

## **Modelagem da Volatilidade Condicional Incorporando o Período Não Regular do Pregão ao Modelo APARCH: Um Estudo com Ações Listadas na BM&FBovespa**

### **Autoria**

Breno Valente Fontes Araujo - brenovfa@gmail.com

Centro de Pós-Grad e Pesquisas em Admin – CEPEAD/UFMG - Universidade Federal de Minas Gerais

Frank Magalhães de Pinho - fpinho@ibmec.edu.br

Centro de Pós-Grad e Pesquisas em Admin – CEPEAD/UFMG - Universidade Federal de Minas Gerais

Marcos Antônio de Camargos - marcosac@face.ufmg.br

Centro de Pós-Grad e Pesquisas em Admin – CEPEAD/UFMG - Universidade Federal de Minas Gerais

### **Resumo**

A volatilidade tem bastante destaque nos estudos de Finanças, pois é um parâmetro fundamental na precificação de derivativos, gestão de risco e de portfólios. Acreditando que durante o período não regular do pregão ocorra a chegada de informações importantes capazes de impactar na volatilidade do dia, o objetivo deste artigo é avaliar como os períodos after-market e pré-abertura impactam na estimação da volatilidade condicional de um dia à frente. Para isso, utilizou-se o modelo APARCH (Asymmetric Power Autoregressive Conditional Heteroskedasticity) incorporando o período after-market, a pré-abertura e o overnight total, para avaliar se os mesmos carregam informações importantes para modelagem da volatilidade. Foram analisados dados intradiários de 20 ações de empresas brasileiras listadas na BM&FBovespa, pertencentes ao índice BR TITANS 20 com ADRs listados nas bolsas de Nova York e na NASDAQ, considerando o período de janeiro de 2010 a julho de 2015. Os critérios utilizados para avaliar os resultados foram: de informação AICc e significância estatística dos coeficientes (in-sample) e RMSE, MAPE e  $R^2$  da regressão de Mincer Zarnowitz (out-of-sample). A análise dos resultados dentro e fora da amostra não permite afirmar o melhor modelo, pois não há unanimidade nas ações, entretanto, em ambas as análises, os períodos não regulares do pregão demonstraram incorporar informações importantes para a maior parte das ações. Ademais, os modelos que incorporaram o período pré-abertura obtiveram, em geral, resultados superiores aos modelos que incorporaram o período after-market, sinalizando que tal período carrega informações importantes para a previsão da volatilidade condicional.

## Modelagem da Volatilidade Condicional Incorporando o Período Não Regular do Pregão ao Modelo APARCH: Um Estudo com Ações Listadas na BM&FBovespa

### Resumo:

A volatilidade tem bastante destaque nos estudos de Finanças, pois é um parâmetro fundamental na precificação de derivativos, gestão de risco e de portfólios. Acreditando que durante o período não regular do pregão ocorra a chegada de informações importantes capazes de impactar na volatilidade do dia, o objetivo deste artigo é avaliar como os períodos *after-market* e pré-abertura impactam na estimação da volatilidade condicional de um dia à frente. Para isso, utilizou-se o modelo *APARCH* (*Asymmetric Power Autoregressive Conditional Heteroskedasticity*) incorporando o período *after-market*, a pré-abertura e o *overnight* total, para avaliar se os mesmos carregam informações importantes para modelagem da volatilidade. Foram analisados dados intradiários de 20 ações de empresas brasileiras listadas na BM&FBovespa, pertencentes ao índice BR TITANS 20 com *ADRs* listados nas bolsas de Nova York e na NASDAQ, considerando o período de janeiro de 2010 a julho de 2015. Os critérios utilizados para avaliar os resultados foram: de informação *AICc* e significância estatística dos coeficientes (*in-sample*) e *RMSE*, *MAPE* e  $R^2$  da regressão de Mincer Zarnowitz (*out-of-sample*). A análise dos resultados dentro e fora da amostra não permite afirmar o melhor modelo, pois não há unanimidade nas ações, entretanto, em ambas as análises, os períodos não regulares do pregão demonstraram incorporar informações importantes para a maior parte das ações. Ademais, os modelos que incorporaram o período pré-abertura obtiveram, em geral, resultados superiores aos modelos que incorporaram o período *after-market*, sinalizando que tal período carrega informações importantes para a previsão da volatilidade condicional.

**Palavras-chave:** Volatilidade Condicional, Dados Intradiários, Modelo *APARCH*, *After-Market*, Pré-Abertura.

### 1. INTRODUÇÃO

Diversos estudos no Brasil e no exterior utilizam dados diários para a previsão da volatilidade condicional do dia seguinte. No Brasil, a maior parte desses estudos se desenvolveu nos últimos vinte anos, apontando para um bom desempenho dos modelos da família *ARCH* (*Autoregressive Conditional Heteroskedasticity*) para previsão da volatilidade (CERETTA e COSTA JR., 2001; GAIO *et al.*, 2007; GALDI e PEREIRA, 2007; SILVA, 2009; MELLO, 2009; CAVALERI; RIBEIRO, 2011). 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* (SOUZA, 2004; ACCIOLY; MENDES, 2015), enquanto no exterior foram encontrados diversos estudos que debatem o tema (GALLO; PACINI, 1997; MARTENS, 2002; BARCLAY; HENDERSHOTT, 2004; TAYLOR, 2007; CHEN; YU; ZIVOT, 2012).

A Instrução Normativa 358/2002 da Comissão de Valores Mobiliários (CVM) 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. Neste 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 (2008) 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. Desta 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 BM&FBovespa e pertencentes ao *Dow Jones Brazil Titans ADR Index BR TITANS 20* (BR TITANS 20).

O presente trabalho se diferencia dos demais principalmente ao fragmentar o retorno *overnight* em subperíodos, assim como feito por Chen, Yu e Zivot (2012). Porém, devido às peculiaridades do mercado brasileiro, o período *overnight* foi dividido em dois subperíodos, ao invés de três: o *after-market* e o pré-abertura. Para isto, foi utilizado 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, contribuir com a literatura de três maneiras: 1. apresentando a análise do período *overnight* ainda pouco estudada no Brasil; 2. utiliza-se subperíodos do horário não regular do pregão como variáveis explicativas para modelar a volatilidade condicional; 3. ao realizar uma 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. 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). Neste sentido, Engle (1982) propôs o modelo *ARCH* em que a variância condicional pode ser

modelada por meio de uma função quadrática. Matematicamente, para um modelo *ARCH* de ordem  $m$  tem-se que a equação da volatilidade será dada pela equação abaixo:

$$\epsilon_t = \sigma_t u_t; e \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2; \quad (1)$$

em que  $u_t$  é uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) com média 0 e variância 1,  $\omega > 0$  e  $\alpha_i \geq 0$  para  $i > 0$  e  $t = 1, \dots, m$ . A restrição dos coeficientes do modelo deve ser satisfeita a fim de se obter uma variância condicional positiva. Além disso, salienta-se que a distribuição de  $\epsilon_t$  segue, comumente na prática, a distribuição normal padronizada, ou a *t-Student* ou ainda a dos erros generalizados - GED (TSAY, 2010). 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 esta 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*. A equação de volatilidade para um modelo *GARCH* tradicional de ordem  $p$  e  $q$  pode ser expresso pela equação 2:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2; \quad (2)$$

em que  $\epsilon_t$  apresenta as mesmas propriedades daquele explicitado formalmente no modelo *ARCH*,  $\omega > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$ , e  $\sum_{i=1}^{\max(l,m)} (\alpha_i + \beta_j) < 1$ . Como no caso dos modelos *ARCH*,  $\epsilon_t$  na prática é usualmente assumido seguindo uma distribuição normal padronizada ou *t-Student* ou dos erros generalizados (GED).

Ao longo dos anos, diversos outros 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 deste ser capaz de se comportar como outros sete modelos da família *ARCH*.

O modelo *Asymetric Power Autorregressive 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 3:

$$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i (|\epsilon_{t-i}| - \gamma_i \epsilon_{t-i})^\delta + \sum_{j=1}^p \beta_j \sigma_{t-j}^\delta; \quad (3)$$

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, representa 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 tem impacto maior sobre a variância condicional de um dia à frente. Caso  $\gamma_i$  seja estatisticamente significativo e negativo, indica que choques positivos tem 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 da família *ARCH*, 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 ( $\gamma_i = 0$  e  $\beta_j = 0$ ); *GARCH* de Taylor/Schwartz's ( $\delta = 1, \gamma_i = 0$ ); *TARCH* ( $\delta = 1$  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 deste estudo não ter como objetivo avaliar os diversos modelos, mas 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 dela 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), Figlewski (1997) e Jorion (1995) constataram que os modelos *GARCH* apresentam resultados ruins para estimação da volatilidade, não por deficiência do modelo, mas sim, 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 a volatilidade igual a zero.

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. Este trabalho utilizou 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 *et al.* (2001a), Andersen *et al.* (2001b), e Bollerslev e Wright (2001):

$$r_{t,i} = P_{t,i} - P_{t,i-1} \quad (4)$$

$$VR_t^2 = \sum_{i=1}^n r_{t,i}^2 \quad (5)$$

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 *et al.* (2001c) indicaram que quanto maior a frequência dos dados intradiários, mais perto a volatilidade realizada se aproxima da volatilidade integrada, que pode ser considerada como 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 se encontra entre 5 e 25 minutos (MOTA; FERNANDES, 2004). Oomen (2001) apresenta como frequência ótima intervalos de 25 minutos. Já Giot e Laurent (2004) encontraram uma frequência ótima de 15 minutos para seu estudo. Optou-se por utilizar a frequência de 15 minutos, devido a disponibilidade dos dados e a utilização da mesma frequência em inúmeros estudos nacionais (SILVA, 2002; MOTA; FERNANDES, 2004; MOREIRA; LEMGRUBER, 2004; MILACH, 2010; REIS, 2011).

### 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: 1. pregão regular: que ocorre entre as 10h e às 17h; e 2. *after-market*: que ocorre entre as 17h30min e 18h, sendo a única forma de negociação após o fechamento do pregão, representando, devido a isso, como 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 BM&FBovespa

Mercado	Cancelamento de Ofertas		Pré-Abertura		Negociação		Call de Fechamento		After-Market			
	Início	Fim	Início	Fim	Início	Fim	Início	Fim	Cancelamento de Ofertas	Negociação		
	Início	Fim	Início	Fim	Início	Fim	Início	Fim	Início	Fim	Início	Fim
Mercado a vista	09:30	09:45	09:45	10:00	10:00	16:55	16:55	17:00	17:25	17:30	17:30	18:00

FONTE - Adaptada da BM&FBovespa (2016).

O limite estipulado de variação no *after* de uma ação de 2%, seja para mais ou para menos, restringe a análise do trabalho, pois pode ocorrer dias em que a variação do *after-market* chegue a 2% e não passa por impossibilidade, não representando a real situação das ordens de compra e venda. Para contornar tal situação, foi verificado o número de vezes em que a variação do período *after-market* se aproximou de 2%. Foi constatado 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 deste período impactam no comportamento do mercado no pregão regular e sobre a volatilidade de ações e índices, conforme se observa na figura 1.

Figura 1 - Evidências Empíricas Analisando o Período *Overnight* na Literatura Nacional e Internacional

Autor(es) Ano	Objetivo	Conclusão
Accioly e Mendes (2015)	Avaliar a inserção da volatilidade realizada como variável exógena ao modelo <i>GARCH</i> 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, Yu e Zivot (2012)	Avaliar o modelo <i>GARCH</i> 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	As informações do <i>overnight</i> têm impacto significativo sobre a volatilidade condicional dos

	agentes que trabalham com gestão de risco.	ativos analisados, promovendo modelos mais precisos para gestão do risco.
Souza (2004)	Analisar se a incorporação do efeito <i>overnight</i> ao modelo <i>GARCH</i> , leva à redução na persistência de volatilidade com dados diários das oito ações mais líquidas da BM&FBovespa.	Foi detectado o efeito redução na persistência da volatilidade para estas 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 <i>GARCH</i> , ao incluir diversas formas da volatilidade <i>after-hours</i> , melhora as previsões da volatilidade.	Concluíram que essa inclusão não apresenta uma melhora significante para o modelo.
Gallo e Pacini (1997)	Avaliar se as variações entre o preço de abertura de um dia e o preço de fechamento do dia anterior tem poder explicativo sobre a volatilidade condicional de diferentes índices.	Destacaram que ao prever a volatilidade fora da amostra ( <i>ou-of-sample</i> ), ao se adicionar essa variável, o modelo apresentou resultados superiores ao modelo <i>GARCH</i> tradicional.

FONTE – Elaborado pelos autores.

Além das evidências apresentadas acima, 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 constitui das empresas listadas na BM&FBovespa e pertencentes ao índice BR TITANS 20, índice calculado desde de 2004 a partir dos *ADRs* de empresas brasileiras mais negociados na bolsa de Nova York e *NASDAQ*, apresentadas na figura 2.

**Figura 2 - Companhias Brasileiras Pertencentes ao Índice BR TITANS 20**

Nome da companhia	Código	Nome da companhia	Código	Nome da companhia	Código
Ambev S/A	ABEV3	COPEL	CPL6	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 Part.	TIMP3
Pão-de-Açúcar	PCAR4	Fibria Celulose S/A	FIBR3	Ultrapar Part. S/A	UGPA3
SABESP	SBSP3	Gerdau S/A	GGBR4	Vale S/A	VALE5
CEMIG	CMIG4	Itaú-Unibanco S/A	ITUB4		

FONTE - Elaborado 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. Eles estão ajustados para proventos e eventos (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 quinze minutos para o pregão regular e *after-market*.

O período de análise foi do dia primeiro 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 deste 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, Yu e Zivot (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 valores 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.

Abaixo, 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 *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 figura 3.

**Figura 3 - Modelos APARCH Avaliados**

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

FONTE - Elaborado pelos autores.

Com o objetivo de refinar os modelos e recomendado por Tsay (2010), foi estimado 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 foi realizado o teste de Ljung-Box 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 i.i.d. (independentes e identicamente distribuídos). Feito isso, foi realizado o teste *ARCH* para comprovação da existência de heterocedasticidade. A partir deste ponto, estimou-se os modelos da volatilidade condicional por meio do modelo *APARCH*.

Inicialmente, foi estimado 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, foram ajustados 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-student* assimétrica e a generalizada (GED) assimétrica. A não avaliação da distribuição normal se deve 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 para uma melhor adequação para essas distribuições. A distribuição assimétrica *t-student* foi a mais apropriada para todas as séries de log-retornos, menos para o papel da Vale, em que a distribuição assimétrica GED se mostrou 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 heterocedasticidade 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  $\tilde{\varepsilon}_t$ , ou resíduos padronizados, é feito da seguinte forma:

$$\tilde{\varepsilon}_t = \frac{\varepsilon_t}{\sigma_t} \quad (6)$$

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

$$AIC = -2L + 2p \quad (7)$$

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*, e pode ser calculada a partir da transformação abaixo:

$$AICc = AIC + \frac{2p(p+1)}{n-p-1} \quad (8)$$

na qual *n* seria o tamanho amostral do conjunto de dados utilizados na estimação do modelo. O *AICc* penaliza 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 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 (*RMSE – Root Mean Squared Error*) e o Erro Percentual Médio Absoluto (*MAPE – Mean Absolut Percentage Error*). A ideia da regressão de Mincer-Zarnowitz é simples, basta regressar a volatilidade realizada (observável)

em função da volatilidade condicional estimada pelos modelos. A regressão é formalmente descrita da seguinte maneira:

$$(VR_{t+k}^2)^{1/2} = \alpha_0 + \alpha_1(\sigma_{t+k|t}^2)^{1/2} + u_{t+k} \quad (9)$$

em que  $VR_{t+k}^2$  é a variância realizada do dia t+k, e  $\sigma_{t+k|t}^2$  se refere a 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*, neste 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. Ela pode ser calculada:

$$RMSE = \left[ \frac{1}{T} \sum_{t=1}^T (VR_t^2 - \sigma_t^2)^2 \right]^{1/2} \quad (10)$$

Por fim, avaliou-se também pelo método do *MAPE*, que tem como uma das vantagens a fácil interpretação, já que a escala é em porcentagem, e tem como principal desvantagem o fato de que se o valor realizado for muito pequeno ou zero, o valor do *MAPE* “explode” ou não é possível se calcular.

$$MAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{VR_t - \sigma_t}{VR_t} \right| \cdot 100 \quad (11)$$

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 ( $\mu$ ,  $AR_1$ ,  $MA_1$ ,  $\alpha$  e  $\beta$  - sem e com defasagens,  $\gamma$  Vexog, Skew e Shape) foi analisado, tendo como referência significância de 5% (p-valor > 0,05).

Novamente, devido ao limite de páginas, optou-se por fazer uma análise geral de todas as ações, ordenadas pelas que tiveram maiores e menores variações nos períodos não regulares do pregão. Nesse sentido, foi calculada a média diária de variação absoluta das variáveis exógenas, ou seja, *AM*, *OP* e *OV*. A tabela 2 indica, portanto, o valor médio absoluto de variação para esses três períodos, respectivamente.

Tabela 2 - Variação Média Diária das Variáveis Exógenas

	<i>AM</i>	<i>OP</i>	<i>OV</i>
<b>CSNA3</b>	0,003548814	0,008745876	0,008873695
<b>OIBR4</b>	0,004211465	0,008800468	0,008755965
<b>PETR4</b>	0,002825213	0,008310654	0,008693282
<b>SANB11</b>	0,003836082	0,008177516	0,007839651
<b>VALE5</b>	0,002373105	0,00739494	0,007547827
<b>GGBR4</b>	0,003033502	0,007377315	0,007535473
<b>TIMP3</b>	0,003867277	0,008036907	0,00747945
<b>ITUB4</b>	0,002608262	0,007004393	0,006958666
<b>CPLE6</b>	0,003427863	0,007405181	0,006894102

<b>FIBR3</b>	0,003542756	0,007154832	0,0067755
<b>BBDC4</b>	0,002594722	0,006728224	0,006651544
<b>SBSP3</b>	0,003714428	0,006968218	0,006639699
<b>EMBR3</b>	0,003731118	0,007009767	0,006549214
<b>CMIG4</b>	0,003006004	0,006169636	0,006231121
<b>BRFS3</b>	0,003310456	0,006302612	0,006002672
<b>CPFE3</b>	0,003231528	0,006195274	0,005724897
<b>VIVT4</b>	0,002954537	0,005788055	0,005591866
<b>UGPA3</b>	0,002983965	0,005906725	0,005504156
<b>PCAR4</b>	0,002938879	0,005786827	0,005415152
<b>ABEV3</b>	0,002528768	0,004793972	0,004690389

FONTE - Elaborada pelos autores.

Conforme se observa na tabela 2, 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. Desta forma, acredita-se que a chegada de informação importante no período *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 3 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 3 - Melhores Critérios de Informação AICc para os Modelos de Cada Ação**

		AICc			AICc
<b>ABEV3</b>	APARCH (1,1)	-5,71159	<b>BBDC4</b>	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>
<b>BRFS3</b>	<b>APARCH (1,1)</b>	<b>-5,37863</b>	<b>CMIG4</b>	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>
<b>CPFE3</b>	<b>APARCH (1,1)</b>	<b>-5,66059</b>	<b>CPLE6</b>	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>
<b>CSNA3</b>	APARCH (1,1)	-4,74253	<b>EMBR3</b>	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>
<b>FIBR3</b>	<b>APARCH (1,1)</b>	<b>-4,67232</b>	<b>GGBR4</b>	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
<b>ITUB4</b>	APARCH (1,1)	-5,31292	<b>OIBR4</b>	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>
<b>PCAR4</b>	<b>APARCH (1,1)</b>	<b>-5,34712</b>	<b>PETRA4</b>	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>
<b>SANB11</b>	<b>APARCH (1,1)</b>	<b>-5,10457</b>	<b>SBSP3</b>	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>
<b>TIMP3</b>	APARCH (1,2)	-4,98892	<b>UGPA3</b>	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
<b>VALE5</b>	APARCH (1,1)	-5,32292	<b>VIVT4</b>	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

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 uma maior variação média diária nos períodos não regulares do pregão. Tal fato pode ser a razão dessas empresas não apresentarem resultados superiores quando incorporadas as variáveis exógenas.

Já para todas outras ações, pelo menos um dos modelos que incorporam as variáveis exógenas apresentaram 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 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 melhores ajustados para 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 2, e corrobora os resultados encontrados por Chen, Yu e Zivot (2012).

A fim de sumarizar os resultados encontrados nessa pesquisa, a tabela 4 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*, destaca-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*, indica os modelos que apresentaram resultados superiores pelo critério  $R^2$  da regressão de MZ e pelos critérios dos erros de estimação.

Tabela 4 – 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	
	AM	OP	AM	OP	AM	OP	AM	OP
ABEV4	**	*	***	**	*	**	**	***
BBDC4	*	**				*		
BRFS3						*		
CMIG4	*	**	***	***	*	**	***	
CPFE3				*			*	
CPLE6	*	**	***	***	*	**	**	**
CSNA3	*	**	*		*	**	**	***
EMBR3	*	**			*	**	*	*
FIBR3					**	*	*	
GGBR4	*	*		**		**	**	**
ITUB4	**					*	*	
OIBR4	*	**	***	***	*	*	*	
PCAR4								
PETR4	*	**	***	***	**	**	*	*
SANB11				***			*	
SBSP3	*	**	***	***	*	**	**	**
TIMP3	**	*	**			**	**	**
UGPA3	**	*			*	**	**	**
VALE5	*	**	***	**	**	**	**	*
VIVT4	*	*	***			*	***	
<b>Total</b>	<b>9</b>	<b>14</b>	<b>2</b>	<b>10</b>	<b>10</b>	<b>12</b>	<b>14</b>	<b>11</b>
<b>O melhor</b>	<b>2</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>5</b>	<b>6</b>	<b>5</b>	<b>7</b>

FONTE - Elaborada pelos autores.

Para critério *In-sample* AICc, \* indica os modelos em que a variável exógena foi superior ao modelo tradicional, e \*\* o modelo que apresentou melhor resultado dentre todos. Para *In-sample* Coef., \*, \*\* e \*\*\* indicam a significância estatística do coeficiente da variável exógena ao nível de 10%, 5% e 1% de significância, respectivamente.

Para o critério *Out-of-sample* R<sup>2</sup> MZ, \* indica os modelos em que a variável exógena apresentou R<sup>2</sup> superior ao modelo tradicional, e \*\* o modelo que apresentou melhor resultado dentre todos.

Para *Out-of-sample* erros, \* indica os modelos em que pelo menos um dos dois critérios (RMSE e MAPE) apresentou resultados superiores ao tradicional, \*\* indica os modelos em que os dois critérios foram superiores ao modelo tradicional, e \*\*\* indica os modelos que apresentaram melhores resultados para os dois critérios dentre todos os modelos.

A tabela 4 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 dez 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 9 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 7 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 foram menores do que estes, ao incorporar os períodos *OP* e *OV*, o que indica para a maior relevância desses dois últimos em relação ao primeiro.

Os resultados encontrados nesse estudo indicam, portanto, que os períodos não regulares do pregão incorporam informações importantes aos modelos de estimação da volatilidade condicional para a maior parte das ações, corroborando com os resultados de Gallo e Pacini (1997) 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 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, Yu e Zivot (2012), em que avaliam as 30 ações mais líquidas da *NASDAQ*, e concluíram que o período não regular do pregão incorpora informações importantes 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 importantes 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 importantes 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 a dificuldade de se trabalhar com os mesmos.

Tanto na análise dentro da amostra (*in-sample*), quanto fora da amostra (*out-of-sample*), não há uma unanimidade com relação ao melhor modelo na análise *out-of-sample*.

Para 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 importantes 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, 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 permitem concluir, portanto, que os períodos não regulares do pregão incorporam informações importantes aos modelos de estimação da volatilidade condicional para a maior parte das ações, corroborando com os estudos internacionais de Gallo e Pacini (1998) e Taylor (2007), e com 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) e sendo mais significativo para a modelagem da volatilidade condicional, conclusão semelhante às de Chen, Yu e Zivot (2012).

Ao constatar a importância do período não regular do pregão para estimação da volatilidade condicional das ações, este estudo oferece informações importantes 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 VaR), na composição e otimização de carteiras de investimentos.

Sugere-se que as causas desses resultados possam ter três origens: 1. 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; 2. a divulgação de informações importantes serem realizadas nesse período, como exigido por lei; e, 3. 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, Yu e Zivot (2012).

Uma limitação desse 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 no 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, ao invés 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 para trabalhos futuros.

## REFERÊNCIAS

- ACCIOLY, V. B.; MENDES, B. V. de M. EGARCH-RR: Realized ranges explaining EGARCH volatilities, *Relatório COPPEAD*, v.416, p.1-23, 2015.
- ANDERSEN, T.G.; BOLLERSLEV, T. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review*, v.39, n.4, p.885-905, 1998.
- ANDERSEN, T. G.; BOLLERSLEV, T.; DIEBOLD, F.X.; EBENS, H. The distribution of realized stock return volatility. *Journal of Financial Economics*, v.61, n.1, p.43-76, 2001a.

- ANDERSEN, T. G.; BOLLERSLEV, T.; DIEBOLD, F.X.; LABYS, P. Modeling and forecasting realized volatility. *National Bureau of Economic Research*, v.71, p.579-625, 2001b.
- ANDERSEN, T.G., BOLLERSLEV, T., DIEBOLD F.X., LABYS, P. The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, v.96, p.42-55, 2001c.
- AKAIKE, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, v.19, n.6, p.716-723, 1974.
- BARCLAY, M. J.; HENDERSHOTT, T. Price discovery and trading after hours. *Review of Financial Studies*, v.16, n.4, p.1041-1073, 2003.
- BARCLAY, M. J.; HENDERSHOTT, T. Liquidity externalities and adverse selection: evidence from trading after hours. *The Journal of Finance*, v.59, n.2, p.681-710, 2004.
- BM&FBOVESPA. Manual de Procedimentos Operacionais da Bolsa de Valores de São Paulo, 2009. Disponível em: [http://www.bmfbovespa.com.br/Pdf/cap\\_06\\_do\\_after\\_market.pdf](http://www.bmfbovespa.com.br/Pdf/cap_06_do_after_market.pdf) Acessado em: outubro de 2016.
- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, v.31, n.3, p.307-327, 1986.
- BOLLERSLEV, T.; WRIGHT, J. H. High-frequency data, frequency domain inference, and volatility forecasting. *Review of Economics and Statistics*, v. 83, n. 4, p. 596-602, 2001.
- BURNHAM, K. P.; ANDERSON, D. R. Multimodel inference understanding AIC and BIC in model selection. *Sociological Methods & Research*, v. 33, n. 2, p. 261-304, 2004.
- CAVALERI, R.; RIBEIRO, E. P. Combinação de previsões de volatilidade: um estudo. *Revista Economia*, v.12, n.2, p.239-261, 2011.
- CERETTA, P. S.; COSTA JÚNIOR, N. C. A. Particularidades do mercado financeiro latino-americano. *RAE-Revista de Administração de Empresas*, v.41, n.2, p.72-77, 2001.
- CHEN, C. H.; YU, W. C.; ZIVOT, E. Predicting stock volatility using after-hours information: evidence from the NASDAQ actively traded stocks. *International Journal of Forecasting*, v.28, n.2, p.366-383, 2012.
- COMISSÃO DE VALORES MOBILIÁRIOS - CVM. Instrução CVM nº 358, 2002.
- CUMBY, R.; FIGLEWSKI S.; HASBROUCK J. Forecasting volatility and correlations with EGARCH models. *Journal of Derivatives*, v.1, n.2, p.51-63, 1993.
- DING, Z.; GRANGER, C.W.J.; ENGLE, R.F. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, v.1, n.1, p.83-106, 1993.
- ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, v.50, n.4, p.987-1007, 1982.
- FIGLEWSKI, S. Forecasting volatility. *Financial Markets, Institutions and Instruments*, v.6, n.1, p.1-88, 1997.
- GAIO, L. E.; PESSANHA, G. R. G.; OLIVEIRA, D. R.; ÁZARA, L. N. Análise da volatilidade do índice Bovespa: um estudo empírico utilizando modelos da classe ARCH. *Revista Contextus*, v.5, n.1, p.7-16, 2007.
- GALDI, F. C.; PEREIRA, L. M. Valor em risco (VaR) utilizando modelos de previsão de volatilidade: EWMA, GARCH e volatilidade estocástica. *Brazilian Business Review*, v.4, n.1, p.74-95, 2007.
- GALLO, G. M.; PACINI, B. Early news is good news: the effects of market opening on market volatility. *Studies in Nonlinear Dynamics & Econometrics*, v.2, n.4, p.115-131, 1998.
- GIOT, P.; LAURENT, S. Modelling daily value-at-risk using realized volatility and ARCH type models. *Journal of Empirical Finance*, v.11, n.3, p.379-398, 2004.
- HURVICH, C. M.; TSAI, C. L. A corrected Akaike information criterion for vector autoregressive model selection. *Journal of Time Series Analysis*, v.14, n.3, p.271-279, 1993.

- JORION, P. Predicting volatility in the foreign exchange market. *The Journal of Finance*, v.50, n.2, p.507-528, 1995.
- LJUNG, G. M.; BOX, G. E. P. On a measure of a lack of fit in time series models. *Biometrika*, v.65, n.2, p.297-303, 1978.
- MARTENS, M. Measuring and forecasting S&P 500 index-futures volatility using high-frequency data. *Journal of Futures Markets*, v.22, n.6, p.497-518, 2002.
- MELLO, A. R. A. F. *Volatilidade implícita das opções de ações: uma análise sobre a capacidade de previsão do mercado sobre a volatilidade futura*. 2009. 44 f. Dissertação (Mestrado em Finanças e Economia) – EAESP-FGV, São Paulo, 2009.
- MILACH, F. T. *Estimação da volatilidade: uma aplicação utilizando dados intradiários*. 2010. 96 f. Dissertação (Mestrado em Administração) – PPGA-UFRGS, Porto Alegre, 2010.
- MINCER, J. A.; ZARNOWITZ, V. The evaluation of economic forecasts. *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*. NBER, p.3-46, 1969.
- MOREIRA, J. M. S.; LEMGRUBER, E. F. O uso de dados de alta frequência na estimação da volatilidade e do valor em risco para o IBOVESPA. *Revista Brasileira de Economia*, v.58, n.1, p.100-120, 2004.
- MOTA, B. S.; FERNANDES, M. Desempenho de estimadores de volatilidade na Bolsa de Valores de São Paulo. *Revista Brasileira de Economia*, v.58, n.3, p.429-448, 2004.
- NG, H. S.; LAM, K. P. *How does the sample size affect GARCH model?* Working Paper, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, 2006.
- NICOLAU, J. *Modelação de séries temporais financeiras*. Coleção Econômicas - II série, n.18. Coimbra: Almedina. 2012.
- OOMEN, R. C. A. Using high frequency stock market index data to calculate, model & forecast realized return variance. *European University, Economics Discussion Paper*, n.2001/6, 2001.
- REIS, D. L. P. E. *Análise de desempenho de indicadores de volatilidade*. 2011. 82 f. Dissertação (Mestrado em Economia) – FEA-UFJF, Juiz de Fora, 2011.
- SILVA, C. A. G. Modelagem de estimação da volatilidade do retorno das ações brasileiras: os casos da Petrobrás e Vale. *Cadernos do IME-Série Estatística*, v.26, n.1, p.15-28, 2009.
- SILVA, J. C. *Estimação do valor em risco usando informação intradiária*. 2002, 24 f. Dissertação (Mestrado em Economia) – EAESP-FGV, Rio de Janeiro, 2002.
- SOUZA, M. N. C. *Redução da persistência de volatilidade nos modelos GARCH para cálculo do valor em risco no mercado brasileiro*. 2004, 98 f. Dissertação (Mestrado em Administração) – COPPEAD - UFRJ, Rio de Janeiro, 2004.
- TAYLOR, N. A note on the importance of overnight information in risk management models. *Journal of Banking and Finance*, v.31, p.161-180, 2007.
- TSAY, R. S. *Analysis of financial time series*. New Jersey: Wiley Series in Probability and Statistics, 2010.
- ZIVOT, E. Practical issues in the analysis of univariate GARCH models. *Working Papers UWEC*, University of Washington, Department of Economics. 2008.