

Testando a Existência de Efeitos Lead-Lag entre os Mercados Acionários Norte-Americano e Brasileiro

Gustavo Rezende de Oliveira †

Universidade de Brasília

Otávio Ribeiro de Medeiros ^Ω

Universidade de Brasília

RESUMO: Este trabalho visa identificar o efeito lead-lag entre o mercado acionário norte-americano (NYSE) e o brasileiro (Bovespa), ou seja, se os movimentos de elevação ou queda de preços na NYSE são seguidos, em média, por movimentos similares na Bovespa, permitindo a previsibilidade do valor dos ativos negociados no mercado brasileiro e eventual possibilidade de arbitragem. A existência deste efeito indica uma relativa segmentação entre os mercados, possibilidade não contemplada nas Hipóteses do Mercado Eficiente (HME), segundo a qual os preços dos ativos são imprevisíveis. Identificou-se a existência de co-integração bem como de causalidade (Granger) bidirecional. Através da análise de regressão com vários modelos (regressão linear múltipla, equações simultâneas, VECM e GARCH), constatou-se que o índice Ibovespa é, em grande parte, explicado pelo movimento do Índice Dow Jones em minutos anteriores, divergindo do pressuposto da HME de não previsibilidade de preços. No entanto, os resultados mostraram que a realização de arbitragem não é economicamente viável, em decorrência dos custos de transação.

Palavras-chave: efeito lead-lag, hipótese do mercado eficiente arbitragem.

Recebido em 30/06/2008; revisado em 10/12/2008; aceito em 18/03/2009.

Correspondência com autores*:

† *Professor da Universidade de Brasília*

*Endereço: Av. Parque Águas Claras, 3825, Apto 808,
Brasília –DF, CEP: 71930-000*

e-mail: gustavorezende@hotmail.com

Telefone: (61) 9999-6502

Ω *Professor da Universidade de Brasília*

*Endereço: SQN 205, Bloco C, Apto. 401, Asa Norte,
Brasília – DF, CEP: 70843-030*

e-mail: otavio@unb.br

Telefone: (61) 9978-9503

Nota do Editor: *Este artigo foi aceito por Antonio Lopo Martinez.*

1. INTRODUÇÃO

A evolução tecnológica nos meios de comunicação e na informática, iniciada em meados do século XX, vem contribuindo de forma decisiva para a integração dos mercados acionários mundiais. A integração de mercados significa que o retorno esperado de ativos de uma mesma classe de risco seja o mesmo em diferentes mercados, conforme Eiteman et al. (1994) apud Costa Júnior e Leal (1997). A maior integração entre os mercados financeiros internacionais tende a fazer com que estes se movimentem em conjunto. Dessa forma, um único evento, se de grande relevância, pode fazer com que os preços gerais das ações – índices gerais de bolsas de valores – se elevem ou decresçam em bolsas de valores de todo o mundo instantaneamente.

A Hipótese do Mercado Eficiente – HME estabelece que os preços das ações não são previsíveis, pois se comportam como um passeio aleatório, não permitindo a arbitragem. Além disso, a HME enfoca o mercado individualmente, um ente único, como se houvesse apenas “o mercado” (FAMA, 1970). Se mercados não completamente integrados – ou seja, mercados segmentados – são analisados de forma conjunta, como se fossem apenas um mercado, a HME pode apresentar falhas por várias razões. No entanto, existem mercados mais desenvolvidos que outros o que os levam a apresentar diferentes características, entre elas, a diferença no tamanho. A Bolsa de Valores de Nova York (NYSE), por exemplo, é pelo menos vinte vezes maior do que a Bolsa de Valores de São Paulo (BOVESPA). O grau de concentração em diversas bolsas de valores também varia a depender do mercado. A Bovespa, por exemplo, ainda é bastante concentrada, pois mais de 60% do volume de negócios se relaciona a apenas 19 empresas (ASSAF NETO, 2007). A existência de mecanismos de proteção ao investidor minoritário em algumas bolsas de valores, em detrimento de outras, também indica que alguns mercados já atingiram maior nível de desenvolvimento. Além disso, em função do poder da mídia, em países mais desenvolvidos a informação é disponibilizada de forma mais rápida e eficiente o que leva o investidor de países menos desenvolvidos a operar em desvantagem. Todas estas características, que levam à diferenciação do grau de desenvolvimento entre mercados, permitem que ocorra a assimetria de informações.

Se a HME falhar, os ativos em um mercado desenvolvido podem incorporar o efeito da informação em seus preços de forma mais eficiente. Em outras palavras, movimentos de preços de ativos em um mercado mais eficiente – Estados Unidos, por exemplo – podem preceder os movimentos de preços de um mercado menos desenvolvido, caso de países emergentes como o Brasil, caso estes sejam parcialmente integrados.

Este é o efeito *lead-lag*: ao se considerar os movimentos de preços de dois ou mais mercados, há um que lidera (*lead*) e outro(s) que o segue(m) com uma defasagem (*lag*). Nessa situação, haveria uma quebra da HME e seria possível prever, com certo nível de confiança, os movimentos de preços no(s) mercado(s) liderado(s) em função dos movimentos de preços no mercado líder, sendo possível obter retornos anormais no(s) mercado(s) liderado(s).

O objetivo deste trabalho é verificar a existência do efeito *lead-lag* entre os mercados acionários norte-americano e brasileiro por meio da determinação das defasagens temporais (*lags*) entre os movimentos de índices de preços das ações da NYSE e da Bovespa. Busca-se verificar, também, a possibilidade de se auferir retornos anormais em decorrência dos efeitos *lead-lag*, considerando os custos de transação.

Para tanto, foi utilizado como *proxy* o índice Ibovespa, que reflete o conjunto de ações desta bolsa. Para a NYSE a *proxy* utilizada foi o índice Dow Jones (*Dow Jones Industrial Average* - DJIA). Verificou-se, também, os efeitos das variações ocorridas no DJIA sobre ações específicas do mercado brasileiro. Em função do elevado volume de negociação,

foram analisadas as doze ações com maior ponderação no Ibovespa. Estas ações representam mais de 50% da participação ponderada no índice.

O artigo está estruturado da seguinte forma: a seção 2 trata das evidências empíricas sobre o efeito *lead-lag*. Na seção 3 são apresentadas as metodologias utilizadas na pesquisa. A seção 4 apresenta a aplicação destas metodologias ao caso prático e a seção 5 traz as conclusões do estudo.

2. EVIDÊNCIAS EMPÍRICAS

Vários trabalhos foram desenvolvidos com o fim de verificar a existência de efeitos *lead-lag* entre mercados correlatos. Por exemplo, Brooks, Rew e Ritson (2001) analisaram a existência de relações entre contratos futuros e índice à vista do FTSE 100, índice de ações da Bolsa de Londres, utilizando observações com frequência de 10 minutos. Foi confirmado o que vários outros trabalhos semelhantes identificaram: os retornos preços conduzem os retornos do mercado à vista, confirmando a hipótese de que as informações são primeiramente absorvidas pelo mercado futuro. A razão disto, segundo eles, é que

o índice futuro de preços deve refletir rapidamente toda informação disponível sobre eventos que afetem o preço do ativo. O índice à vista deve responder da mesma maneira, mas, para que o índice reaja à informação completamente, todas as ações que o compõem devem ser reavaliadas de acordo com a nova informação. Como nem todas as ações são negociadas com frequência de 10 minutos, o índice à vista irá responder à informação com defasagem (BROOKS, REW E RITSON, 2001, p. 34).

Nesse estudo, a realização de arbitragem não se mostrou viável porque os custos de transação superariam os retornos anormais

Li, Greco e Chavis (2001) observaram a existência de movimentos prévios nos valores das ações negociadas em Hong Kong, as chamadas ações H, que se seguiam nos valores das ações de mesmas empresas negociadas na China, as chamadas ações A. O efeito *lead-lag* nestes mercados é atribuído ao diferente grau de divulgação de informações nos dois mercados. Os autores perceberam que o efeito *lead-lag* aumentava de acordo com a volatilidade do retorno do ativo.

Foi verificada a presença de efeitos *lead-lag* no mercado acionário indiano, que segmenta o mercado em termos de capitalização de mercado. Poshakwale e Theobald (2002) observaram que empresas com elevada capitalização conduzem o movimento das ações de empresas com baixa capitalização. A velocidade de ajustamento no índice de empresas de grande capitalização é maior que no índice de empresas de baixa capitalização

Em estudo realizado no mercado acionário de Nova York, para o período de agosto de 1984 a setembro de 1987, Chan (1992) concluiu que o índice futuro intradiário influencia o mercado à vista de forma forte, conforme já previsto pela teoria. Ao contrário, há fraca evidência de que o índice à vista influencia o índice futuro. A influência se dá tanto quando o movimento decorre de informações boas, quanto de informações ruins.

Com vistas a verificar o real fluxo de informações, Jong e Nijman (1997) observaram a ocorrência de influência do índice futuro do S&P500 sobre o mercado à vista, tomando como base informações colhidas minuto a minuto. Mais uma vez foi detectado que a informação é absorvida pelo mercado futuro minutos antes do mercado à vista. No caso, a defasagem observada foi, em média, de dez minutos. Nas poucas situações em que o índice à vista liderou o índice futuro, a defasagem foi de, no máximo, dois minutos.

Nesta mesma linha, Kawaller, Koch e Koch (1987) observaram a presença de *lead-lag* entre os índices futuro e à vista do S&P500. Utilizando dados colhidos minuto a minuto, concluíram que a influência do primeiro sobre o segundo se dá em um período entre vinte e quarenta e cinco minutos. Nos poucos momentos em que foi observada a antecipação do índice à vista em relação ao índice futuro, a defasagem foi de no máximo um minuto.

Kofman e Martens (1997) examinaram a correlação nas variações de preços dos índices futuros de ações de Londres e Nova York. Os resultados indicaram que o efeito *lead-lag* no mercado norte-americano prepondera sobre o mercado britânico, embora em alguns momentos ocorra o inverso. Foi também avaliada a possibilidade de ganhos por meio de simulações com cotações reais e considerando custos de transação. O estudo concluiu que existe possibilidade de ganho para aquisição de índice futuro no mercado britânico poucos minutos após a elevação do índice no mercado norte-americano.

3. METODOLOGIA

Esta seção apresenta os procedimentos metodológicos utilizados para atingir os objetivos da pesquisa. Com o propósito de verificar a existência do efeito *lead-lag* entre as bolsas, foram utilizados diversos modelos econométricos: análise de regressão, análise de co-integração, modelo VAR/VECM, análise de causalidade Granger e modelo GARCH.

3.1. Os dados

As variáveis foram obtidas no sistema CMA, com frequência de um minuto, na cotação de fechamento. Todos os testes envolvendo as séries dos índices DJIA e Ibovespa foram realizados com dados de 19 de julho de 2006 a 28 de setembro de 2007. Cada uma das duas séries contém 105.443 observações. Foram consideradas apenas cotações nos períodos em que as duas bolsas estiveram operando simultaneamente. Como os dados são disponibilizados em horário local, fez-se necessário realizar ajustes para que os movimentos das bolsas fossem analisados de forma contemporânea considerando, inclusive, os períodos de horário de verão adotados em cada um dos dois países.

As análises envolvendo empresas que compõem o Ibovespa compreenderam o período de 14 de março de 2007 a 28 de setembro de 2007. Foram utilizadas doze empresas com maior participação ponderada no índice Ibovespa. São as ações Petrobrás PN (PETR4), Vale do Rio Doce PNA (VALE5), Bradesco PN (BBDC4), Usiminas PNA (USIM5), Banco Itaú PN (ITAU4), Telemar PN (TNLP4), Vale do Rio Doce ON (VALE3), Gerdau PN (GGBR4), Petrobrás ON (PETR3), Siderúrgica Nacional (CSNA3), Itausa PN (ITSA4) e Unibanco UNT (UBBR11).

3.2. Análise de Regressão

Parte-se da hipótese de que há efeitos *lead-lag* entre a NYSE e a Bovespa. Esses efeitos podem ser capturados pela regressão:

$$P^{ib}_t = \beta_0 + \sum_{i=1}^n \beta_i P^{ib}_{t-i} + \sum_{j=0}^n \gamma_j P^{dj}_{t-j} + u_t \quad (1)$$

onde P^{ib} é o índice de preços das ações listadas no Ibovespa; P^{dj} é o índice de preços das ações listadas na NYSE; γ_j e β_i são parâmetros a serem estimados na equação e u_t são resíduos aleatórios $\sim N(0, \sigma^2)$.

Como é sabido que os preços das ações são variáveis não-estacionárias¹, é necessário utilizar os retornos dos índices de preços das ações, dados por:

$$R_t = \log \frac{P_t}{P_{t-1}} = \Delta \log P_t \quad (2)$$

Assim, a equação (1) transforma-se em:

$$R^{ib}_t = \beta_0 + \sum_{i=1}^n \beta_i R^{ib}_{t-i} + \sum_{j=0}^n \gamma_j R^{dj}_{t-j} + u_t \quad (3)$$

onde R^{ib} é o retorno do Ibovespa; R^{dj} é o retorno do DJIA; γ_j e β_i são parâmetros a serem estimados na equação e u_t é o erro aleatório $\sim N(0, \sigma^2)$.

A equação (3), porém, pode apresentar dois problemas econométricos. Em primeiro lugar, se R^{dj} é também uma variável endógena, então há um problema de simultaneidade e a equação (3) não pode ser estimada por mínimos quadrados (OLS), pois resultaria em coeficientes viesados e inconsistentes. O segundo problema é que, se P^{ib} e P^{dj} forem séries co-integradas, então, a equação (3) deve incluir um termo de correção de erro.

A solução para o primeiro problema é estimar a equação (3) por um método que leve em consideração a simultaneidade, tal como o TSLS (Mínimos Quadrados em Dois Estágios). Para tanto, é necessário definir as variáveis instrumentais a serem utilizadas que, no caso, serão as próprias variáveis R^{ib} e R^{dj} defasadas.

A solução para o segundo problema é realizar uma análise de co-integração e, caso verificada a sua existência, introduzir na equação (3) o termo de correção de erro. Porém, embora os parâmetros estimados por OLS sejam inconsistentes quando os regressores são co-integrados, isto não ocorre quando a estimação se dá por TSLS: para construção de modelos empíricos, em uma abordagem estrutural, é necessário se preocupar com a questão de identificação e viés de simultaneidade, mas não é necessário se preocupar com questões de não-estacionariedade e co-integração. Na prática, ainda que as séries sejam co-integradas, não é necessário incluir o termo de correção de erro quando se utiliza TSLS (HSIAO, 1997, p. 395).

3.3. Análises de Co-integração

Segundo Harris (1995), o conceito de co-integração indica a existência de um equilíbrio de longo prazo, para o qual o sistema econômico converge no tempo. Para testar a existência de co-integração, neste trabalho, foi empregada a técnica por equação simples de Engle e Granger (1987), por haver apenas duas variáveis estocásticas e, no máximo, um vetor de co-integração.

A regressão de co-integração pelo método de Engle-Granger é dada por:

$$\log P^{ib}_t = \beta_0 + \log \beta_1 P^{ib}_t + u_t \quad (4)$$

onde P^{ib} é o índice de preços das ações listadas no Ibovespa; P^{dj} é o índice de preços das ações listadas na NYSE; β_0 e β_1 são parâmetros a serem estimados na equação e u_t são erros aleatórios $\sim N(0, \sigma^2)$.

Para haver co-integração é necessário que as duas séries tenham a mesma ordem de não-estacionariedade e que os resíduos desta regressão sejam estacionários. Foi utilizado o teste *Ampliado de Dickey Fuller* (ADF) para verificar a existência de não-estacionariedade.

¹ Regressões que utilizam variáveis não-estacionárias são consideradas regressões espúrias, onde podem surgir relações aparentemente significativas, mas que são, na verdade, inexistentes. Vide Brooks (2002, p.367)

Segundo Groppo e Bacchi (2003), as estatísticas mais difundidas para testar a existência de estacionariedade destinam-se a séries que têm, no máximo, uma raiz unitária. No modelo desenvolvido por Dickey e Fuller, conhecido como modelo DF, tem-se que:

$$X_t = \rho X_{t-1} + u_t \quad (5)$$

em que u_t é um erro $\sim N(0, \sigma^2)$. Testa-se $H_0: \rho = 1$ contra a $H_1: \rho < 1$. Se $|\rho| < 1$, X_t é estacionária e descrita por um processo AR(1). Já, se $\rho = 1$, X_t é não-estacionária e descrita por passeio aleatório (*random walk*).

Utiliza-se o método dos mínimos quadrados ordinários para a realização do teste. Os testes de raízes unitárias não empregam a distribuição-padrão t de Student, e sim as estatísticas cujos valores críticos foram tabulados por Dickey e Fuller. Para que as variáveis sejam co-integradas, os resíduos devem ser $I(0)$, ou seja, estacionários. Assim é estimada a regressão entre os logaritmos naturais dos níveis das duas séries para testar a estacionariedade dos resíduos.

Em se verificando a co-integração entre as duas séries, deve ser utilizado o modelo de correção de erro (ECM), que pode ser expresso, conforme abaixo:

$$R_t^{ib} = \beta_0 + \beta_1 R_t^{dj} + \beta_2 (R_{t-1}^{ib} - \beta_3 R_{t-1}^{dj}) + u_t \quad (6)$$

onde R^{ib} é o retorno do índice de preços das ações listadas no Ibovespa; R^{dj} é o retorno do índice de preços das ações listadas na NYSE; β é a constante da equação; β_1 representa a relação de curto prazo entre as duas variáveis; β_2 representa a velocidade de ajustamento do equilíbrio e u_t são resíduos aleatórios $\sim N(0, \sigma^2)$.

Como visto anteriormente, porém, em casos em que a estimação da regressão seja feita por TSLS, não é necessário incluir o termo de correção de erro, ainda que as séries sejam co-integradas.

3.4. O modelo VAR (Vetores auto-regressivos)

Pelo modelo VAR, desenvolvido por Sims (1980), o valor de uma variável é expresso como uma função linear dos valores passados, ou defasados, dessa variável e de todas as outras variáveis incluídas no modelo. Se cada equação contiver o mesmo número de variáveis defasadas no sistema, ela pode ser estimada por OLS, sem que seja preciso recorrer a qualquer método de sistemas, tais como os mínimos quadrados em dois estágios. O modelo efetivo a ser estimado é:

$$R_t^{ib} = \beta_0 + \sum_{i=1}^n \beta_i R_{t-i}^{ib} + \sum_{j=1}^n \gamma_j R_{t-j}^{dj} + u_t \quad (7)$$

$$R_t^{dj} = \alpha_0 + \sum_{i=1}^n \alpha_i R_{t-i}^{dj} + \sum_{j=1}^n \delta_j R_{t-j}^{ib} + v_t \quad (8)$$

onde R^{ib} é o retorno do Ibovespa; R^{dj} é o retorno do DJIA; β_i , γ_j , α_i , δ_j , são parâmetros a serem estimados na equação e u_t e v_t são os erros aleatórios $\sim N(0, \sigma^2)$.

3.4.1. O modelo VECM (Vector Error Correction Model)

A estimação do Vetor de Correção de Erros, VECM, foi popularizada por Engle e Granger (1987) e tem por objetivo analisar os ajustamentos de curto prazo que ocorrem nas séries co-integradas. Se os mercados funcionam perfeitamente, os movimentos de preço nesses mercados devem ser mais bem descritos pelo Vetor de Correção de Erro - VECM de primeira ordem, com o termo de correção de erro sendo a diferença do retorno entre os dois mercados. O modelo VECM é dado por:

$$R_t^{dj} - R_{t-1}^{dj} = \beta_1 [R_{t-1}^{dj} - R_{t-1}^{ib}] + u_t \quad (9)$$

$$R_t^{ib} - R_{t-1}^{ib} = \beta_2 [R_{t-1}^{dj} - R_{t-1}^{ib}] + v_t \quad (10)$$

onde R^{ib} é o retorno do Ibovespa; R^{dj} é o retorno do DJIA; β_1 e β_2 são parâmetros a serem estimados na equação e u_t e v_t são os erros aleatórios $\sim N(0, \sigma^2)$.

3.5. Causalidade de Granger

A fim de verificar o sentido de causalidade entre as séries, será realizado um teste de causalidade Granger (1969). O conceito de causalidade entre as variáveis diz respeito à precedência temporal, ou seja, se a variável R^{dj} causa a variável R^{ib} , então os valores defasados da variável R^{dj} , ajudam a prever R^{ib} . As seguintes equações compõem o modelo VAR, utilizado para testar a direção de causalidade entre os dois índices.

$$R_t^{ib} = \beta_0 + \sum_{i=1}^n \beta_i R_{t-i}^{ib} + \sum_{j=1}^n \gamma_j R_{t-j}^{dj} + u_t \quad (11)$$

$$R_t^{dj} = \alpha_0 + \sum_{i=1}^n \alpha_i R_{t-i}^{dj} + \sum_{j=1}^n \delta_j R_{t-j}^{ib} + v_t \quad (12)$$

onde R^{ib} é o retorno do Ibovespa; R^{dj} é o retorno do DJIA; $\beta_i, \gamma_j, \alpha_i, \delta_j$, são parâmetros a serem estimados na equação e u_t e v_t são erros aleatórios $\sim N(0, \sigma^2)$.

As hipóteses a serem testadas são de que os coeficientes dos valores passados da variável exógena das equações acima são iguais a zero, isto é:

$$\gamma_1 = \gamma_2 = \dots = \gamma_p = 0 \quad (13)$$

$$\beta_1 = \beta_2 = \dots = \beta_p = 0 \quad (14)$$

O teste F é utilizado para determinar o sentido de causalidade. No caso em que as duas hipóteses são rejeitadas, haverá relação bicausal; se as duas hipóteses não são rejeitadas, há ausência de causalidade.

3.6. O modelo GARCH

Se os resíduos de uma regressão apresentam uma variância que não é constante, mas sim variável no tempo, isto é, apresentam heteroscedasticidade condicional autoregressiva (efeitos ARCH), é necessário especificar um modelo da família ARCH (*Autoregressive Conditional Heteroscedasticity*) e estimar a referida regressão pelo método de máxima verossimilhança. No presente trabalho, utilizou-se um modelo GARCH (*Generalized Autoregressive Conditional Heteroscedasticity*). O modelo GARCH pode ser expresso como:

$$R_t^{ib} = \alpha + \beta R_t^{dj} + u_t \quad u_t \sim N(0, \sigma_t^2) \quad (15)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (16)$$

onde R^{ib} é o retorno do Ibovespa, R^{dj} é o retorno do DJIA e σ_t^2 é a variância dos erros u_t . Este modelo assume que a variância dos erros não é constante, o que o torna mais adequado para análise de séries de tempo financeiras cujas variâncias dos erros não são constantes ao longo do tempo.

4. RESULTADOS

De acordo com o esperado, a co-integração foi identificada entre os dois mercados. Primeiramente, verificou-se a existência de raízes unitárias para o logaritmo natural das duas séries conforme tabelas 1 e 2:

Tabela 1: Teste de raiz unitária – Dow Jones, no nível

Null Hypothesis: LOG(DJ) has a unit root		
Lag Length: 1 (Automatic based on SIC, MAXLAG=68)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.559289	0.5035
Test critical values: 1% level	-3.430243	
5% level	-2.861377	
10% level	-2.566723	

Observou-se que a hipótese nula – de que o logaritmo natural do DJIA, no nível, possui raiz unitária – não foi rejeitada.

Tabela 2: Teste de raiz unitária – Ibovespa, no nível

Null Hypothesis: LOG(IBOV) has a unit root		
Lag Length: 4 (Automatic based on SIC, MAXLAG=68)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.567703	0.8753
Test critical values: 1% level	-3.430243	
5% level	-2.861377	
10% level	-2.566723	

Da mesma forma, a hipótese nula de que o logaritmo natural do Ibovespa tem uma raiz unitária no nível, não foi rejeitada.

Na primeira diferença, porém, em ambos os casos, a hipótese nula de haver raiz unitária foi rejeitada, conforme tabelas 3 e 4, abaixo. Esta é a primeira condição para que haja co-integração: que as duas séries possuam o mesmo nível de integração. Identificou-se integração de primeira ordem (I(1)) nos logaritmos naturais das duas séries.

Tabela 3: Teste de raiz unitária – Dow Jones, primeira diferença

Null Hypothesis: D(LOG(DJ)) has a unit root		
Lag Length: 0 (Automatic based on SIC, MAXLAG=68)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-316.7573	0.0001
Test critical values: 1% level	-3.430243	
5% level	-2.861377	
10% level	-2.566723	

Tabela 4: Teste de raiz unitária – Ibovespa, primeira diferença

Null Hypothesis: D(LOG(IBOV)) has a unit root		
Lag Length: 3 (Automatic based on SIC, MAXLAG=68)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-149.6146	0.0001
Test critical values: 1% level	-3.430243	
5% level	-2.861377	
10% level	-2.566723	

Assim, na primeira diferença dos logaritmos naturais, tanto a série DJIA quanto a série Ibovespa são estacionárias.

Em seguida, os resíduos da regressão entre as duas séries foram isolados a fim de se verificar a existência de raiz unitária, já que o método de Engle-Granger prevê que, caso os resíduos da regressão de duas séries não-estacionárias de mesma ordem não possuam raiz unitária, ou seja, sejam estacionários, haverá co-integração entre as duas séries. A tabela 5 apresenta os resultados do teste de raiz unitária dos resíduos da regressão:

Tabela 5: Teste de raiz unitária – Resíduos, no nível

Null Hypothesis: U has a unit root		
Lag Length: 8 (Automatic based on SIC, MAXLAG=68)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.043122	0.0311
Test critical values: 1% level	-3.430243	
5% level	-2.861377	
10% level	-2.566723	

Como a hipótese nula foi rejeitada a 5% de significância, pôde-se concluir que há co-integração entre as séries DJIA e Ibovespa. A existência de co-integração implica a adoção do termo de correção de erro para estimar a regressão por OLS. Porém, como visto anteriormente, a inclusão do termo de correção de erro é desnecessária quando a estimação da regressão se dá por meio de TSLS.

A fim de verificar o nível de dependência do valor corrente do índice do Ibovespa aos valores prévios das duas variáveis (R^{ib} e R^{dj}), foi necessário utilizar um modelo VAR. Como exposto anteriormente, quando a relação entre duas variáveis não-estacionárias é importante, a sua transformação em séries estacionárias por meio de sua diferença não é suficiente para que o resultado de sua regressão seja válido.

Por definição, no longo prazo, as variáveis convergem para um mesmo valor e não se alteram. Assim, a diferença entre suas variáveis passa a ser zero (BROOKS, 2002). Dessa forma, para que este erro fosse eliminado, foi utilizado o VECM (*Vector Error Correction Model*).

Dos resultados puderam-se concluir vários aspectos importantes da relação entre as séries, inclusive o câmbio.

A primeira conclusão importante é a de que os períodos de defasagem da variável DJIA significativas vão até os 9 minutos anteriores ao minuto corrente. Os coeficientes para períodos de defasagem anteriores se mostraram nulos. Este resultado é extremamente importante na medida em que demonstra que movimentos passados na NYSE exercem influência no valor corrente das ações que compõem o índice Ibovespa.

À exceção do período $t-6$, os *lags* do próprio Ibovespa que exercem influência sobre o valor corrente vão até o minuto $t-8$, porém, todos estes coeficientes apresentaram valores absolutos baixos.

Outra conclusão relevante é a de que a variável câmbio mostrou-se insignificante em quase todos os tempos de defasagem. Apenas para os períodos defasados $t-13$, $t-14$, $t-15$ e $t-18$ a hipótese nula, de que os coeficientes são iguais a zero, foi rejeitada, a 5%. Embora sejam significativos, todos estes coeficientes apresentam valores extremamente baixos. Para os períodos defasados $t-13$, $t-14$, $t-15$ e $t-18$ os valores dos coeficientes são de 0,0008, 0,0007, 0,0009 e 0,0007, respectivamente.

Por ter se mostrado pouco influente na variação do Ibovespa, a variável câmbio foi excluída do modelo nos testes seguintes. Por esta mesma razão o modelo VECM foi novamente estimado sem a variável câmbio.

Os *lags* da variável DJIA com coeficientes significativos pouco variaram do modelo anterior, em que o câmbio estava presente. A diferença é que o tempo de influência se elevou para dez minutos. Para os minutos $t-1$ a $t-10$ os coeficientes – todos significativos – foram de 0,367, 0,225, 0,111, 0,046, 0,0364, 0,035, 0,026, 0,021, 0,013 e 0,021, respectivamente. Interessante notar que todos estes coeficientes são positivos. O significado destes números é interessante, pois indica em que proporção cada movimento, em instantes anteriores, na NYSE reflete no nível do índice da Bovespa. Assim, por este modelo, se, por exemplo, há um minuto houve uma queda de 0,02% na bolsa norte-americana, este evento contribuirá para que a bolsa brasileira caia 0,074% ($0,02\% \times 0,36950$ – coeficiente do *lag* $t-1$) no presente minuto. Naturalmente que o valor das ações do índice Bovespa recebe outras influências, inclusive aquelas de abrangência apenas local, mas é inegável que a relação entre as duas bolsas é muito forte.

Esses coeficientes indicam que a informação não é refletida instantaneamente no preço de todas as ações que compõem a cesta do Ibovespa, assim como não refletiria em qualquer outro mercado de ativos. Informações relevantes, que forçam uma variação no preço de determinado ativo ou conjunto de ativos em nível mundial, não estão disponíveis para todos imediatamente. Este lapso de dez minutos pode ser entendido como tempo médio necessário para que o efeito de certa informação esteja incorporado ao valor de determinado ativo.

É interessante observar que, como era de se esperar, os coeficientes das variáveis defasadas do índice da NYSE vão se reduzindo em valores absolutos ao longo de cada minuto que antecede o valor atual do índice Bovespa. Os três minutos mais recentes são cruciais para a assimilação da informação pelo valor do ativo enquanto os sete minutos mais distantes exercem uma influência um pouco menor, até que, após 10 minutos, o efeito da informação já está refletido no valor da ação.

Apenas quatro, dos vinte coeficientes defasados da variável Ibovespa testados, foram significativos. Foram eles os minutos $t-1$, $t-2$, $t-3$ e $t-7$, com os coeficientes de 0,016, -0,009, -0,103 e -0,117, respectivamente. À exceção do minuto imediatamente anterior ao minuto corrente, todos os outros coeficientes significativos foram negativos, denotando uma relação inversa entre os minutos passados e presentes da variável Ibovespa. Nota-se que todos os quatro coeficientes apresentaram valores absolutos baixos, sendo o maior de 11,74%.

A fim de confirmar a forte relação existente entre as duas variáveis, foi utilizado o modelo de Causalidade Granger, com 20 valores defasados. Os resultados estão na tabela 6. O teste foi realizado com a diferença do logaritmo de cada elemento da amostra.

Tabela 6: Teste de causalidade de Granger

Pairwise Granger Causality Tests			
Date: 03/11/08 Time: 17:14			
Sample: 1 105443			
Lags: 20			
Null Hypothesis:	Obs	F-Statistic	Probability
DLOG(DJ) does not Granger Cause DLOG(IBOV)	105422	281,870	0.00000
DLOG(IBOV) does not Granger Cause DLOG(DJ)		17,3089	3.5E-61

A significância das relações entre as séries é dada pela estatística F, que testa a hipótese de que os coeficientes das variáveis das equações são nulas.

Ambas as hipóteses nulas – DJIA não causa Ibovespa e Ibovespa não causa DJIA – foram rejeitadas. Este resultado, de causalidade bi-direcional, confirma a forte relação entre as duas variáveis.

Para confirmar as conclusões obtidas por meio das metodologias já apresentadas, foram realizadas duas outras regressões, desta vez utilizando o método TSLS (*Two-Stage Least Squares*), cujos resultados serão apresentados a seguir.

No sub-item 2.2 foram apresentados dois problemas na estimação da equação (3) por OLS. Um desses problemas era que, se R^{dj} fosse influenciado por R^{ib} , ou seja, a variável R^{dj} ser endógena, a regressão deveria ser estimada por TSLS.

Pelas tabelas abaixo, é possível verificar que, nas regressões em que R^{dj} é variável dependente, a série R^{ib} é significativa. Com a utilização de *dummy* para retornos positivos, tanto a variável R^{ib} contemporânea quanto as defasagens até o quarto minuto são significativas.

Tabela 7: Regressão OLS das variáveis R^{ib} (dependente) e R^{dj} com *dummy* para retornos positivos

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7,71E-06	7,08E-07	10,88269	0,000000
DLOG(DJ(-1))	0,004417	0,002063	2,140630	0,032300
DLOG(DJ(-2))	-0,010404	0,002078	-5,006047	0,000000
DLOG(DJ(-3))	-0,008022	0,002081	-3,855710	0,000100
DLOG(DJ(-4))	-0,009918	0,002077	-4,776043	0,000000
DLOG(IBOV)	0,011604	0,001086	10,68425	0,000000
DLOG(IBOV(-1))	0,014103	0,001126	12,52486	0,000000
DLOG(IBOV(-2))	0,007130	0,001125	6,337884	0,000000
DLOG(IBOV(-3))	0,003526	0,001117	3,157902	0,001600
DLOG(IBOV(-4))	0,004800	0,001093	4,392987	0,000000

(continua na próxima página)

Tabela 7: Regressão OLS das variáveis R^{ib} (dependente) e R^{dj} com *dummy* para retornos positivos (continuação)

R-squared	0,003958	Mean dependent var	7.86E-06
Adjusted R-squared	0,003873	S.D. dependent var	0,00023
S.E. of regression	0,00023	Akaike info criterion	-13,91801
Sum squared resid	0,005571	Schwarz criterion	-13,91710
Log likelihood	733753,7	F-statistic	46,54395
Durbin-Watson stat	2,012031	Prob(F-statistic)	0,00000

Com a utilização de variável *dummy* para retornos negativos, apenas a variável R^{ib} em sua defasagem de quatro minutos não apresentou significância estatística a 5%, conforme tabela abaixo.

Tabela 8: Regressão OLS das variáveis R^{ib} (dependente) e R^{dj} com *dummy* para retornos negativos

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3,21E-05	6,61E-07	-48,59469	0,00000
DLOG(DJ(-1))	0,008849	0,001927	4,592801	0,00000
DLOG(DJ(-2))	-0,004084	0,001941	-2,104309	0,03540
DLOG(DJ(-3))	-0,002276	0,001943	-1,171414	0,24140
DLOG(DJ(-4))	-0,002620	0,001939	-1,351224	0,17660
DLOG(IBOV)	0,011994	0,001014	11,82640	0,00000
DLOG(IBOV(-1))	0,013078	0,001051	12,43855	0,00000
DLOG(IBOV(-2))	0,004216	0,001051	4,013318	0,00010
DLOG(IBOV(-3))	0,002336	0,001043	2,239933	0,02510
DLOG(IBOV(-4))	-0,000209	0,001020	-0,204878	0,83770
R-squared	0,004335	Mean dependent var	-3,20E-05	
Adjusted R-squared	0,00425	S.D. dependent var	0,000215	
S.E. of regression	0,000215	Akaike info criterion	-14,05503	
Sum squared resid	0,004858	Schwarz criterion	-14,05412	
Log likelihood	740969,9	F-statistic	50,99688	
Durbin-Watson stat	1938121	Prob(F-statistic)	0,00000	

Os resultados destas regressões indicaram que a série DJIA é endógena. Por esta razão, novas regressões foram realizadas, desta vez utilizando-se o método TSLS.

Das quinze variáveis da regressão utilizando *dummy* de retorno positivo, apenas cinco não foram estatisticamente significativas, a 5%, conforme pode ser observado na tabela abaixo:

Tabela 9: Regressão TSLS das variáveis R^{ib} (dependente) e R^{dj} com *dummy* para retornos positivos

Variable	Coefficient	Std. Error	t-Statistic	Prob.	Non-signific. at 5%
C	0,000149	8,64E-07	172,8133	0,0000	
DLOG(IBOV(-1))	-0,000484	0,001377	-0,351568	0,7252	*
DLOG(IBOV(-2))	-0,007845	0,001377	-5,698065	0,0000	
DLOG(IBOV(-3))	-0,007536	0,001377	-5,472939	0,0000	
DLOG(IBOV(-4))	-0,002102	0,001377	-1,526518	0,1269	*
DLOG(DJ)	0,037001	0,002383	15,5286	0,0000	
DLOG(DJ(-1))	0,193812	0,002472	78,41562	0,0000	
DLOG(DJ(-2))	0,114543	0,002521	45,43227	0,0000	
DLOG(DJ(-3))	0,050621	0,00254	19,92839	0,0000	
DLOG(DJ(-4))	0,024446	0,002543	9,614649	0,0000	
DLOG(DJ(-5))	0,013615	0,002457	5,542427	0,0000	
DLOG(DJ(-6))	0,009642	0,002406	4,008208	0,0001	
DLOG(DJ(-7))	0,009991	0,002386	4,186484	0,0000	
DLOG(DJ(-8))	0,002969	0,002382	1,246399	0,2126	*
DLOG(DJ(-9))	-0,002996	0,002381	-1,258213	0,2083	*
DLOG(DJ(-10))	-0,003171	0,00238	-1,332615	0,1827	*

Apenas as variáveis de primeiro e último *lags* do Ibovespa, bem como as variáveis dos três últimos *lags* do DJIA não se apresentaram estatisticamente significativos. Todos os coeficientes significativos das variáveis defasadas do índice da bolsa americana foram positivos. Esta correlação positiva se reduz progressivamente até o minuto $t-7$, com resultados semelhantes ao obtido pelo modelo VEC.

A principal diferença entre este resultado e aquele obtido pelo modelo VECM, para retornos positivos, é que o coeficiente da variável do minuto contemporâneo, neste modelo TSLS, é inferior ao da variável de um minuto de defasagem.

A soma de todos os coeficientes é de 0,454. Isto indica que o retorno positivo do minuto atual da NYSE e os sete minutos que o antecedem, explicam 45,36% do retorno no minuto $t-0$ da Bovespa.

Os coeficientes defasados do próprio índice Ibovespa pouco explicam o seu valor contemporâneo. Como visto anteriormente, apenas os coeficientes dos minutos $t-2$ e $t-3$ foram significativos, com valores de -0,008 e -0,007 respectivamente, apresentando o total pouco expressivo de -1,53%. Como se verá à frente, os coeficientes defasados do Ibovespa para retornos negativos apresentam sinais positivos, ou seja, para retornos positivos, a relação dos *lags* do Ibovespa é inversa, para retornos negativos, é direta.

Marcados com asteriscos na 6ª coluna da tabela abaixo, vêm-se os coeficientes que não apresentaram significância estatística a 5%:

Tabela 10: Regressão TSLS das variáveis R^{ib} (dependente) e R^{dj} com *dummy* para retornos negativos

Variável	Coefficiente	Erro Padrão	Estatística t	P-valor	Não-signif. a 5%
C	-2,64E-05	9,78E-07	-26,96551	0,0000	
DLOG(IBOV(-1))	0,005501	0,001559	3,529149	0,0004	
DLOG(IBOV(-2))	-0,001928	0,001559	-1,237119	0,216	*
DLOG(IBOV(-3))	-0,00173	0,001559	-1,109718	0,2671	*
DLOG(IBOV(-4))	0,004351	0,001559	2,790832	0,0053	
DLOG(DJ)	0,042383	0,002698	15,71064	0,0000	
DLOG(DJ(-1))	0,172321	0,002798	61,58221	0,0000	
DLOG(DJ(-2))	0,107115	0,002854	37,5268	0,0000	
DLOG(DJ(-3))	0,060432	0,002876	21,0136	0,0000	
DLOG(DJ(-4))	0,019266	0,002879	6,693029	0,0000	
DLOG(DJ(-5))	0,017923	0,002781	6,444348	0,0000	
DLOG(DJ(-6))	0,005979	0,002723	2,195263	0,0281	
DLOG(DJ(-7))	0,013378	0,002702	4,951372	0,0000	
DLOG(DJ(-8))	0,012126	0,002697	4,496686	0,0000	
DLOG(DJ(-9))	0,004597	0,002696	1,705191	0,0882	*
DLOG(DJ(-10))	0,006779	0,002694	2,516146	0,0119	

Os coeficientes com defasagens de dois e três minutos do Ibovespa, bem como a nove minutos do DJIA foram as únicas hipóteses nulas de coeficientes igual a zero não rejeitadas. Todos os outros coeficientes apresentaram resultados positivos.

A soma de todos os coeficientes é de 0,458, valor muito semelhante à soma dos coeficientes da NYSE para retornos positivos. Este valor sugere que o retorno negativo do minuto atual da NYSE e os dez minutos que o antecedem, explicam 45,77% do retorno no minuto $t-0$ da Bovespa.

Algumas premissas têm que ser consideradas para que a regressão TSLS seja válida. Foram realizados os testes de autocorrelação, de normalidade e de heterocedasticidade. A seguir são apresentados os gráficos dos resíduos das regressões TSLS.

A autocorrelação ocorre em análises de séries temporais quando os erros associados com observações em um dado período de tempo se mantêm por transferência nos períodos de tempos futuros. O teste Durbin-Watson, dos mais populares, não pode ser utilizado para regressões em que defasagens da variável dependente sejam regressores. Por isso, utilizou-se o teste Breusch-Godfrey. Conforme pode ser observado abaixo, ambas as regressões, com variáveis *dummy* para retornos negativos e positivos, apresentaram autocorrelação nos resíduos, pois a hipótese nula de que não há autocorrelação foi rejeitada nos dois casos.

Tabela 11: Teste de autocorrelação dos resíduos BG para regressão TSLS com *dummy* para retornos positivos

Breusch-Godfrey Serial Correlation LM Test:			
Obs*R-squared	4.058,581	Probability	0,000000

Tabela 12: Teste de autocorrelação dos resíduos BG para regressão TSLS com *dummy* para retornos negativos

Breusch-Godfrey Serial Correlation LM Test:			
Obs*R-squared	1.312,707	Probability	0,000000

A presença de autocorrelação dos resíduos indica que os coeficientes estimados sejam ineficientes, isto é, não tenham a menor variância possível. Eles, porém, não são viesados (BROOKS, 2002, p. 166)

Uma das propriedades dos estimadores de mínimos quadrados é a de que os erros seguem uma distribuição normal. Dessa forma, é razoável assumir esta premissa se partirmos do significado do termo de erro: uma soma de fatores que não foram incluídos no modelo. Se forem muitos estes fatores, a sua soma deverá seguir uma distribuição normal, de acordo com o teorema do limite central.

Para certificar a normalidade dos resíduos, foi realizado o teste Jarque-Bera, que utiliza os resultados para os momentos da distribuição normal, em particular os coeficientes de assimetria – que é zero para a distribuição normal – e de curtose – que é igual a 3 na distribuição normal (SARTORIS, 2007). Em ambos os casos, ou seja, para as regressões com variáveis *dummy* para retornos positivos e negativos, a hipótese nula de que os resíduos são normalmente distribuídos foi rejeitada. Embora os resíduos não tenham seguido uma distribuição normal, a violação a esta premissa é virtualmente inconseqüente, para amostras grandes (BROOKS, 2002, p. 182).

Outra propriedade dos estimadores de mínimos quadrados é a de que a variância dos erros deve ser constante, o que é conhecido como homocedasticidade, o que significa ausência de heteroscedasticidade. A conseqüência da violação desta hipótese é a de que, embora os estimadores de mínimos quadrados ordinários não sejam viesados, estes não são o de menor variância.

Tanto para a regressão com variáveis *dummy* para retornos positivos, quanto para retornos negativos, a hipótese nula de que os erros são homocedásticos foi rejeitada.

Tabela 13: Teste White de heteroscedasticidade dos resíduos para regressão TSLS com *dummy* para retornos positivos

White Heteroskedasticity Test:			
F-statistic	107,6836	Probability	0,000000
Obs*R-squared	3.135,362	Probability	0,000000

Tabela 14: Teste White de heteroscedasticidade dos resíduos para regressão TSLS com *dummy* para retornos negativos

White Heteroskedasticity Test:			
F-statistic	127,4389	Probability	0,000000
Obs*R-squared	3.690,430	Probability	0,000000

Em razão da presença de heteroscedasticidade nas regressões, foram realizadas outras regressões pela metodologia GARCH, que incorpora a inconstância da variância no próprio modelo.

A resposta à pergunta de pesquisa só será possível se o número de acertos na previsão de retornos positivos, com base nos retornos do DJIA, for superior ao número de erros. Para realizar este teste, foi selecionado, aleatoriamente, um intervalo de trinta dias (14/03/07 a 12/04/07) com dados reais das doze ações com maior ponderação na composição do Ibovespa. Para cada empresa foram utilizadas 7.598 observações.

Apenas cinco *lags* foram utilizados neste teste. Os valores do minuto contemporâneo do DJIA e dos quatro minutos anteriores foram ponderados pelos respectivos coeficientes para identificar um retorno de alta para o minuto seguinte. A previsão para dois minutos a frente baseou-se na cotação corrente do índice americano e nos três minutos anteriores ponderados pelos seus coeficientes. Assim, sucessivamente, até a previsão para cinco minutos no futuro, quando a única informação disponível era a cotação corrente do DJIA.

Os resultados confirmaram a hipótese de que é possível prever, com razoável grau de acerto, os retornos positivos da Bovespa com base, apenas, nos movimentos da NYSE.

Conforme pode ser observado na tabela abaixo, o percentual de acertos para previsões de retornos positivos chegou a mais de 60% para previsões de um a quatro minutos futuros. Para retorno positivo para o quinto minuto futuro, o percentual de acerto chegou a 59,40%. O menor índice de acerto foi de 54,53%, para o quinto minuto. Os dados relativos aos retornos iguais a zero não foram computados. Isto ocorre quando não há variação do valor entre as duas cotações em análise ou quando o dado não está disponível para determinado minuto.

Tabela 15: Previsão de retornos positivos para doze ações do Ibovespa, por intervalo de tempo.

Ações	1 minuto			2 minutos			3 minutos			4 minutos			5 minutos		
	certas	falhas	% ac.	certas	falhas	% ac.	certas	falhas	% ac.	certas	falhas	% ac.	certas	falhas	% ac.
Petrobrás PN (PETR4)	1.755	1.238	58,64 %	1.854	1.288	59,01 %	1.846	1.412	56,66 %	1.870	1.432	56,63 %	1.870	1.430	56,67 %
Vale do Rio Doce PNA (VALE5)	1.488	936	61,39 %	1.607	933	63,27 %	1.613	991	61,94 %	1.597	1.042	60,52 %	1.570	1.073	59,40 %
Bradesco PN (BBDC4)	1.181	779	60,26 %	1.301	838	60,82 %	1.387	898	60,70 %	1.392	956	59,28 %	1.391	998	58,23 %
Usiminas PNA (USIM5)	885	616	58,96 %	960	653	59,52 %	1.005	652	60,65 %	982	745	56,86 %	994	755	56,83 %
Banco Itaú PN (ITAU4)	921	623	59,65 %	986	678	59,25 %	1.003	705	58,72 %	1.039	742	58,34 %	1.021	763	57,23 %
Telemar PN (TNLP4)	717	498	59,01 %	771	505	60,42 %	824	553	59,84 %	822	582	58,55 %	818	577	58,64 %
Vale do Rio Doce ON (VALE3)	828	538	60,61 %	886	552	61,61 %	855	595	58,97 %	867	566	60,50 %	851	585	59,26 %
Gerdau PN (GGBR4)	1.018	677	60,06 %	1.118	713	61,06 %	1.191	755	61,20 %	1.178	852	58,03 %	1.161	854	57,62 %
Petrobrás ON (PETR3)	729	446	62,04 %	741	515	59,00 %	738	542	57,66 %	741	581	56,05 %	753	548	57,88 %
Siderúrgica Nacional (CSNA3)	929	583	61,44 %	997	598	62,51 %	1.027	657	60,99 %	1.021	706	59,12 %	1.022	721	58,63 %
Itausa PN (ITSA4)	554	419	56,94 %	634	450	58,49 %	735	506	59,23 %	736	567	56,49 %	792	558	58,67 %
Unibanco UNT (UBBR11)	657	527	55,49 %	806	543	59,75 %	810	625	56,45 %	801	667	54,56 %	813	678	54,53 %
Acerto máximo			62,04 %			63,27 %			61,94 %			60,52 %			59,40 %
Acerto mínimo			55,49 %			58,49 %			56,45 %			54,56 %			54,53 %
Acerto médio			59,54 %			60,39 %			59,42 %			57,91 %			57,80 %

Resta, porém, avaliar a possibilidade de realização de arbitragem, só possível caso o custo de transação seja inferior ao retorno proporcionado pela negociação. Para cada operação, seja de compra, seja de venda, a Bovespa cobra 0,019% de taxa de liquidação mais 0,006% de emolumentos. Como a arbitragem envolve a compra e a venda do ativo, o custo de transação é de 0,05% do valor negociado. Há, também, o custo de corretagem (em média, R\$ 10,00 por operação). Quanto maior o volume financeiro menos significativo se torna este custo, razão pela qual este valor pode ser desconsiderado da análise.

Com base nos retornos positivos previstos pelo teste acima (período de 14/03/07 a 12/04/07), foi obtida a média dos retornos positivos. Nenhum superou o retorno de 0,05%. A possibilidade de arbitragem é inviabilizada exclusivamente pelos custos de transação e não pela possibilidade de previsibilidade de retornos positivos com base na variação do índice da bolsa americana.

Como as regressões por TSLS apresentaram heterocedasticidade foram realizadas outras regressões pelo método GARCH, modelo que não assume que a variância é constante.

O modelo GARCH também foi utilizado com o fim de verificar o grau de influência de variáveis defasadas no valor corrente do índice da Bovespa. Nesta fase, a análise foi dividida em duas partes: apuração dos coeficientes para retornos de alta e para retornos de baixa. Para tanto, foram criadas variáveis *dummy* para anular os efeitos dos intervalos de alta e intervalos de baixa, alternadamente.

Outra observação importante refere-se às altas volatilidades verificadas entre o fechamento da bolsa e abertura no dia seguinte. Para excluir este efeito, outra variável *dummy* foi criada para anular o resultado destes retornos anormais.

Como o resultado do modelo VECM demonstrou que apenas as dez primeiras defasagens da variável DJ eram significantes, a regressão com utilização de GARCH foi realizada com este número de *lags*. Para a variável Ibovespa, foram utilizados quatro *lags*.

Para retornos de alta, todos os coeficientes das quatro defasagens da variável Ibovespa foram estatisticamente significantes e apresentaram os seguintes valores de $t-1$ a $t-4$, respectivamente: 0,038, -0,038, 0,006 e 0,039. Para os retornos de baixa, apenas os *lags* $t-1$ e $t-4$ se mostraram significativos, com coeficientes de 0,011 e 0,003, nesta ordem. Nos retornos positivos, a influência do DJIA foi estatisticamente significativa em todos os dez *lags*.

Para os retornos negativos, apenas um dos coeficientes não apresentou significância estatística: o de defasagem de nove minutos. Destaca-se o fato do coeficiente do minuto $t-1$ ser superior ao do minuto $t-0$, o que pode significar certa cautela do investidor para movimentos negativos. Diferentemente do que ocorre com os coeficientes com a variável *dummy* para retornos positivos, os cinco primeiros *lags* da regressão com a *dummy* para retornos de baixa são positivos.

5. CONCLUSÕES

O objetivo deste trabalho foi testar a existência de efeitos *lead-lag* entre a NYSE e a Bovespa. Este objetivo foi verificado por meio da utilização de vários testes.

Primeiramente, a forte relação existente entre as duas bolsas de valores foi percebida pela identificação de co-integração (teste de Engle-Granger) entre as séries não-estacionárias de primeira ordem que compõem essas bolsas.

O teste de Causalidade Granger detectou causalidade bidirecional entre os mercados, confirmando esta relação entre as séries.

A integração identificada, porém, não é total já que os movimentos defasados de ambas as séries exercem influência sobre o minuto $t-0$ do índice Ibovespa. Se a integração não é total, há segmentação entre os mercados, o que abre margem à arbitragem.

A fim de verificar em que intensidade os períodos defasados das duas séries explicam o índice Ibovespa no minuto contemporâneo, alguns modelos foram utilizados no logaritmo natural do retorno destas séries: um modelo de Vetor Auto-regressivo com correção de erro (VECM), com sessenta *lags*, um modelo GARCH e um modelo de equações simultâneas estimado por TSLS. Para os dois últimos foram utilizados dez *lags* para série DJIA e quatro para a série Ibovespa. Em todos os testes foi observada a influência dos *lags* da NYSE no minuto corrente do Ibovespa.

Com base nos resultados obtidos em todos estes testes, foi possível afirmar que há evidências significativas de que é possível fazer alguma previsão sobre o Ibovespa com base em informações sobre a trajetória precedente do DJIA.

Estes resultados revelam que o retorno da Bovespa é, em grande parte, explicado pelo movimento da NYSE em minutos anteriores (e também, em pequena escala, por seus próprios movimentos defasados).

Ainda que óbvio, é interessante lembrar que o valor dos ativos que compõem o Ibovespa não está sujeito apenas às movimentações da NYSE. Por um lado, há informações que afetam apenas aquele mercado. Por outro lado, outros eventos, de abrangência local, afetam apenas o mercado nacional. Além disso, a reação dos investidores à determinada informação, ainda que de alcance global, é atenuada ou acentuada por circunstâncias locais. Por exemplo, a elevação do preço do petróleo no mercado externo pode não ser refletida na mesma proporção em mercados diversos por estarem eles sujeitos a, por exemplo, sua dependência externa pelo produto, existência de combustíveis alternativos, etc.

Independentemente da reação aos eventos não ser exata, o fato é que os movimentos dos dois mercados analisados são extremamente semelhantes, com uma certa precedência dos eventos no mercado da bolsa norte-americana.

Dessa forma, pode-se inferir que um dos pressupostos da Hipótese do Mercado Eficiente (HME) é questionável: as informações não estão sendo incorporadas nos preços dos ativos instantaneamente.

De fato, se é possível prever, em parte, o movimento de uma determinada bolsa de valores, também é possível que se possa realizar arbitragem em decorrência deste lapso de tempo que vai da constatação do movimento na bolsa externa até a incorporação do efeito da informação no preço do ativo local. E, se é possível fazer arbitragem, o mercado não é, definitivamente, eficiente.

Pretendeu-se, neste trabalho, testar a Hipótese do Mercado Eficiente, abordando a inter-relação entre dois mercados, com o fim de agregar evidências que possam colaborar com futuros trabalhos acerca do tema.

Com relação à arbitragem, cabe destacar que os custos de corretagem cobrados pelas corretoras, além dos custos de liquidação e emolumentos cobrados pela Bovespa, tornam, aparentemente, inviáveis as negociações com fins de obtenção de lucros. Os retornos dos ativos para períodos inferiores a dez minutos são, em média, inferiores aos custos de transação, inviabilizando o negócio.

REFERÊNCIAS

ASSAF NETO, A.; LIMA F. G.; DE ARAÚJO, A. M. P. **A realidade da concentração do capital no Brasil: Um estudo comparativo com duas outras economias emergentes.**

- Available at <http://www.fipecafi.com.br/palestras/assaf/pdf/Artigo_Concentracao_Capital_Brasil_Dividendos.pdf>. Accessed on February 20, 2008.
- BOVESPA– <www.bovespa.com.br>. Accessed on March 20, 2008.
- BROOKS, C. **Introductory Econometrics for Finance**. 1st Edition, Cambridge, Cambridge University Press. 2002.
- BROOKS, C. REW, A. G.; RITSON, S. A Trading Strategy Based on the Lead-lag Relationship Between the Spot Index and Futures Contract for the FTSE 100. **International Journal of Forecasting**, V. 17, 2001.
- CAMPBELL, J. Y., LO, A.W.; MACKINLAY, A. C. **The Econometrics of Financial Markets**. Princeton University Press. New Jersey. 1997.
- CARNEIRO, FRANCISCO GALRÃO. **A Metodologia dos Testes de Causalidade em Economia**. Available at < <http://www.unb.br/face/eco/textos/causal.pdf>> Accessed on January 10, 2008.
- CHAN, K. A Further Analysis of the Lead--Lag Relationship Between the Cash Market and Stock Index Futures Market. **The Review of Financial Studies**, Vol. 5, No. 1 (1992), pp. 123-152
- COSTA JUNIOR, N.C.A.; LEAL, R.P.C. Mercosul e a Globalização dos Mercados de Capitais: Testes de Causalidade. IN: **Revista de Administração**, vol. 32, 1997.
- CUNHA, M. S. **Raiz Unitária e Co-integração: Três Aplicações**. Available at <<http://www.unb.br/face/eco/textos/cointegracao.pdf>> Accessed on September 30, 2007.
- ENGLE, R. F.; GRANGER, C. W. E. Cointegration and error correction: Representation, estimation and testing. **Econometrica**, V.55. 1987.
- FAMA, E. F. Efficient capital markets: A review of empirical work, **Journal of Finance**, 417. 1970.
- GRANGER, C. W. J. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. **Econometrica**, V. 37, 1969.
- GROPPO, G. S.; BACCHI, M. R. P. Causalidade entre mercados acionários: estudo sobre Ibovespa, Dow Jones e Merval. In: **Resenha BM&F** n. 166. 2005.
- GUJARATI, D. N. **Econometria Básica**. São Paulo. Makion Books. 2000.
- HARRIS, R. I. D. **Using Cointegration Analysis in Econometric Modelling**. London: Prentice Hall, 1995.
- HSIAO, C. Statistical Properties of the Two-Stage Least Squares Estimator under Cointegration. **Review of Economic Studies**, Vol. 64, n. 3. 1997
- JONG, F.; NIJMAN, T. High frequency analysis of lead-lag relationships between financial markets. **Journal of Empirical Finance**, Volume 4, Issues 2-3, June 1997, pages 259-277.
- KAWALLER, I.G.; KOCH, P.D.; KOCH, T.W. The Temporal Price Relationship Between S&P 500 Futures and the S&P 500 Index. **The Journal of Finance**, Vol. 42, No. 5 (Dec., 1987), pp. 1309-1329.
- KOFMAN, P.; MARTENS, M. Interaction between stock markets: an analysis of the common trading hours at the London and New York stock exchange. **Journal of International Money and Finance**, Volume 16, Issue 3, June 1997, pages 387-414.

LI, Y.; GRECO, J.F.; CHAVIS, B. **Lead-lag relations between A shares and H shares in the Chinese stock markets**. California State University of Fullerton Working Paper, 2001.

POSHAKWALE, S.; THEOBALD, M. Market capitalisation, cross-correlations, the lead/lag structure and microstructure effects in the Indian stock market. **Journal of International Financial Markets, Institutions and Money**, Volume 14, Issue 4, October 2004, Pages 385-400.

SARTORIS, A. **Estatística e Introdução à Econometria**. São Paulo: Saraiva, 2007.

SIMS, C. A. Macroeconomics and Reality. **Econometrica**, n. 48. 1980.