

Decidiu-se, então, não avaliar tal fato nos modelos estimados, devido ao baixo número de observações relevantes.

O período *overnight*, ou seja, o período entre o fechamento de um dia e a abertura do dia seguinte, vem sendo foco de muitos estudos ao longo dos anos. A maior parte dos estudos avalia como as informações desse período impactam no comportamento do mercado no pregão regular e sobre a volatilidade de ações e índices, conforme se observa na Tabela 2.

**Tabela 2**

*Evidências empíricas analisando o período overnight na literatura nacional e internacional*

Autor (ano)	Objetivo	Conclusão
Accioly e Mendes (2015)	Avaliar a inserção da volatilidade realizada como variável exógena ao modelo GARCH e incorporar o retorno ao quadrado do período <i>overnight</i> .	Concluíram que o retorno do período <i>overnight</i> tem poder explicativo em alguns casos, mas apresentou menor poder do que a abordagem de um fator apresentada por eles.
Chen et al. (2012)	Avaliar o modelo GARCH tradicional na previsão da volatilidade condicional, a partir de dados intradiários, para um modelo que abrange também o período não regular do pregão para as 30 ações mais líquidas da <i>Nasdaq</i> .	O pós-fechamento e a variação <i>overnight</i> apresentam pouca explicação sobre a volatilidade condicional, enquanto o período pré-abertura apresenta significância estatística sobre tal variável.
Taylor (2007)	Avaliar o valor econômico das informações do período <i>overnight</i> para agentes que trabalham com gestão de risco.	As informações do <i>overnight</i> têm impacto significativo sobre a volatilidade condicional dos ativos analisados, promovendo modelos mais precisos para gestão do risco.
Souza (2004)	Analisar se a incorporação do efeito <i>overnight</i> ao modelo GARCH leva à redução na persistência de volatilidade com dados diários das oito ações mais líquidas da BM&FBOVESPA.	Detectou-se o efeito redução na persistência da volatilidade para essas ações. Entretanto, não foi possível concluir sobre o melhor modelo de estimação, devido aos diferentes resultados para cada uma.
Barclay e Hendershott (2003)	Avaliar como as informações divulgadas durante as 24 horas de um dia impactam na quantidade, no preço e quando as negociações são realizadas, tendo como foco o <i>after-market</i> .	Os preços são mais eficientes e mais informações são reveladas por hora durante o pregão regular do que no <i>after-market</i> . Entretanto, a pequena negociação no <i>after-market</i> pode evidenciar significativas explicações para o preço das ações.
Martens (2002)	Examinar se o modelo GARCH, ao incluir diversas formas da volatilidade <i>after-hours</i> , melhora as previsões da volatilidade.	Concluíram que essa inclusão não apresenta melhora significativa para o modelo.
Gallo e Pacini (1998)	Avaliar se as variações entre o preço de abertura de um dia e o preço de fechamento do dia anterior têm poder explicativo sobre a volatilidade condicional de diferentes índices.	Destacaram que, ao prever a volatilidade fora da amostra ( <i>out-of-sample</i> ), ao se adicionar essa variável, o modelo apresentou resultados superiores ao modelo GARCH tradicional.

BM&FBOVESPA = Bolsa de Valores, Mercadorias e Futuros de São Paulo.

**Fonte:** Elaborada pelos autores.

Além das evidências aqui apresentadas, Barclay e Hendershott (2004) destacam que, após o pregão regular, as ações tendem a representar mais informações privadas do que durante o pregão, e que os agentes que negociam no *after-market* tendem a ser mais profissionais e representam instituições. Os autores ainda indicaram que as transações

realizadas no período após o pregão regular se tornam importantes apenas quando apresentam atividades transacionais suficientes. Ou seja, não faz sentido analisar ações que apresentem pouca variação fora do pregão regular, o que também justifica a escolha pela amostra das empresas brasileiras pertencentes ao índice BR Titans 20.

### 3. METODOLOGIA

A amostra analisada é constituída das empresas listadas na BM&FBOVESPA e pertencentes ao índice BR Titans 20, calculado desde 2004 a partir dos ADRs de empresas

brasileiras mais negociados na bolsa de Nova York e Nasdaq, apresentadas na Tabela 3.

**Tabela 3**

*Companhias brasileiras pertencentes ao índice Dow Jones Brazil Titans 20 ADR Index (BR Titans 20)*

Nome da companhia	Código	Nome da companhia	Código	Nome da companhia	Código
Ambev S/A	ABEV3	Copel	CPLE6	Oi S/A	OIBR4
Banco Bradesco	BBDC4	CSN	CSNA3	Petrobras S/A	PETR4
Banco Santander	SANB11	CPFL S/A	CPFE3	Telefônica S/A	VIVT4
BRF S/A	BRFS3	Embraer S/A	EMBR3	Tim Participações	TIMP3
Pão de Açúcar	PCAR4	Fibria Celulose S/A	FIBR3	Ultrapar Participações S/A	UGPA3
Sabesp	SBSP3	Gerdaul S/A	GGBR4	Vale S/A	VALE5
Cemig	CMIG4	Itaú-Unibanco S/A	ITUB4		

**Fonte:** Elaborada pelos autores.

Os dados utilizados nesta pesquisa são intradiários, obtidos junto a um grupo de pesquisa vinculado a um departamento de computação da universidade de origem dos autores. Estão ajustados para proventos e outros eventos corporativos (grupamento/desdobramentos) que afetam o preço diretamente sem necessidade de transações, e foram tratados de forma a conter informações do preço das ações a cada 15 minutos para o pregão regular e *after-market*.

O período de análise foi do dia 1 de janeiro de 2010 ao dia 20 de março de 2015 para nove ações, totalizando cerca de 1.290 observações diárias, e até 24 de julho de 2015 para as outras 11 ações, totalizando 1.375 observações diárias. Essa diferença temporal é decorrente da disponibilidade dos dados obtidos. A escolha desse período está relacionada com a contemplação de momentos de crise no sistema financeiro, tal como a crise da solvência de dívidas na Europa e os problemas políticos associados à dívida pública no Brasil, o que permite incorporar períodos de alta e baixa volatilidade.

De maneira semelhante à janela utilizada por Chen et al. (2012), o período *out-of-sample* analisado neste estudo foi aproximadamente de um ano (260 dias). A análise *in-sample* foi realizada com todos os demais dados da amostra, pois, segundo Ng e Lam (2006), janelas *in-sample* com aproximadamente mil observações minimizam os impactos na estimação dos coeficientes dos modelos da família GARCH.

Para a avaliação *out-of-sample*, ou seja, da previsão da volatilidade condicional um dia à frente, definiu-se,

como estratégia de análise de dados, a *rolling* (rolagem) recursiva; a cada nova previsão da variância condicional, o modelo foi estimado novamente, contemplando a nova observação. Por exemplo, para previsão da volatilidade do período 1.001, são utilizadas as 1.000 observações anteriores para estimar o modelo; já para a previsão da volatilidade do período 1.002, são utilizadas as 1.001 observações anteriores, e assim sucessivamente.

A seguir destacam-se as variáveis exógenas avaliadas e como foram calculadas:

- Variação do *after-market* (AM): a variação logarítmica do preço de fechamento do *after-market* em relação ao preço de fechamento do pregão regular;
- Variação do período pré-abertura (OP): a variação logarítmica do preço de abertura de um dia em relação ao preço de fechamento do *after* do dia anterior;
- Variação do *overnight* total (OV): a variação logarítmica do preço de abertura de um dia em relação ao preço de fechamento do pregão regular do dia anterior.

Como o interesse da análise é na variação absoluta dos valores, ou seja, independe se a variação é positiva ou negativa, foram utilizadas as variações elevadas ao quadrado. Foram avaliados, portanto, quatro diferentes modelos para cada ação: um APARCH tradicional e outros três incorporando cada uma das variáveis exógenas, como indicado na Tabela 4.

**Tabela 4**

Modelos APARCH (asymmetric power autoregressive conditional heteroskedasticity) avaliados

APARCH	$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i ( \epsilon_{t-i}  - \gamma_i \epsilon_{t-i})^\delta + \sum_{j=1}^p \beta_j \sigma_{t-j}^\delta$
APARCH + AM	$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i ( \epsilon_{t-i}  - \gamma_i \epsilon_{t-i})^\delta + \sum_{j=1}^p \beta_j \sigma_{t-j}^\delta + \partial AM_{t-1}^2$
APARCH + OP	$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i ( \epsilon_{t-i}  - \gamma_i \epsilon_{t-i})^\delta + \sum_{j=1}^p \beta_j \sigma_{t-j}^\delta + \partial OP_t^2$
APARCH + OV	$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i ( \epsilon_{t-i}  - \gamma_i \epsilon_{t-i})^\delta + \sum_{j=1}^p \beta_j \sigma_{t-j}^\delta + \partial OV_t^2$

As variáveis estão descritas no texto.

**Fonte:** Elaborada pelos autores.

Com objetivo de refinar os modelos e seguindo a recomendação de Tsay (2010), estimou-se cada um deles até a ordem 2, ou seja, (1,1), (1,2), (2,1) e (2,2), a fim de identificar aquele que melhor se adequa às séries. Antes das estimações, realizou-se o teste proposto por Ljung e Box (1978) para verificação de autocorrelação nas séries de *log*-retornos diários. Um modelo autorregressivo – ARMA (1,1) foi ajustado para equação da média, de forma a garantir que os resíduos (inovações) fossem independentes e identicamente distribuídos (i.i.d.). Feito isso, realizou-se o teste ARCH para comprovação da existência de heteroscedasticidade. A partir desse ponto, estimaram-se os modelos da volatilidade condicional por meio do modelo APARCH.

Inicialmente, estimou-se o modelo APARCH (1,1) para as diferentes distribuições das “inovações” (resíduos), de forma a identificar aquela que melhor se adequa à série de dados. Posteriormente, ajustaram-se os diversos modelos para estimação da volatilidade condicional, com e sem as variáveis exógenas. Os modelos que não apresentaram resultados satisfatórios nos testes para seu ajuste não foram avaliados, enquanto os demais foram avaliados, inicialmente, dentro da amostra (*in-sample*). Aqueles que apresentaram melhor resultado na análise dentro da amostra em cada grupo foram selecionados para análise fora da amostra (*out-of-sample*) de forma a identificar os que apresentam melhores resultados para previsão da volatilidade de um dia à frente.

Para estimação dos modelos, foram consideradas duas diferentes distribuições dos resíduos em torno da equação da média: *t* de Student assimétrica e generalizada (GED) assimétrica. A não avaliação da distribuição normal deve-se ao fato de que ela é um caso particular da distribuição GED. De forma similar, as distribuições assimétricas abrangem ambas as possibilidades, com e sem assimetria. Além disso, um dos fatos estilizados já comprovados na literatura de finanças é o comportamento assimétrico das séries de dados, o que sugere uma melhor adequação para essas distribuições. A distribuição assimétrica *t* de Student foi a mais apropriada para todas as séries de *log*-retornos, menos para o papel da Vale, em que a distribuição assimétrica GED mostrou-se mais adequada.

Para checar se os modelos estão bem especificados, as “inovações” padronizadas devem formar uma sequência de variáveis aleatórias i.i.d., em que o efeito da heteroscedasticidade tenha sido eliminado. Como indicado por Tsay (2010), utilizou-se o teste de Ljung-Box para verificar se os resíduos (inovações) padronizados são i.i.d., comprovando que a equação para média está adequada, e o multiplicador de Lagrange (LM) ou o Ljung-Box dos resíduos padronizados ao quadrado para testar se o efeito ARCH foi controlado, comprovando que a equação da variância condicional é adequada. O cálculo das “inovações” padronizadas  $\tilde{\epsilon}_t$ , ou resíduos padronizados, é feito da seguinte forma:

$$\tilde{\varepsilon}_t = \frac{\varepsilon_t}{\sigma_t} \quad 4$$

Para análise dentro da amostra, utilizou-se o critério de Akaike corrigido (AICc), proposto por Akaike (1974), uma função em que se penaliza a qualidade do modelo ajustado segundo o número de parâmetros estimados. A equação do cálculo do AICc é:

$$AIC = -2L + 2p \quad 5$$

em que  $L$  representa o valor máximo da função de *log* verossimilhança do modelo e  $p$  o número de parâmetros estimados.

Hurvich e Tsai (1993) propuseram uma correção no AIC, denominada AICc, que pode ser calculada a partir da seguinte transformação:

$$AICc = AIC + \frac{2p(p+1)}{n-p-1} \quad 6$$

em que  $n$  seria o tamanho amostral do conjunto de dados utilizados na estimação do modelo. O AICc penaliza a

$$(VR_{t+k}^2)^{1/2} = \alpha_0 + \alpha_1(\sigma_{t+k|t}^2)^{1/2} + u_{t+k} \quad 7$$

em que  $VR_{t+k}^2$  é a variância realizada do dia  $t+k$  e  $\sigma_{t+k|t}^2$  se refere à variância condicional estimada para o dia  $t+k$  com base nas informações disponíveis no dia  $t$ . Na regressão de Mincer-Zarnowitz, se a volatilidade condicional está bem estimada, deve-se ter  $\alpha_0$  e  $\alpha_1$  iguais a 0 e 1, respectivamente, com boa significância estatística. Entretanto, esses coeficientes podem sofrer com problema de erros de medida das variáveis, dificultando sua interpretação (Andersen & Bollerslev, 1998). Não obstante, os autores indicam que o  $R^2$  da regressão pode ser utilizado para avaliar a medida utilizada como volatilidade *ex post*, nesse caso expressa pela volatilidade realizada, explicada pela volatilidade condicional.

Outra métrica de avaliação utilizada é a RMSE dos valores estimados pelos modelos em relação à volatilidade realizada do dia. A RMSE é utilizada para indicar quão distante, em média, o conjunto de estimativas está do parâmetro previsto:

$$RMSE = \left[ \frac{1}{T} \sum_{t=1}^T (VR_t^2 - \sigma_t^2)^2 \right]^{1/2} \quad 8$$

Avaliou-se, também, pelo método do MAPE, que tem como uma das vantagens a fácil interpretação, já que a

incorporação de parâmetros e, caso a amostra seja muito grande, o AICc tende para o AIC tradicional. Burnham e Anderson (2004) recomendam a utilização do AICc em detrimento do AIC, principalmente em casos em que o número de observações é pequeno. Portanto, será utilizado o AICc para análise dentro da amostra e determinação dos melhores modelos que se adequem à amostra.

Para ambos os critérios, quanto menor o valor, melhor, ou seja, serão selecionados como melhores aqueles que apresentarem menor valor para o AICc. Após a verificação dos modelos que melhor se adequam a cada série de dados, foram utilizados os melhores de cada grupo para estimação da volatilidade condicional um período à frente para a análise *out-of-sample*.

Na análise *out-of-sample*, foram utilizadas três técnicas: a regressão de Mincer e Zarnowitz (1969), a raiz dos erros quadráticos médios (*root mean squared error* – RMSE) e o erro percentual médio absoluto (*mean absolut percentage error* – MAPE). A ideia da regressão de Mincer-Zarnowitz (MZ) é simples, basta regredir a volatilidade realizada (observável) em função da volatilidade condicional estimada pelos modelos. A regressão é formalmente descrita da seguinte maneira:

escala é em porcentagem, e como principal desvantagem o fato de que se o valor realizado for muito pequeno ou 0, o valor do MAPE “explode” ou não é possível se calcular.

$$MAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{VR_t - \sigma_t}{VR_t} \right| \cdot 100 \quad 9$$

Por fim, realizou-se o teste de Diebold Mariano (1995) para comparar se os erros de previsão, ou seja, a diferença entre valor estimado e observado dos modelos alternativos eram estatisticamente diferentes do modelo tradicional.

A hipótese nula é que não há diferença estatística entre os erros, ou seja, os modelos são, em média, igualmente precisos. Neste trabalho, utilizaram-se, como hipótese alternativa, dois cenários: (i) os modelos são estatisticamente diferentes; (ii) o modelo alternativo é superior ao modelo tradicional. Dessa forma, permitiu-se avaliar não somente se há diferença estatística, mas também se o modelo alternativo gera resultados superiores ou inferiores.

Neste estudo foi possível, então, avaliar os modelos *in-sample* e *out-of-sample*.



## 4. ANÁLISE DOS RESULTADOS

Os resultados dos testes de adequação do modelo ajustado, AICc e o p-valor dos testes realizados para diferentes defasagens em cada ação (teste Ljung-Box – resíduos padronizados e padronizados ao quadrado, teste ARCH), foram omitidos. Foram selecionados os modelos com melhor (menor) valor do AICc. O comportamento dos coeficientes de cada ação ( $\mu$ , AR1, MA1,  $\alpha$  e  $\beta$  – sem e com defasagens, Gamma Vexog, Skew e Shape) foi analisado, tendo como referência a significância de

5% (p-valor > 0,05).

Novamente, devido ao limite de páginas, optou-se por fazer uma análise geral de todas ações, ordenadas pelas que tiveram maiores e menores variações nos períodos não regulares do pregão. Nesse sentido, calculou-se a média diária de variação absoluta das variáveis exógenas, ou seja, AM, OP e OV. A Tabela 5 indica, portanto, o valor médio absoluto de variação para esses três períodos, respectivamente.

**Tabela 5**

*Variação média diária das variáveis exógenas*

	AM	OP	OV
CSNA3	0,003548814	0,008745876	0,008873695
OIBR4	0,004211465	0,008800468	0,008755965
PETR4	0,002825213	0,008310654	0,008693282
SANB11	0,003836082	0,008177516	0,007839651
VALE5	0,002373105	0,00739494	0,007547827
GGBR4	0,003033502	0,007377315	0,007535473
TIMP3	0,003867277	0,008036907	0,00747945
ITUB4	0,002608262	0,007004393	0,006958666
CPL6	0,003427863	0,007405181	0,006894102
FIBR3	0,003542756	0,007154832	0,0067755
BBDC4	0,002594722	0,006728224	0,006651544
SBSP3	0,003714428	0,006968218	0,006639699
EMBR3	0,003731118	0,007009767	0,006549214
CMIG4	0,003006004	0,006169636	0,006231121
BRFS3	0,003310456	0,006302612	0,006002672
CPFE3	0,003231528	0,006195274	0,005724897
VIVT4	0,002954537	0,005788055	0,005591866
UGPA3	0,002983965	0,005906725	0,005504156
PCAR4	0,002938879	0,005786827	0,005415152
ABEV3	0,002528768	0,004793972	0,004690389

*As variáveis estão descritas no texto.*

**Fonte:** Elaborada pelos autores.

Conforme se observa na Tabela 5, o período pré-abertura tem uma variação média superior ao período *after-market*, sugerindo que a mesma é responsável por maior parte da variação do período *overnight* total. Os papéis que apresentam maior variação média nos períodos não regulares foram CSNA3, OIBR4, PETR4, SANB11 e VALE5, enquanto os que apresentam menor variação média foram ABEV3, PCAR4, UGPA3, VIVT4 e CPFE3. Barclay e Hendershott (2004) indicam que a variação dos períodos não regular é importante apenas quando há movimentação suficiente para tanto. Dessa forma, acredita-se que a chegada de informação importante nos

períodos *after-market* e pré-abertura para modelagem da volatilidade ocorre principalmente nas ações que apresentaram maior variação média diária.

Ainda avaliando os modelos dentro da amostra, a Tabela 6 indica os modelos que apresentaram melhor (menor) critério de informação AICc para cada uma das categorias, sejam elas sem variável exógenas, incorporando o período *after-market* (AM), incorporando o período pré-abertura (OP) e incorporando o período *overnight* total (OV). Em destaque estão os modelos que apresentaram melhores critérios dentre todos.

Tabela 6

Melhores critérios de informação de Akaike corrigido (AICc) para os modelos de cada ação

AICc			AICc		
ABEV3	APARCH (1,1)	-5,71159	BBDC4	APARCH (1,1)	-5,41311
	APARCH (1,1) + AM	-5,70970		APARCH (1,2) + AM	-5,42162
	<b>APARCH (2,2) + OP</b>	<b>-5,72134</b>		APARCH (1,2) + OP	-5,42163
	APARCH (2,2) + OV	-5,72012		<b>APARCH (1,2) + OV</b>	<b>-5,42163</b>
BRFS3	<b>APARCH (1,1)</b>	<b>-5,37863</b>	CMIG4	APARCH (1,1)	-5,29872
	APARCH (1,1) + AM	-5,37741		APARCH (1,1) + AM	-5,29874
	APARCH (1,1) + OP	-5,37741		APARCH (1,2) + OP	-5,32857
	APARCH (1,1) + OV	-5,37741		<b>APARCH (2,2) + OV</b>	<b>-5,33154</b>
CPFE3	<b>APARCH (1,1)</b>	<b>-5,66059</b>	CPLE6	APARCH (1,1)	-5,32234
	APARCH (1,2) + AM	-5,65720		APARCH (1,1) + AM	-5,32034
	APARCH (2,2) + OP	-5,65860		APARCH (1,1) + OP	-5,32281
	APARCH (2,2) + OV	-5,65969		<b>APARCH (1,1) + OV</b>	<b>-5,33485</b>
CSNA3	APARCH (1,1)	-4,74253	EMBR3	APARCH (1,1)	-5,08514
	APARCH (1,1) + AM	-4,74532		APARCH (1,2) + AM	-5,08201
	APARCH (2,1) + OP	-4,75104		APARCH (1,1) + OP	-5,09151
	<b>APARCH (1,1) + OV</b>	<b>-4,75197</b>		<b>APARCH (1,1) + OV</b>	<b>-5,09456</b>
FIBR3	<b>APARCH (1,1)</b>	<b>-4,67232</b>	GGBR4	APARCH (1,1)	-4,91230
	APARCH (1,1) + AM	-4,67038		APARCH (2,2) + AM	-4,91280
	APARCH (1,1) + OP	-4,67037		<b>APARCH (1,2) + OP</b>	<b>-4,93230</b>
	APARCH (1,1) + OV	-4,67032		APARCH (1,2) + OV	-4,93212
ITUB4	APARCH (1,1)	-5,31292	OIBR4	APARCH (1,2)	-4,58497
	<b>APARCH (1,2) + AM</b>	<b>-5,32008</b>		APARCH (1,2) + AM	-4,58314
	APARCH (1,2) + OP	-5,31190		APARCH (1,2) + OP	-4,60736
	APARCH (2,2) + OV	-5,30970		<b>APARCH (1,2) + OV</b>	<b>-4,61311</b>
PCAR4	<b>APARCH (1,1)</b>	<b>-5,34712</b>	PETR4	APARCH (2,2)	-5,06764
	APARCH (1,1) + AM	-5,34530		APARCH (2,2) + AM	-5,06565
	APARCH (1,1) + OP	-5,34532		APARCH (2,1) + OP	-5,09839
	APARCH (1,1) + OV	-5,34533		<b>APARCH (2,1) + OV</b>	<b>-5,10900</b>
SANB11	<b>APARCH (1,1)</b>	<b>-5,10457</b>	SBSP3	APARCH (1,1)	-5,11231
	APARCH (1,1) + AM	-5,10262		APARCH (1,2) + AM	-5,11720
	APARCH (1,1) + OP	-5,10255		APARCH (1,2) + OP	-5,11810
	APARCH (1,1) + OV	-5,10285		<b>APARCH (1,2) + OV</b>	<b>-5,12030</b>
TIMP3	APARCH (1,2)	-4,98892	UGPA3	APARCH (2,2)	-5,72721
	APARCH (2,2) + AM	-4,98592		<b>APARCH (1,2) + AM</b>	<b>-5,74570</b>
	<b>APARCH (2,2) + OP</b>	<b>-5,00678</b>		APARCH (1,2) + OP	-5,73839
	APARCH (2,2) + OV	-5,00448		APARCH (1,2) + OV	-5,74565
VALE5	APARCH (1,1)	-5,32292	VIVT4	APARCH (2,2)	-5,61880
	APARCH (1,2) + AM	-5,32486		APARCH (1,2) + AM	-5,62073
	APARCH (1,1) + OP	-5,33919		<b>APARCH (1,2) + OP</b>	<b>-5,62728</b>
	<b>APARCH (2,2) + OV</b>	<b>-5,34683</b>		APARCH (1,2) + OV	-5,62076

As variáveis estão descritas no texto.

Valores em negrito indicam os modelos que apresentaram melhor AICc dentre todos para aquela ação.

Fonte: Elaborada pelos autores.

Os resultados indicam que, para o critério de informação AICc, o modelo sem as variáveis exógenas é o melhor apenas para as ações BRFS3, CPFE3, FIBR3, PCAR4 e SANB11. Vale ressaltar que, dessas cinco empresas, apenas SANB11 apresenta maior variação média diária nos períodos não regulares do pregão. Tal fato pode ser devido a essas empresas não apresentarem resultados superiores quando incorporadas as variáveis exógenas.

Já para todas as outras ações, pelo menos um dos modelos que incorporam as variáveis exógenas apresentou resultados superiores para a análise *in-sample*. Para os papéis ITUB4 e UGPA3, o modelo que

apresentou melhor resultado dentro da amostra foi o que incorpora o período *after-market*, indicando que, para essas ações, as negociações que ocorrem nesse período são importantes para estimação da volatilidade condicional. Para ABEV3, GGBR4, TIMP 3 e VIVT4, a variação entre o preço de abertura e o fechamento do *after-market*, ou seja, as informações incorporadas no período pré-abertura, se mostra a variável exógena mais importante, já que o modelo OP apresentou melhor resultado em relação aos demais. Já para as demais ações, BBDC4, CMIG4, CPLE6, CSNA3, EMBR3, OIBR4, PETR4, SBSP3 e VALE5, o modelo que apresentou

melhor resultado foi o OV, ou seja, o que incorpora o período total *overnight* (variação do preço de abertura em relação ao preço de fechamento do pregão regular do dia anterior).

Constata-se, portanto, que, ao se avaliar o critério de informação AICc (análise *in-sample*), as variáveis exógenas proporcionaram modelos mais bem ajustados para a maior parte das ações. Destaca-se, também, que o período pré-abertura (OP) e *overnight* total (OV) aparentam incorporar mais informações que o período *after-market*, o que pode ser justificado pela menor variação do *after* em relação aos demais, como visto na Tabela 6, e corroborando os resultados encontrados por Chen et al. (2012).

Um ponto de ressalva é que a diferença nos valores dos critérios para os diferentes modelos é pequena para

boa parte das ações. Tal fato é indício de que não existe unanimidade sobre o impacto das variáveis exógenas. Portanto, é necessária a utilização de outros critérios para complementar as análises da pesquisa.

A fim de sumarizar os resultados encontrados nessa pesquisa, a Tabela 7 apresenta os principais indicadores tanto para estimação do modelo dentro da amostra (*in-sample*) quanto fora da amostra (*out-of-sample*). Para análise *in-sample*, destacam-se os modelos que, incorporando as variáveis exógenas, apresentaram resultados superiores para o critério AICc e os modelos em que os coeficientes das variáveis exógenas foram estatisticamente significativos. Já os critérios para a análise *out-of-sample* indicam os modelos que apresentaram resultados superiores pelo critério R<sup>2</sup> da regressão de MZ e pelos critérios dos erros de estimação.

**Tabela 7**

*Síntese dos resultados encontrados na pesquisa*

Ação	In-sample AICc			In-sample Coef.			Out-of-sample R <sup>2</sup> MZ			Out-of-sample erros			Out-of-sample teste DM		
	AM	OP	OV	AM	OP	OV	AM	OP	OV	AM	OP	OV	AM	OP	OV
ABEV4		**	*	*	***	**	*	**	*	**	***		+	+	
BBDC4	*	*	**												-
BRFS3							**	*	*						
CMIG4	*	*	**		***	***	*	**		***			+	-	
CPFE3						*				*					
CPL6		*	**		***	***	*	*	**	**	**	**			
CSNA3	*	*	**		*		*		**	**	**	***	+	+	+
EMBR3		*	**					*	**	*	*	*			
FIBR3							**	*		*					-
GGBR4	*	**	*			**		**		**	***	**		+	
ITUB4	**									*					-
OIBR4		*	**		***	***	*	**	*		*		-	-	-
PCAR4															
PETR4		*	**		***	***	**				*	*	-		
SANB11						***	**			*				-	
SBSP3	*	*	**	***	***	***	*	*	**	**	**	**	+	+	+
TIMP3		**	*		**			**		**	**				
UGPA3	**	*	*				*	*	**	**	**	**		+	
VALE5	*	*	**		***	**	**				*		-		-
VIVT4	*	**	*		***			**	*	***					
Total	9	14	14	2	10	10	12	12	9	14	11	7	4	5	2
O melhor	2	4	9				5	6	5						

**Notas:** Para critério de informação *in-sample* de Akaike corrigido (AICc): \* = modelos em que a variável exógena foi superior ao modelo tradicional; \*\* = modelo que apresentou melhor resultado dentre todos.

Para *in-sample* coef.: \*, \*\*, \*\*\* = significância estatística do coeficiente da variável exógena ao nível de 10, 5 e 1%, respectivamente.

Para critério *out-of-sample* R<sup>2</sup> de Mincer-Zarnowitz (MZ): \* = modelos em que a variável exógena apresentou R<sup>2</sup> superior ao modelo tradicional; \*\* = modelo que apresentou melhor resultado dentre todos.

Para critério *out-of-sample* erros: \* = modelos em que pelo menos um dos dois critérios [raiz dos erros quadráticos médios (root mean squared error – RMSE) e erro percentual médio absoluto (mean absolut percentage error – MAPE)] apresentou resultado superior ao tradicional; \*\* = modelos em que os dois critérios foram superiores ao modelo tradicional; \*\*\* = modelos que apresentaram melhores resultados para os dois critérios dentre todos os modelos.

Para teste de Diebold Mariano (DM): + = modelos estatisticamente significativos e superiores; - = modelos estatisticamente significativos e inferiores (nível de 5% de significância).

As variáveis estão descritas no texto.

**Fonte:** Elaborada pelos autores.

A Tabela 7 apresenta, de forma resumida, os principais resultados desta pesquisa. Destaca-se o fato de, para todos os critérios, os modelos que incorporam o período não regular do pregão apresentaram resultados interessantes. Os critérios utilizados para avaliação dentro da amostra indicam que os períodos pré-abertura e *overnight* total apresentam resultados superiores aos modelos que incorporam o período *after-market*, já que o coeficiente da variável exógena foi significativo apenas para duas ações para esse período e para 10 ações para os outros dois. Além disso, o critério AICc apresentou melhor resultado para os dois períodos, principalmente para o período *overnight* total, em que nove vezes apresentou melhor resultado em relação aos demais modelos.

A análise fora da amostra indica números gerais semelhantes para os modelos incorporando as variáveis exógenas. Para o critério  $R^2$ , os modelos que incorporaram o período AM e OP apresentaram resultados superiores ao tradicional em 12 vezes, enquanto os modelos que incorporaram a variável OV apresentaram resultados superiores em nove vezes. Além disso, cada um deles foi o melhor dentre todos, cinco, seis e cinco vezes, respectivamente. O último critério que avalia os erros de estimação (o quanto o valor estimado se distancia dos valores realizados) indicou que a incorporação do período AM melhorou os resultados em 14 vezes para pelo menos um dos critérios, a incorporação do período OP em 11 vezes e a incorporação do período OV em sete vezes. Tal resultado se contrapõe, em certo nível, aos resultados da análise dentro da amostra. Entretanto, vale destacar que o impacto dos resultados superiores ao incorporar o período AM, que foi menor do que o impacto ao incorporar o período OP, indica a maior relevância deste último em relação ao primeiro.

Por fim, com o objetivo de realizar uma análise ainda mais rigorosa, o teste de Diebold Mariano indica os modelos alternativos que apresentaram erros de previsão

estatisticamente diferentes dos modelos tradicionais. O teste indica que quatro ações (ABEV3, CMIG4, CSNA3 e SBSP3), para os modelos que incorporaram a variável AM, tiveram resultados superiores aos tradicionais, enquanto quatro tiveram resultados inferiores e 12 a hipótese nula, de que os erros de previsão são iguais; portanto, ela não é rejeitada. Para os modelos que incorporaram a variável OP, cinco ações (ABEV3, CSNA3, GGBR4, SBSP3 e UGPA3) tiveram resultados superiores, enquanto três tiveram resultados inferiores e 12 a hipótese nula, de que os erros de previsão são iguais; sendo assim, ela não é rejeitada. Já para os modelos que incorporaram a variável OV, apenas dois modelos tiveram resultados superiores (CSNA3 e SBSP3), enquanto quatro tiveram resultados negativos e outras 14 a hipótese nula, de que os erros de previsão são iguais; dessa forma ela, não é rejeitada.

Os resultados encontrados neste estudo não permitem concluir sobre um único modelo para todas as empresas, mas permitem afirmar que os períodos não regulares do pregão incorporam informações relevantes aos modelos de estimação da volatilidade condicional para parte das ações, corroborando os resultados de Gallo e Pacini (1998) e Taylor (2007) para o mercado internacional. Os estudos do mercado brasileiro de Souza (2004) e Accioly e Mendes (2015) não chegaram a uma conclusão única para todas as ações, porém, ambos indicam para a significância do período *overnight* na modelagem da volatilidade condicional para a maioria dos casos, assim como esta pesquisa.

Além disso, os resultados são semelhantes aos encontrados por Chen et al. (2012), em que avaliam as 30 ações mais líquidas da Nasdaq e concluem que o período não regular do pregão incorpora informações relevantes aos modelos de volatilidade condicional para a maior parte das empresas, mas não para todas. Ademais, os autores também evidenciam que o período pré-abertura incorpora mais informações do que os demais ao ser incorporado aos modelos da família GARCH.

## 5. CONSIDERAÇÕES FINAIS

Partindo da premissa de que informações relevantes para previsão da volatilidade chegam ao mercado em horários não regulares do pregão, esta pesquisa avaliou como os períodos *after-market* e pré-abertura impactam sobre a estimação da volatilidade condicional de empresas brasileiras listadas na BM&FBOVESPA e pertencentes ao índice BR Titans 20.

Por meio de uma análise ainda não realizada no Brasil, avaliou-se o período não regular do pregão subdividido em dois, de forma a identificar informações relevantes

para estimação da volatilidade. Além disso, utilizou-se a medida de volatilidade realizada (percebida) como *proxy* para volatilidade de um dia, para comparação entre valores estimados e realizados. Tal fato ainda é pouco explorado no Brasil, devido à necessidade de dados em alta frequência e à dificuldade de se trabalhar com eles.

Tanto na análise dentro da amostra (*in-sample*) quanto fora da amostra (*out-of-sample*), não há unanimidade com relação ao melhor modelo. Para a grande maioria das ações, pelos diferentes critérios, os modelos que incorporaram

as variáveis exógenas superaram o modelo tradicional, sinalizando que, durante o período não regular de pregão, há a chegada de informações relevantes para previsão da volatilidade de um dia à frente. Apenas para duas ações o modelo tradicional, sem incorporar as variáveis exógenas, foi superior nos três critérios de avaliação utilizados. Além deles, em outras três ações, apesar de os modelos incorporando os períodos não regulares do pregão apresentarem resultados superiores em alguns critérios, a melhora foi muito baixa ou insignificante.

Os resultados encontrados neste estudo não são conclusivos para um único modelo superior aos demais, mas nos permitem afirmar que os períodos não regulares do pregão incorporam informações relevantes aos modelos de estimação da volatilidade condicional para boa parte das ações, corroborando os estudos internacionais de Gallo e Pacini (1998) e Taylor (2007) e os estudos do mercado brasileiro, como os de Souza (2004) e Accioly e Mendes (2015). Ademais, os resultados indicam que o período pré-abertura tem maior impacto sobre o período não regular como um todo (*overnight* total), sendo mais significativo para a modelagem da volatilidade condicional, conclusão semelhante às de Chen et al. (2012).

Ao constatar a importância do período não regular do pregão para estimação da volatilidade condicional das ações, este estudo oferece informações relevantes para que agentes de investimentos possam refinar os modelos de previsão da volatilidade e, conseqüentemente, obter melhores resultados na precificação de derivativos, na gestão de risco (cálculo do *value at risk* – VaR) e na composição e otimização de carteiras de investimentos.

Sugere-se que as causas desses resultados possam ter três origens: (i) a cointegração entre os mercados, já que durante o período não regular do pregão, no Brasil, outros mercados estão em funcionamento e podem impactar no mercado nacional; (ii) a divulgação de informações relevantes serem realizadas nesse período, conforme legislação vigente; e (iii) a possibilidade de assimetria informacional, em que negociadores com informações privilegiadas emitem ordens de compra e venda no período não regular do pregão e essas informações são absorvidas pelo mercado durante as primeiras horas do pregão regular, como também sugerido por Chen et al. (2012).

Uma limitação deste estudo é que o mercado de capitais brasileiro ainda sofre muita oscilação e, em alguns momentos, a variação do período não regular pode ser muito baixa. Para estudos futuros, sugere-se uma análise do impacto do período não regular em diferentes períodos intradiários do pregão regular, principalmente nas primeiras horas de negociação. Nesse caso, não apenas avaliando o impacto sobre a volatilidade, mas também sobre o retorno em si. Outra sugestão é a análise de outras variáveis como *proxy* para a volatilidade observada de um dia para confrontar os resultados encontrados utilizando a volatilidade realizada como *proxy*. Sugere-se, também, como forma de confirmar/confrontar os resultados encontrados nesta pesquisa, que seja realizada uma análise com modelos de volatilidade estocástica, em vez de modelos de volatilidade condicional. Ademais, uma análise qualitativa dos *market makers* (formadores de mercado) e dos operadores do período não regular do pregão pode gerar *insights* e respostas interessantes em trabalhos futuros.

## REFERÊNCIAS

- Accioly, V. B., & Mendes, B. V. M. (2015). EGARCH-RR: realized ranges explaining EGARCH volatilities. *Relatório COPPEAD*, 416, 1-23.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.
- Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4), 885-905.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61(1), 43-76.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001b). Modeling and forecasting realized volatility. *National Bureau of Economic Research*, 71(2), 579-625.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001c). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96(453), 42-55.
- Barclay, M. J., & Hendershott, T. (2003). Price discovery and trading after hours. *The Review of Financial Studies*, 16(4), 1041-1073.
- Barclay, M. J., & Hendershott, T. (2004). Liquidity externalities and adverse selection: evidence from trading after hours. *The Journal of Finance*, 59(2), 681-710.
- Bolsa de Valores, Mercadorias e Futuros de São Paulo. (2014). Manual de Procedimentos Operacionais do Segmento Bovespa: Ações, Futuros e Derivativos de Ações. Recuperado de <http://www.bmfbovespa.com.br/lumis/portal/file/fileDownload.jsp?fileId=8AA8D0976075EB9901607A6ACDE8352E>.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.
- Bollerslev, T., & Wright, J. H. (2001). High-frequency data, frequency domain inference, and volatility forecasting. *The Review of Economics and Statistics*, 83(4), 596-602.
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2), 261-304.

- Cavaleri, R., & Ribeiro, E. P. (2011). Combinação de previsões de volatilidade: um estudo. *Revista Economia*, 12(2), 239-261.
- Ceretta, P. S., & Costa, N. C. (2001). Particularidades do mercado financeiro latino-americano. *Revista de Administração de Empresas*, 41(2), 72-77.
- Chen, C. H., Yu, W. C., & Zivot, E. (2012). Predicting stock volatility using after-hours information: evidence from the Nasdaq actively traded stocks. *International Journal of Forecasting*, 28(2), 366-383.
- Comissão de Valores Mobiliários. (2002). *Instrução CVM 358, de 3 de janeiro de 2002*. Dispõe sobre a divulgação e uso de informações sobre ato ou fato relevante relativo às companhias abertas, disciplina a divulgação de informações na negociação de valores mobiliários e na aquisição de lote significativo de ações de emissão de companhia aberta, estabelece vedações e condições para a negociação de ações de companhia aberta na pendência de fato relevante não divulgado ao mercado, revoga a Instrução CVM no 31, de 8 de fevereiro de 1984, a Instrução CVM no 69, de 8 de setembro de 1987, o art. 3º da Instrução CVM nº 229, de 16 de janeiro de 1995, o parágrafo único do art. 13 da Instrução CVM 202, de 6 de dezembro de 1993, e os arts. 3o a 11 da Instrução CVM no 299, de 9 de fevereiro de 1999, e dá outras providências. Recuperado de <http://www.cvm.gov.br/export/sites/cvm/legislacao/instrucoes/anexos/300/inst358.pdf>.
- Cumby, R., Figlewski, S., & Hasbrouck, J. (1993). Forecasting volatilities and correlations with EGARCH models. *The Journal of Derivatives*, 1(2), 51-63.
- Diebold, F. X., & Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3), 253-265.
- Ding, Z., Granger, C. W., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1), 83-106.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987-1007.
- Figlewski, S. (1997). Forecasting volatility. *Financial Markets, Institutions and Instruments*, 6(1), 1-88.
- Gaio, L. E., Pessanha, G. R. G., Oliveira, D. R. de, & de Ázara, L. N. (2007). Análise da volatilidade do índice Bovespa: um estudo empírico utilizando modelos da classe ARCH. *Contextus*, 5(1), 7-16.
- Galdi, F. C., & Pereira, L. M. (2007). Valor em Risco (VaR) utilizando modelos de previsão de volatilidade: EWMA, GARCH e Volatilidade Estocástica. *Brazilian Business Review*, 4(1), 74-95.
- Gallo, G. M., & Pacini, B. (1998). Early news is good news: the effects of market opening on market volatility. *Studies in Nonlinear Dynamics & Econometrics*, 2(4), 115-131.
- Giot, P., & Laurent, S. (2004). Modelling daily value-at-risk using realized volatility and ARCH type models. *Journal of Empirical Finance*, 11(3), 379-398.
- Hurvich, C. M., & Tsai, C. L. (1993). A corrected Akaike information criterion for vector autoregressive model selection. *Journal of Time Series Analysis*, 14(3), 271-279.
- Jorion, P. (1995). Predicting volatility in the foreign exchange market. *The Journal of Finance*, 50(2), 507-528.
- Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297-303.
- Martens, M. (2002). Measuring and forecasting S&P 500 index-futures volatility using high-frequency data. *Journal of Futures Markets*, 22(6), 497-518.
- Milach, F. T. (2010). *Estimação da volatilidade: uma aplicação utilizando dados intradiários* (Dissertação de Mestrado). Escola de Administração, Universidade Federal do Rio Grande do Sul, Porto Alegre.
- Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In: *Economic forecasts and expectations: analysis of forecasting behavior and performance*. Cambridge, MA: National Bureau of Economic Research.
- Moreira, J. M. D. S., & Lemgruber, E. F. (2004). O uso de dados de alta frequência na estimação da volatilidade e do valor em risco para o IBOVESPA. *Revista Brasileira de Economia*, 58(1), 100-120.
- Mota, B. D. S., & Fernandes, M. (2004). Desempenho de estimadores de volatilidade na bolsa de valores de São Paulo. *Revista Brasileira de Economia*, 58(3), 429-448.
- Ng, H. S., & Lam, K. P. (2006). *How does the sample size affect GARCH model?* [Working Paper]. Department of Systems Engineering and Engineering Management, Chinese University of Hong Kong.
- Nicolau, J. (2012). *Modelação de séries temporais financeiras*. Coimbra: Almedina. Coleção Econômicas II, n. 18.
- Oomen, R. C. A. (2001). *Using high frequency stock market index data to calculate, model & forecast realized return variance* [Discussion Paper]. European University, Economics.
- Reis, D. L. P. E. (2011). *Análise de desempenho de indicadores de volatilidade* (Dissertação de Mestrado). Faculdade de Economia e Administração, Universidade Federal de Juiz de Fora, Juiz de Fora.
- Silva, C. A. G. (2009). Modelagem de estimação da volatilidade do retorno das ações brasileiras: os casos da Petrobrás e Vale. *Cadernos do IME-Série Estatística*, 26(1), 5-28.
- Silva, J. C. (2002). *Estimação do valor em risco usando informação intradiária* (Dissertação de Mestrado). Escola de Administração de Empresas, Fundação Getúlio Vargas, Rio de Janeiro.
- Souza, M. N. C. (2004). *Redução da persistência de volatilidade nos modelos GARCH para cálculo do valor em risco no mercado brasileiro* (Dissertação de Mestrado). Instituto Coppead de Administração, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- Taylor, N. (2007). A note on the importance of overnight information in risk management models. *Journal of Banking & Finance*, 31(1), 161-180.
- Tsay, R. S. (2010). *Analysis of financial time series*. New Jersey: Wiley. Series in Probability and Statistics.
- Zivot, E. (2009). Practical issues in the analysis of univariate GARCH models. In: *Handbook of financial time series*. Heidelberg: Springer.