



Revista Brasileira de Finanças

ISSN: 1679-0731

rbfin@fgv.br

Sociedade Brasileira de Finanças

Brasil

Ferreira Fonseca, Nelson; Moura Lamounier, Wagner; Bressan, Aureliano Angel  
Retornos Anormais no Ibovespa Utilizando Modelos para Dados de Alta Frequência  
Revista Brasileira de Finanças, vol. 10, núm. 2, 2012, pp. 243-265

Sociedade Brasileira de Finanças

Rio de Janeiro, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=305824777004>

- ▶ Como citar este artigo
- ▶ Número completo
- ▶ Mais artigos
- ▶ Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica

Rede de Revistas Científicas da América Latina, Caribe , Espanha e Portugal  
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

# **Retornos Anormais no Ibovespa Utilizando Modelos para Dados de Alta Frequência**

**(Abnormal returns in the Ibovespa Using Models for High-Frequency Data)**

**Nelson Ferreira Fonseca\***

**Wagner Moura Lamounier\*\***

**Aureliano Angel Bressan\*\*\***

## **Resumo**

Este artigo tem como objetivo identificar estratégias de negociação lucrativas com base nos efeitos de liderança e na defasagem entre os mercados acionários à vista e futuro no Brasil, utilizando dados de alta frequência. Para alcançar esse objetivo e com base nos dados históricos do índice Bovespa e do índice Bovespa Futuro construíram-se quatro modelos econôméticos de previsão: ARIMA, ARFIMA, VAR e VECM. As estratégias de negociação testadas foram: estratégia de negociação líquida, estratégia de compra e manutenção da posição e estratégia de filtro com a média. O período de análise desse artigo estende-se de 1 de agosto de 2006 a 16 de outubro de 2009. Neste trabalho, foi possível obter retornos anormais com a utilização de estratégias de negociação com o modelo VAR sobre os efeitos de liderança e defasagem entre o índice Bovespa e o índice Bovespa Futuro.

**Palavras-chave:** estratégias de negociação; dados de alta frequência; IBOVESPA.

**JEL code:** G14.

## **Abstract**

This article aims to identify profitable trading strategies based on the effects of leads and lags between the spot and futures equity markets in Brazil, using high frequency data. To achieve this objective and based on historical data of the Bovespa and the Bovespa Future indexes, four forecasting models have been built: ARIMA, ARFIMA, VAR, and VECM. The trading strategies tested were: net trading strategy, buy and hold strategy, and filter strategy - better than average predicted return. The period of analysis of this paper extends from August 1, 2006 to October 16, 2009. In this work, it was possible to obtain abnormal returns using trading strategies with the VAR model on the effects of leads and lags between the Bovespa index and Bovespa Future index.

**Keywords:** trading strategies; high-frequency data; IBOVESPA.

Submetido em 30 de outubro de 2011. Reformulado em 29 de fevereiro de 2012. Aceito em 21 de março de 2012. Publicado on-line em 25 de junho de 2012. O artigo foi avaliado segundo o processo de duplo anonimato além de ser avaliado pelo editor. Editor responsável: Márcio Laurini. Reprodução parcial ou total e trabalhos derivativos permitidos com a citação apropriada da fonte.

\*Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil. E-mail: nfonseca02@gmail.com

\*\*Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil. E-mail: wagner@face.ufmg.br

\*\*\*Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil. E-mail: bressan@face.ufmg.br

## 1. Introdução

O desejo dos agentes econômicos em prever os valores futuros dos ativos é inerente ao comportamento do ser humano. Desde a criação dos mercados organizados, os pesquisadores e participantes tentam descobrir maneiras de utilizar o histórico dos preços dos ativos para obter informações relevantes capazes de propiciar a obtenção de lucros por meio do uso dessas informações.

O surgimento do mercado organizado de derivativos fez surgir novos campos de pesquisa que estudam a capacidade do comportamento deste mercado auxiliar na previsão do comportamento futuro dos preços de seus respectivos ativos subjacentes.

A previsão dos valores dos ativos tem importância para as atividades de investimento, como *trading* e *hedging*, e para identificação de preços anômalos no mercado. Dessa maneira, o conhecimento da relação entre o preço à vista e o preço futuro permite que o mercado corrija possíveis falhas entre estes dois preços.

Investidores frequentemente assumem posições casadas no mercado à vista e no mercado futuro, sendo esta a motivação desse artigo: a relação entre os preços destes dois mercados.

A antecipação do índice Bovespa Futuro ocorre porque a negociação de um contrato futuro de índice refere-se ao fato de que em uma única operação o investidor pode manter posições diversificadas, líquidas e de baixo custo (Brooks *et al.*, 2001). Para obter a mesma posição sem um contrato futuro de índice, o investidor terá que buscar dezenas de ações individualmente, incorrendo em altos custos de operação e em uma baixa liquidez em algumas ações. Essa antecipação pode ser usada para prever os valores correntes do índice. Este artigo tem por objetivo identificar a existência de precedência do índice Bovespa Futuro sobre o índice Bovespa, de forma a possibilitar a previsibilidade do índice para que se possam criar estratégias lucrativas que possibilitem ganhos anormais. Entende-se como ganho anormal, o ganho acima de um *benchmark* escolhido. Para este estudo, o *benchmark* é a estratégia passiva.

A abordagem de análise no presente artigo visa analisar a possibilidade de ganhos com estratégias de negociação com dados de alta frequência. Ainda pouco discutidas academicamente no Brasil, esse tipo de operação vem crescendo cada vez mais no dia a dia das operações no mercado financeiro. Basicamente uma estratégia de negociação em alta frequência é caracterizada pela escolha e manutenção de portfólios por curtos períodos de tempo. Nessas escolhas e alocações geralmente há intensivo uso de métodos computacionais e estatísticos para processamento, previsões e definições de *timing* de entrada e saída do mercado, aproveitando-se de oportunidades momentâneas de arbitragens e de ganhos.

O uso de estratégias com dados de alta frequência teve sua origem em 1998 quando a SEC autorizou o uso de transações por meio eletrônico no mercado financeiro americano, e teve o seu crescimento elevado substancialmente em conjunto com a evolução na capacidade computacional dos processadores atuais e no desenvolvimento e aumento da velocidade dos meios de comunicação, particular-

mente no caso da internet, com transações globais e interligação instantânea dos mercados.

Segundo Grant (2010), as transações com dados de alta frequência se configuram como uma das maiores inovações recentes do mercado financeiro e envolvem atualmente mais de 56% das transações com ações nos Estados Unidos e 38% na Europa. As estratégias de negociação em alta frequência são motivadas principalmente pelo surgimento de possibilidades de ganhos com pequenas discrepâncias e desvios que os preços e retornos dos ativos, commodities e títulos em geral apresentam em curtos espaços de tempo em relação aos seus valores “justos”. A rápida identificação e operação quando do surgimento dessas oportunidades momentâneas pode permitir aos investidores ganhos expressivos.

## 2. Revisão da Literatura

A previsão do comportamento do preço à vista de um contrato de índice de ações insere-se em duas grandes abordagens distintas de pesquisa. A primeira implica a previsão do preço à vista com base no histórico de preços do mesmo índice, em que se utilizam modelos econômétricos univariados. A segunda implica a utilização dos preços dos contratos futuros dos índices de ações como variável explicativa junto com o histórico dos preços à vista para explicar o comportamento futuro dos preços à vista. A seguir, são revisados alguns dos principais estudos relacionados ao tema deste artigo, que, por sua vez, permeia estes dois enfoques de pesquisa.

Um importante trabalho que investiga a relação do mercado à vista e o mercado futuro norte americano é o trabalho de Stoll & Whaley (1990). Os autores examinaram a relação causal entre o retorno do índice S&P500 futuro e o retorno do índice S&P500 à vista, bem como a relação entre o retorno do índice MMI (Major Market Index) futuro e o retorno do índice MMI à vista. Outra relação investigada foi à existente entre o índice S&P500 futuro e o índice MMI futuro em relação às ações à vista da IBM. Os resultados indicaram que o índice S&P 500 futuro e o índice MMI futuro são antecipatórios em relação aos seus respectivos índices subjacentes em um intervalo de tempo médio de cinco minutos, e ocasionalmente, em um período de mais de 10 minutos. Os índices futuros também se mostram antecipatórios em relação à ação da IBM à vista. Outro importante resultado aponta que os índices à vista S&P500 e o MMI foram adequadamente modelados conforme o modelo univariado ARMA(2,3) e as ações da IBM por um modelo MA(3).

Considerando ainda o índice S&P500, Tse & Chan (2010) examinaram as interações de liderança e defasagem entre os mercados à vista e futuro do índice. Como metodologia, os autores utilizaram o *Threshold Regression Model* (TRM) para modelar as mudanças na estrutura da regressão em diferentes condições de mercado. Os resultados encontrados foram que o mercado futuro é fortemente antecipatório ao mercado à vista quando existe mais de uma variável explicativa no modelo, enquanto que quando existe mais de uma variável explicativa o mercado à vista é fracamente antecipatório ao mercado futuro.



Para o mercado financeiro asiático, Tse (1995) examinou o comportamento dos preços do índice Nikkei médio e o seu correspondente contrato futuro. O modelo de correção de erros foi utilizado para investigar a relação entre estas duas séries. O estudo encontrou que as mudanças nas defasagens do preço do contrato futuro afetam os ajustes de curto prazo do preço futuro do ativo subjacente. Porém, o inverso não foi observado. O estudo ainda comparou o modelo de correção de erro com o modelo univariado autoregressivo (AR) ajustado para a série do índice à vista. O modelo de correção de erro apresentou um desempenho preditivo melhor que o modelo AR, enquanto que o modelo AR apresentou um desempenho melhor que o modelo *martingale*.

Ainda em relação ao índice Nikkei de ações médio, Lien & Tse (1999) examinaram a performance preditiva do modelo de correção de erro fracionário e integrado (*Fractionally integrated error correction model*) e compararam com o modelo *martingale*, modelo de vetor autoregressivo e o convencional modelo de correção de erro. Neste estudo, foram considerados modelos com e sem heterocedasticidade. Os resultados mostraram que para um horizonte de 20 dias o modelo com melhor desempenho preditivo foi o modelo de correção de erro fracionário e integrado com heterocedasticidade. Estes resultados reafirmaram a noção de que a cointegração e a cointegração fracionária são importantes para previsões de horizontes de longa duração.

No mercado financeiro do Reino Unido, existem vários trabalhos que investigaram as relações de liderança e defasagem das cotações do índice à vista FTSE 100 e do seu respectivo índice futuro. Wahab & Lashgari (1993) investigaram a relação entre os preços do índice FTSE100 futuro e os preços do ativo subjacente. Para esta investigação, os autores usaram como metodologia o modelo de correção de erros (VEC). Os resultados revelaram que existem efeitos antecipatórios tanto do preço futuro em relação ao preço à vista quanto o contrário. Porém, os resultados mais significativos foram sobre os efeitos do preço futuro sobre o preço à vista, como era de se esperar pela literatura precedente.

Outro trabalho para o mercado financeiro do Reino Unido é o do Brooks *et al.* (2001), cujo objetivo é examinar as relações de liderança e defasagem entre o índice FTSE100 e o índice de preço futuro deste mesmo índice. Os autores utilizaram quatro metodologias para verificar a relação entre os preços à vista e os preços futuros: modelo ARIMA; modelo VEC; modelo VAR; e modelo VEC com o modelo teórico COC (*cost of carry theory*). O melhor modelo de previsão encontrado foi o VEC (correção de erros) associado com o teórico (COC). Em todos os modelos ajustados foi verificado que as mudanças nas defasagens do preço futuro podem ajudar a prever as mudanças no preço à vista. A habilidade de previsão do melhor modelo foi então utilizada para encontrar estratégias de negociação, as quais por sua vez, foram testadas sobre condições reais para encontrar oportunidades de negociações lucrativas. Os resultados mostraram que o modelo apresentou retornos significativamente altos quando comparado a um modelo passivo. Porém, quando se consideraram os custos da transação, os retornos do modelo não

conseguiram superar mais este modelo passivo. Este modelo consiste em um investimento no índice no instante inicial e a venda dele no final do período, não realizando mais nenhuma outra transação.

Para o mercado financeiro brasileiro, Galvão *et al.* (2000) investigaram a relação entre o mercado futuro e o mercado à vista com base no comportamento da volatilidade destes mercados para poder inferir sobre as relações de causalidade. Os resultados apontaram que a transmissão de volatilidade e de novas informações entre o mercado futuro e o mercado à vista do índice Bovespa tem origem no mercado à vista. Com isso, o estudo argumenta que o mercado futuro tem funções econômicas importantes de previsibilidade de preços e repartição de riscos.

Ainda em relação ao mercado brasileiro, Silva (2006) investigou a relação entre o índice à vista Bovespa e o índice futuro Bovespa. A metodologia usada neste foi a de cointegração de Johansen. Os resultados encontrados mostraram que o mercado futuro brasileiro é antecipatório em relação ao mercado à vista.

### **3. Referencial Teórico**

#### **3.1 Eficiência de Mercado**

O mercado de capitais pode ser entendido como um sistema de distribuição de valores mobiliários que proporciona liquidez aos títulos de emissão de empresas e viabiliza o processo de capitalização. Este mercado tem como principal função a alocação da propriedade do suprimento de capital da economia e os participantes deste mercado interagem entre si e reagem as informações com o objetivo de determinar o melhor preço para um determinado ativo (Oliveira, 2008). A interação entre os agentes forma então um complexo sistema de formação de preços dos ativos que acontece através da incorporação das informações disponíveis na série temporal de preços.

A incorporação das informações disponíveis na série de preços está associada à Teoria de Eficiência de Mercado. Esta teoria pode ser atribuída de maneira pioneira ao trabalho de Bachelier (1964) e a pesquisa empírica de Cowles (1933).

A Teoria de Eficiência de Mercado fez surgir a Hipótese de Mercado Eficiente (HME). Esta hipótese pode ser atribuída então ao trabalho de Bachelier (1964) e a pesquisa empírica de Cowles (1933). Basicamente a hipótese afirma que se o mercado é eficiente às mudanças dos preços não poderão ser previstas, já que todas as informações estarão contidas nos preços. Com isso não existe possibilidade de se obter lucros anormais em mercados eficientes (Oliveira, 2008).

O mercado financeiro é dito eficiente se os preços dos títulos refletem integralmente a todas as informações disponíveis. Sendo essas informações passadas ou que aquelas que ainda vão acontecer. E ainda em um mercado eficiente os preços se ajustam rapidamente quando surgirem novas informações relevantes que não foram previstas (Haugen, 1997). Fama (1970) por sua vez distinguiu três formas de eficiência de mercado: Fraca, Semiforte e Forte.

A confirmação da hipótese dos mercados eficientes traz consequências para a



administração de carteiras. Conforme sugere Elton *et al.* (2004), se a forma de eficiência na Forma Fraca estiver correta, então táticas de negociação baseadas em preços passados tais como aquelas preconizadas pela análise técnica seriam inúteis. Se a Forma Semiforte for válida, então as estratégias de negociações baseadas em informações publicamente disponíveis são suspeitas. Finalmente, se a validade da Forma Forte for confirmada pelos estudos empíricos, a própria utilidade da análise de títulos é colocada em dúvida.

A principal conclusão da Hipótese do Mercado Eficiente é que os preços dos ativos são imprevisíveis porque seguem um passeio aleatório e que com isso é impossível que algum agente do mercado consiga obter ganhos acima do mercado, tenha retornos anormais (Oliveira, 2008). Dessa maneira apenas o valor atual da variável é relevante para a previsão de valores futuros, já que no processo de passeio aleatório os incrementos ao valor atual da variável são aleatórios.

### 3.2 Formação de preços futuros

Com base nos trabalhos de Kawaller *et al.* (1987), Stoll & Whaley (1990), Lien & Tse (1999) e Brooks *et al.* (2001), é possível apresentar o processo de formação de preços futuros em consonância com a abordagem de alta frequência utilizada no presente estudo. Nesta perspectiva, a possibilidade de arbitragem existente entre o preço à vista e o preço futuro é determinante no processo de formação dos preços futuros, pois se a relação for rompida existe a possibilidade de auferir ganhos sem risco (Lien & Tse, 1999). Então, se os mercados à vista e futuro são eficientes e livres de impedimentos e de período contínuo, tem-se a seguinte relação entre o preço do contrato futuro de um índice e o nível de preço do índice subjacente (Stoll & Whaley, 1990):

$$F_t = S_t e^{(r-d)(T-t)}$$

em que  $F_t$  é o preço do contrato futuro de um índice no tempo  $t$ ;  $S_t$  é o nível de preço do índice subjacente no tempo  $t$ ;  $r - d$  é o custo líquido do carregamento do índice de ações subjacente ao contrato futuro até a data do vencimento deste contrato, em que  $r$  é o custo da taxa de juros, dado pela taxa de retorno contínuo do ativo livre de risco, e  $d$  é o rendimento dos dividendos (*dividend yield*) da carteira de ações. Este custo está associado com a compra do ativo à vista e a manutenção do ativo subjacente ao contrato futuro até o vencimento (Lien & Tse, 1999).  $T$  é o tempo de vencimento do contrato futuro. Então,  $T - t$  é o tempo de vida remanescente do contrato futuro. Na equação, a taxa de juros livre de risco,  $r$ , e o dividend yield,  $d$ , do índice de ações subjacente ao contrato futuro são taxas conhecidas a priori, com valores constantes (Stoll & Whaley, 1990).

Transformando a relação anterior em um modelo de log-retorno, em vez do nível da série, tem-se a seguinte relação:

$$R_{S,t} = (r - d) + R_{(F,t)}$$

em que  $R_{S,t} = \ln(S_t/S_{t-1})$ ; e  $R_{F,t} = \ln(F_t/F_{t-1})$ . A equação implica que, sob a HME e na ausência de atritos no mercado, os retornos do mercado futuro e à vista devem ser perfeitamente correlacionados contemporaneamente e não podem ser correlacionados de maneira cruzada ao longo do tempo. Isto é, os retornos destes dois mercados não podem ser correlacionados em pontos diferentes do tempo a não ser no mesmo instante. Com isso, nenhum destes mercados pode ser antecipatório ao outro. Brooks *et al.* (2001) destacam que esta condição implica que os preços do índice à vista e futuro simultaneamente refletem toda informação nova anunciada ao mercado, e o eventual desequilíbrio da primeira relação geraria oportunidades de arbitragem.

A segunda relação tem como suposição que a taxa de juros livre de risco de curto prazo,  $r$ , e a taxa de *dividend yield*,  $d$ , são constantes e o mercado futuro e à vista são eficientes e o tempo é contínuo. Estas suposições possuem diversas implicações que são baseadas na suposição de que a relação do custo de carregamento é mantida por todo o tempo (Stoll & Whaley, 1990). Porém, muitos estudos mostraram que esta relação é rompida em diferentes mercados financeiros do mundo em pontos distintos do tempo, como pode ser observado nos trabalhos de Tse & Chan (2010) para o mercado americano e Silva (2006) para o mercado brasileiro. O rompimento da relação cria oportunidades de arbitragem. Este rompimento é observado na literatura apenas para os dados intradiários, tal como evidenciado no trabalho de Brooks *et al.* (2001). Com este rompimento criam-se as relações de liderança (*lead*) e defasagem (*lag*) entre o mercado à vista e o mercado futuro.

Com isso, as negociações poderão, primeiro, mudar os preços futuros e então antecipar o movimento de preço do índice de ações enquanto a operação de arbitragem não corrige os desvios do modelo do custo de carregamento. Preços futuros exigem um indicador de mudanças nos preços das ações. Então, este é seguido por aqueles investidores que não investem em futuros ou que não têm condições. Com isso, incorporam as informações em suas carteiras das informações oriundas do mercado futuro (Kawaller *et al.*, 1987).

Brooks *et al.* (2001) destacam que as mudanças no índice de ações podem também antecipar as mudanças no preço futuro, já que o índice pode representar um conjunto menor de informações que afetam o preço futuro. Com isso, existe a possibilidade de o mercado futuro antecipar o mercado à vista. Este mercado antecipa os movimentos do mercado futuro, o que provoca no modelo do custo de carregamento uma relação de dupla causalidade. Com isso, a base (diferença absoluta entre o preço futuro e o preço à vista subjacente) deve estar inserida em um intervalo em que não existe a possibilidade de arbitragem. Mudanças no mercado poderão afetar tanto o mercado futuro como o mercado à vista na mesma direção. Este intervalo pode ser definido para os preços futuros e para os preços do índice no tempo  $t$  como (Kawaller *et al.*, 1987):

$$e_{L,t} < (F_t - S_t) < e_{U,t}$$



em que  $e_{L,t}$  é o limite inferior do intervalo de negociação em que não existe a possibilidade de arbitragem no tempo  $t$ ; e  $e_{U,t}$  é o limite superior do intervalo de não arbitragem no tempo  $t$ . Nos momentos em que a base estiver fora deste intervalo haverá possibilidades de arbitragem. Com isso, será possível realizar operações sem risco, em que o lucro será a diferença do preço que estiver mais alto em relação ao preço que estiver mais baixo (Brooks *et al.*, 2001).

Em síntese, o modelo do custo de carregamento, normalmente, é violado, e as discrepâncias são explicadas pelos custos de transação, pela falta de liquidez de algumas ações que compõem o índice e pela defasagem de tempo no cálculo do índice (Brooks *et al.*, 2001).

#### **4. Metodologia**

A metodologia empregada neste trabalho baseia-se na construção de modelos econométricos univariados e multivariados de previsão de preços com suporte em dados históricos de séries temporais. Existe uma diversidade grande destes modelos. Para os fins desta pesquisa, optou-se por selecionar os modelos univariados ARIMA e ARFIMA, e os modelos multivariados VAR e VECM Sendo que para este artigo o objetivo é modelar o comportamento do índice Bovespa com base apenas nas suas defasagens nos modelos univariados; e para os modelos multivariados, nas defasagens do índice e do índice Bovespa Futuro.

As variáveis utilizadas são as cotações de fechamento do índice Bovespa à vista e do índice Bovespa Futuro, sendo que o valor deste é em relação ao 1º vencimento dos contratos. Estas variáveis foram obtidas no sistema de cotação em tempo real de ativos da Consultoria, Métodos, Assessoria e Mercantil S/A (CMA), com frequência de quinze minutos.

O período de análise dessa pesquisa estende-se de 1 de agosto de 2006 a 16 de outubro de 2009, consistindo em 26.947 observações. O período para estimar os modelos estende-se de 1 de agosto de 2006 a 30 de dezembro de 2008, consistindo em 20.451 observações. O período de previsão foi de 2 de janeiro de 2009 a 16 de outubro de 2009, consistindo em 6.496 observações. Analisou-se também o período que se estendeu de 1 de agosto de 2006 a 29 de agosto de 2008 para a estimativa dos modelos e o período de previsão estendeu-se de 01 de setembro de 2008 a 16 de outubro de 2009. Esta segunda análise tem por objetivo testar os modelos de previsão no período da crise de 2008.

## 4.1 Modelo ARIMA

A ideia geral deste modelo é transformar uma série não estacionária em estacionária por meio de  $d$  diferenciações. Em seguida, são inseridos na série componentes auto-regressivos e componentes média móvel.

O processo de modelagem ARIMA( $p, d, q$ ) foi proposto por Box & Pierce (1970). Acontece em três fases: a) identificação/seleção do modelo; b) estimação; e c) diagnóstico ou verificação. No presente estudo, o software R<sup>1</sup> é utilizado na estimação do modelo ARIMA. Com os parâmetros do modelo estimado, deve-se verificar a estatística de Ljung-Box  $Q(k)$  dos resíduos para confirmar se o modelo está bem ajustado aos dados.

Conforme Tsay (2005), a previsão  $m$  passos à frente pode ser feita mediante o uso da equação representativa do modelo ARIMA( $p, d, q$ ), utilizando, quando necessário, a esperança matemática das previsões como subsídio para as outras novas previsões.

## 4.2 Modelo ARFIMA

Um modelo ARFIMA (*AutoRegressive Fractional Integrated Moving Average*) é utilizado quando a série apresenta autocorrelação significativa em intervalos muitos. Em um processo ARFIMA( $p, d, q$ ), as autocorrelações, excluindo a defasagem zero, decaem hiperbolicamente.

A especificação de um modelo ARFIMA( $p, d, q$ ) é definida da seguinte maneira:

$$\phi_p(B)(1 - B)^d y_t = \theta_0 + C\epsilon_t$$

As raízes dos polinômios  $\phi_p(B)$  e  $\theta_q(B)$  estão fora do círculo unitário, que, por sua vez, faz com que o processo  $(1 - B)^d y_t$  seja estacionário e invertível, e  $d \in (-0.5, 0.5)$  é um número real ao contrário dos modelos ARIMA( $p, d, q$ ) usuais, em que  $d$  é um número inteiro. O filtro  $(1 - B)^d$  pode ser expandido pelo binômio de Newton ou em uma série de Taylor.

Se  $d = 0$ ,  $y_t$  é um processo autoregressivo médias móveis ARMA( $p, q$ ). Se  $d \neq 0$  e  $d \in \mathbb{Z}$ ,  $y_t$  é um processo autoregressivo médias móveis integrado ARIMA( $p, d, q$ ). Por último, se  $d \neq 0$  e  $d \in \mathbb{R}$ ,  $y_t$  é um processo autoregressivo médias móveis fracionalmente integrado ARFIMA( $p, d, q$ ), em que a função de autocorrelação  $\rho(k)$  tem um decaimento hiperbólico,  $\rho(k) \sim e^{-|k|^{2d-1}}$  com  $|k| \rightarrow \infty$ , enquanto que as autocorrelações de um processo ARMA( $p, q$ ) têm um decaimento exponencial  $\rho_k \sim ca^k$ ,  $0 < a < 1$ .

O teste  $R/S$  permite identificar se a série apresenta uma longa dependência temporal. A estatística  $R/S$ , representada por  $Q_n$ , está em uma faixa de somas parciais dos desvios em relação à média da série temporal, redimensionada pelo seu desvio padrão.

---

<sup>1</sup>Disponível em <http://www.r-project.org/>.



Para as séries financeiras, Peters (1994) demonstra que para  $Q_n$  maior que 0,5 a série é caracterizada como uma série de memória longa, sendo a hipótese nula dada como  $H_0 : Q_n = 0$  e a hipótese alternativa  $H_A : Q_n \neq 0$ .

Geweke & Porter-Hudak (1983) propõem um método, também conhecido como “método GPH”, de estimação não paramétrica do parâmetro  $d$ . Este método permite testar a presença de memória longa e obter estimativas preliminares do parâmetro de memória  $d$  e dos polinômios  $\phi_p(B)$  e  $\theta_q(B)$ . A vantagem deste método está em poder estimar o parâmetro  $d$  sem explicitar especificações dos parâmetros de curto prazo da série, sendo estes parâmetros especificados pelo modelo ARMA( $p, q$ ).

A metodologia de Geweke & Porter-Hudak (1983) de estimação do parâmetro  $d$  é feita em dois estágios: O primeiro é a estimação de  $\tilde{d}$ . O segundo consiste em aplicar o filtro apresentado na equação (XX) com  $\tilde{d}$  no lugar de  $d$  truncado no tamanho da amostra e aplicar os procedimentos usuais de estimação dos modelos ARMA( $p, q$ ).

#### 4.3 Modelo VAR

O modelo VAR (Vector AutoRegression) é um modelo econométrico utilizado com o objetivo de capturar a evolução das relações de interdependência de múltiplas séries temporais. Pode ser entendido como uma generalização do modelo univariado AR(AutoRegressive). Todas as variáveis no VAR são tratadas simetricamente. Com isso, cada variável possui uma equação explicativa, em que são incluídas nesta as defasagens da própria variável e as defasagens das outras variáveis (Enders, 2004).

Outro aspecto importante a se observar no modelo VAR é sobre a causalidade das variáveis envolvidas no modelo. Com esse objetivo, Granger (1969) desenvolve um teste que tem por objetivo verificar o sentido da causalidade entre as séries.

#### 4.4 Modelo VECM

O modelo VECM (*Vector Error Correction Model*), consiste em uma evolução do VAR, já que este pressupõe que as variáveis sejam estacionárias.

Para entender melhor o VECM, é necessário que se apresente antes o conceito de cointegração. Este conceito implica que a regressão entre duas variáveis que não são estacionárias e são cointegradas não pode ser considerada espúria, e os resultados dessa regressão passam a fazer sentido. Nesse sentido, a teoria de cointegração tem dois pontos centrais: testar os resíduos,  $u_t$ , da regressão entre variáveis não estacionárias e integradas de ordem ( $d$ ), para saber se é uma variável estacionária. Se o ponto anterior for verdadeiro, os resíduos devem ser inseridos na equação para ajustar melhor o VAR. Com essa informação inserida no VAR, o modelo passa a ser chamado de VECM, em que se incorpora o erro de equilíbrio do curto prazo (Enders, 2004).

As relações de equilíbrio econômico entre as variáveis não estacionárias im-

plicam a existência de uma combinação linear que retira a tendência estocástica e faz com que essas variáveis não movam independentemente uma das outras. Isso significa que a combinação das variáveis não estacionárias resulta em uma série estacionária.

As previsões de um passo à frente são semelhantes ao VAR, com a diferença de que no VECM é inserido o termo de correção de erro na equação característica do modelo. A próxima seção apresenta os critérios para a avaliação das previsões dos modelos apresentados nas seções anteriores.

#### 4.5 Critérios para a avaliação de desempenho dos modelos

Os modelos univariados e multivariados apresentados são construídos com o objetivo de captar o comportamento dos índices ao longo do período analisado e, com isso, poder calcular as previsões *ex-post* um passo à frente.

Os modelos são estimados por toda a amostra de estimação e verificados se os resultados encontrados são válidos. Depois disso, é necessário medir o desempenho destes modelos, o que será feito por meio dos critérios: estatístico e operacional. O primeiro é feito conforme a estatística para a avaliação dos modelos de previsão; o segundo, por estratégias de negociação baseadas nas previsões estatísticas.

##### Estatística para a avaliação dos modelos de previsão

Para medir o desempenho dos modelos de previsão, o critério estatístico utilizado foi o erro médio percentual absoluto (*Mean Absolute Percentage Error – MAPE*), dado por:

$$MAPE = \frac{1}{n_f} \sum_{i=1}^{n_f} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (1)$$

em que  $n_f$  significa o número de observações da amostra usada para a previsão,  $y_i$  os valores observados e os termos  $\hat{y}_i$  são os valores preditos. O modelo que apresentar o menor valor desta estatística é aquele que gera as previsões mais próximas do valor observado, ou seja, com menor erro percentual médio.

##### Estratégias de negociação baseada nas previsões estatísticas

O período de negociação das estratégias é o mesmo usado para as previsões. Assume-se que o investimento original é de 1.000 unidades monetárias. A estratégia passiva é usada como referência de desempenho das outras estratégias.

As estratégias de investimento serão testadas considerando o valor dos custos de transação e o efeito destes quando estão presentes na negociação e quando não estão dentro da negociação.



### **Estratégia de negociação líquida (Liquid Trading Strategy – LTS)**

A estratégia LTS envolve a negociação de compra e venda a cada intervalo de quinze minutos sempre que o retorno predito pelo modelo for positivo. Caso contrário, retorno predito negativo, não haverá negociação, e o dinheiro permanecerá na carteira.

### **Estratégia de compra e manutenção da posição (Buy and Hold Strategy – BHS)**

A estratégia BHS tem por objetivo reduzir os custos de transação, a partir da diminuição do volume de transações. Com isso, o investidor tem a possibilidade de manter a sua posição no índice se o retorno predito for positivo no próximo período. Com isso, o investidor carrega a posição em aberto até que os retornos preditos se tornem negativos.

### **Estratégia de filtro com a média (Filter Strategy - better predicted return than average - MFS)**

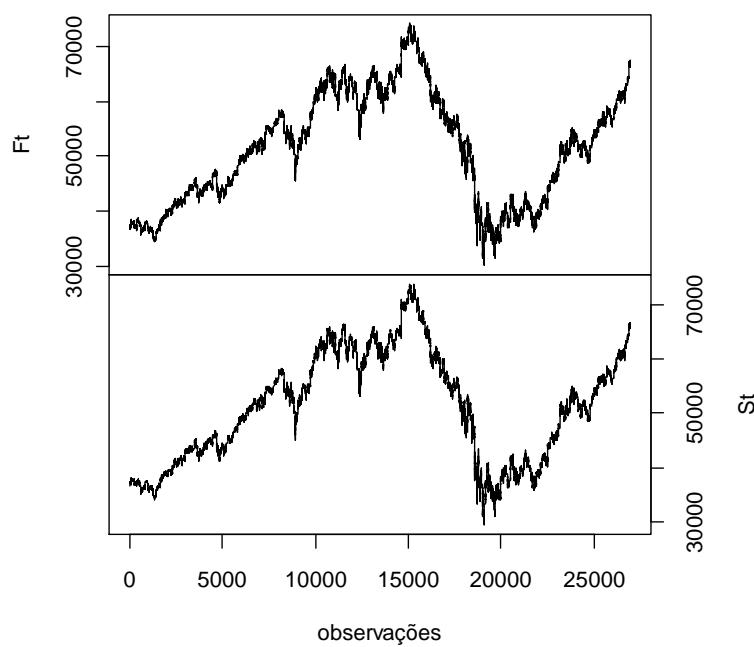
A estratégia MFS envolve a compra do índice se o retorno predito for maior que a média dos retornos preditos positivos. Dessa maneira, os ganhos podem se tornar mais elevados, ao filtrarem apenas negociações de alto potencial de ganhos. Esta estratégia se diferencia das outras por ser mais seletiva na negociação. O investidor mantém a sua posição se o retorno predito for maior que a média dos retornos preditos positivos. Caso contrário, ele liquida a posição. Estratégia semelhante à estratégia BHS, porém usando a média dos retornos preditos positivos como referência ao invés de usar o sinal do retorno predito.

## **5. Apresentação e Análise dos Resultados**

Com base nos dados históricos do índice Bovespa e do índice Bovespa futuro obtidos da CMA, construíram-se quatro modelos econometríticos, cuja metodologia foi apresentada no capítulo anterior. Com os modelos ajustados, calcularam-se as estatísticas de avaliação dos modelos de previsão. Em seguida, estes foram usados com a intenção de construir estratégias lucrativas de negociação.

A série do índice Bovespa é denominada  $St$  quando está sendo apresentada no nível;  $\ln St$ , quando é o logaritmo natural da série  $St$ ; e  $st$ , quando é o retorno logarítmico da série. Para a série do índice Bovespa Futuro, a denominação é  $Ft$  para o nível;  $\ln Ft$ , quando é o logaritmo natural da série  $Ft$ ; e  $ft$ , para o retorno logarítmico da série.

A análise descritiva, ou exploratória, das séries se inicia com o exame gráfico dos valores destas ao longo do tempo (Figura 1).



**Fonte:** elaborado pelos autores

**Figura 1**

Cotações dos índices Bovespa ( $S_t$ ) e Bovespa futuro ( $F_t$ ) do período de 1 de agosto de 2006 a 16 de outubro de 2009 – cotações a cada 15 minutos

Conforme se pode observar na Figura 1, as séries não apresentam ser estacionárias na média. Outro fato a se observar é que a série  $F_t$  apresenta um comportamento semelhante ao da série  $S_t$  por todo o período, como pode ser comprovado pela correlação de 0,9993.

A próxima etapa consiste em tomar a diferença das duas séries logaritmizadas, para que estas possam se tornar estacionárias. Em todos os testes de raiz unitária ADF rejeita-se a hipótese nula de que as séries em diferença possuem raiz unitária. Com o objetivo de confirmar os resultados obtidos pelo teste ADF, empregou-se o teste de raiz unitária KPSS, que tem como hipótese nula a estacionariedade da série. Pelos testes KPSS não se rejeita a hipótese nula de que as séries são estacionárias. Dessa forma, os resultados obtidos pelo teste ADF são confirmados pelo teste KPSS. Assim, as séries  $S_t$  e  $F_t$  são integradas de ordem 1.

O primeiro modelo estimado é o modelo ARIMA e a melhor ordem encontrada para o primeiro período e o segundo período foi  $p = 2$  e  $q = 2$  sem o intercepto, considerando o menor valor do AIC.

Depois de identificados os parâmetros do modelo ARIMA, a próxima fase consiste na estimação destes e na verificação se o modelo está bem ajustado aos dados



por meio da análise dos resíduos. Todos os coeficientes estimados são significativos a um nível de significância de 5% e os resíduos dos modelos não são autocorrelacionados.

Outra propriedade dos estimadores de mínimos quadrados é a de que os erros seguem uma distribuição normal. A hipótese nula de que os resíduos são distribuídos normalmente é rejeitada. Porém, a violação desta premissa não apresenta nenhuma consequência importante para amostras grandes, como as usadas neste trabalho (Brooks *et al.*, 2001).

O segundo modelo estimado é o modelo ARFIMA e a principal característica dos modelos ARFIMA é que a variável de interesse apresente autocorrelações significativas em intervalos muito distantes no tempo e isto ocorre para a série  $St$ . Antes de estimar o parâmetro  $d$ , deve-se fazer um teste inicial na série para saber se esta apresenta uma longa dependência temporal. Para o período de 1 de agosto de 2006 a 16 de outubro de 2009 o resultado da estatística  $R/S$  é que o  $H$  estimado é igual a 1,0228 com um valor-p de 0,0000 e um erro padrão de 0,0269. Desta maneira a série  $St$  apresenta uma memória longa. Rejeita-se a hipótese nula de que a estatística  $H$  seja igual a zero. Peters (1994) demonstra que quando a estatística  $H$  de uma série financeira é maior que 0,5 a série apresenta uma longa dependência temporal.

Com isso, o próximo passo é estimar o parâmetro  $d$ , em que, neste caso, o parâmetro é um número não inteiro. O parâmetro  $d$  é estimado pelo método de Geweke & Porter-Hudak (1983) também conhecido como “método GPH” para os dois períodos de análise e este também permite testar a presença de memória longa na série, um teste complementar ao teste  $R/S$  apresentado acima. Sendo para o primeiro período o valor de  $d$  é igual a 0,921866 e para o segundo é igual a 0,9647304.

Com a série diferenciada, os procedimentos de estimação do modelo são os procedimentos usuais de estimação dos modelos ARMA( $p, q$ ). A melhor ordem encontrada para o modelo do primeiro período foi  $p = 2$  e  $q = 2$  sem o intercepto e do segundo período foi  $p = 2$  e  $q = 1$  sem o intercepto.

Depois de identificados os parâmetros do modelo, a próxima fase será proceder à estimação destes e à verificação se o modelo está bem ajustado aos dados, por meio da análise dos resíduos.

Os resíduos dos modelos não são autocorrelacionados, o que faz com que os modelos estejam bem ajustados aos dados. Outra propriedade dos estimadores de mínimos quadrados é a de que os erros seguem uma distribuição normal. A hipótese nula de que os resíduos são distribuídos normalmente é rejeitada. Porém, a violação desta premissa não apresenta nenhuma consequência importante para amostras grandes, como as usadas neste trabalho (Brooks *et al.*, 2001).

O terceiro modelo estimado é o modelo VAR que tem como objetivo confirmar a relação existente entre os valores correntes do índice Bovespa com os valores prévios do próprio índice e do índice Bovespa futuro.

A primeira condição para estimar um modelo VAR é que as variáveis presentes

no modelo sejam estacionárias. No início deste capítulo, mostrou-se que as séries  $St$  e  $Ft$  para os dois períodos analisados não são estacionárias, mas que o logaritmo do retorno destas,  $st$  e  $ft$ , respectivamente, é estacionário.

A melhor ordem encontrada para o modelo do primeiro período foi  $p = 27$  sem o intercepto e do segundo período foi  $p = 13$  sem o intercepto. Pelos resultados encontrados, pode-se perceber que a equação explicativa de  $ft$  apresenta muitos coeficientes estatisticamente não significativos. Porém, na equação de  $st$  todos os coeficientes são significativos, com um nível de confiança de 95%. Os valores destes coeficientes indicam que a existência da relação de liderança e defasagem entre o índice Bovespa e o índice Bovespa futuro não pode ser rejeitada. Com isso, o índice Bovespa futuro e as defasagens do índice Bovespa podem prever os valores correntes do índice Bovespa, conforme a literatura já havia constatado para outros mercados financeiros.

O teste de causalidade de Granger foi utilizado com o objetivo de confirmar a forte relação entre as duas variáveis. A significância das relações é dada pela estatística  $F$ , que testa a hipótese de que os coeficientes das variáveis das equações são conjuntamente nulos. A Tabela 1 apresenta às estatísticas  $F$  e o valor-p para os dois períodos de análise.

**Tabela 1**

Resultado do teste de causalidade de Granger para as séries  $st$  e  $ft$  – cotações a cada 15 minutos

Hipótese nula	Período de 1 de agosto de 2006 a 30 de dezembro de 2008	Período de 1 de agosto de 2006 a 29 de agosto de 2008
	Estatística F valor-p	Estatística F valor-p
$ft$ não Granger causa $st$	175,276 0	248,5585 0
$st$ não Granger causa $ft$	5,7305 0	2,2267 0,006676

Fonte: elaborado pelo autor.

As duas hipóteses nulas não podem ser rejeitadas. O Ibovespa Futuro não causa no sentido de Granger o Ibovespa e o Ibovespa à vista não causa no sentido de Granger o Ibovespa Futuro. Estes resultados significam que existe uma relação bicausal, o que, por sua vez, confirma a forte relação entre as duas variáveis.

Cabe lembrar que os resultados do teste de causalidade de Granger não significam necessariamente que uma variável “causa” a outra. Esta possível causalidade pode ser devido à outra variável que não esteja presente no modelo, mas que causa as duas variáveis do modelo.

Depois de estimar os coeficientes, a próxima etapa consiste em testar a normalidade dos resíduos e se estes são autocorrelacionados. A autocorrelação ocorre em análises de séries temporais quando os resíduos associados com observações em um dado período de tempo se mantêm por transferência nos períodos de tempos futuros.

A hipótese nula de que os resíduos não são autocorrelacionados é rejeitada. Segundo Brooks *et al.* (2001), a presença de autocorrelação dos resíduos indica



que os coeficientes estimados são inefficientes e não apresentam a menor variância possível, mas não são viesados.

Outra propriedade dos estimadores de mínimos quadrados é a de que os erros seguem uma distribuição normal. A hipótese nula de que os resíduos são distribuídos normalmente é rejeitada. Porém, a violação desta premissa não apresenta nenhuma consequência importante para amostras grandes (Brooks *et al.*, 2001).

O último modelo a ser estimado é o modelo VECM. Segundo Tsay (2005), o maior problema com o uso das variáveis em diferenças é que o uso destas integradas pode remover as relações de longo prazo, apesar de as relações de curto prazo ainda serem captadas pelo modelo. Com isso, a transformação das séries não estacionárias em estacionárias por meio de sua diferença não é suficiente para que o resultado de sua regressão seja válido.

Por este motivo, foi utilizado o VECM para verificar a relação entre os valores prévios das variáveis com o valor atual do Ibovespa. Mas, antes de estimar o modelo, devem-se realizar os testes de cointegração nas variáveis de estudo. Estes testes são necessários para que se possa identificar a existência de equilíbrio de longo prazo entre as variáveis, em que estas podem se desviar no curto prazo, mas convergem ao equilíbrio no tempo.

O primeiro teste consiste no teste ADF de raiz unitária nos resíduos do modelo de equilíbrio de longo prazo, em que a variável  $S_t$  é a variável dependente e a variável  $F_t$  é a variável independente do modelo. A metodologia de Engle & Granger (1987) é utilizada neste teste. O primeiro passo consiste em verificar se as variáveis presentes no modelo são integradas de mesma ordem. As variáveis  $F_t$  e  $S_t$  são integradas de mesma ordem para os dois períodos de análise e são  $I(1)$ , conforme apresentado no início deste capítulo.

O segundo passo consiste no ajuste do modelo de equilíbrio de longo prazo entre as variáveis  $F_t$  e  $S_t$  e no cálculo dos resíduos deste modelo. Segundo a teoria de Engle e Granger, se os resíduos do modelo das duas séries não possuam raiz unitária, haverá co-integração entre as duas séries. O terceiro passo consiste no teste de raiz unitária nos resíduos do modelo.

Pelo teste de Engle e Granger, a hipótese nula não pode ser rejeitada. Com isso, os resíduos não têm raiz unitária.

A existência de cointegração entre as duas séries pode ser confirmada pelo teste de Cointegração de Johansen. Este é feito por meio de duas estatísticas: estatística do traço; estatística do máximo autovalor.

As hipóteses nulas de que não existe vetor de cointegração são rejeitadas. A existência de cointegração implica a adoção do termo de correção de erro no modelo VAR para os dois períodos de análise.

A comparação dos modelos de previsão foi feita pela estatística de desempenho MAPE e pelo número de vezes que o modelo acertou e errou a direção do movimento do índice.

A Tabela 2 apresenta as estatísticas MAPE para os quatro modelos estimados neste trabalho e para os dois períodos de análise. Estas estatísticas foram calculadas sobre as previsões da série no nível, e não sobre o log-retorno da série.

**Tabela 2**

Avaliação das previsões dos modelos estimados para a série  $S_t$  do via erro percentual absoluto médio

Estatísticas	ARIMA	ARFIMA	VAR	VECM
Período 2 de janeiro de 2009 a 16 de outubro de 2009				
MAPE	0,19%	0,18%	0,23%	0,21%
Período 01 de setembro de 2008 a 16 de outubro de 2009				
MAPE	0,27%	0,25%	0,32%	0,29%

Fonte: elaborado pelos autores.

Para os dois períodos, o modelo ARFIMA é o que teve melhor desempenho. A Tabela 3 apresenta o número de vezes que cada modelo acertou e errou a direção do movimento do índice Bovespa.

**Tabela 3**

Número de vezes que cada modelo acertou e errou a direção do movimento da série  $S_t$

Modelo	Período 2 de janeiro de 2009 a 16 de outubro de 2009			Período 01 de setembro de 2008 a 16 de outubro de 2009		
	Previsões certas	Previsões erradas	% de acertos	Previsões certas	Previsões erradas	% de acertos
ARIMA	3321	3175	51,12	6753	2539	72,68
ARFIMA	2861	3635	44,04	4701	4591	50,59
VAR	3183	3313	49	4854	4438	52,24
VECM	3261	3235	50,2	4779	4513	51,43

Fonte: elaborado pelos autores.

Para o primeiro período, os modelos apresentaram desempenho preditivo próximo dos 50% de acertos. O modelo ARIMA é o que apresentou melhor desempenho nesse critério, porém todos os modelos não acertaram significativamente a direção, mais de 50% de acertos. Isto implica que os resultados podem ser aleatórios ou imprevisíveis, como o lance de uma moeda.

Para o segundo período, o desempenho dos modelos foi semelhante ao primeiro período, com exceção do modelo ARIMA, que teve 72,68% de acertos. Cabe destacar que a principal diferença das previsões entre os dois períodos está no fato de que no segundo período as observações possuem uma alta volatilidade, devendo à crise financeira de 2008. Porém, os modelos conseguem captar de maneira semelhante à direção do movimento do índice nos dois períodos.

## 5.1 Estratégias de negociação

As três estratégias de negociação foram feitas com as previsões dos quatro modelos para os dois períodos de análise, considerando inicialmente um cenário ausente de custos de negociação e outro com a presença destes custos.



Os custos considerados foram a taxa de liquidação de 0,019% cobrada pela Bovespa e mais 0,006% de emolumentos cobrado pela CBLC. A arbitragem é possível apenas quando os custos de transação são inferiores aos retornos proporcionados pela transação do ativo. Como a arbitragem envolve a compra e a venda do ativo, o custo de transação total é de 0,05% do valor envolvido.

Além destes, existe o custo de corretagem. Como este é normalmente uma taxa fixa e quanto maior volume financeiro menos significativo se torna o custo de corretagem, este foi desconsiderado na análise de arbitragem. A Tabela 4 apresenta as estratégias de negociação com base nos retornos preditos dos modelos de previsão.

**Tabela 4**

Estratégias de negociação com base nos retornos preditos dos modelos de previsão sem os custos de transação: investimento final; retorno percentual mensal (21 dias)

Modelo	Período 2 de janeiro de 2009 a 16 de outubro de 2009				Período 01 de setembro de 2008 a 16 de outubro de 2009			
	Invest.	LTS	BHS	MFS	Invest.	LTS	BHS	MFS
Passivo								Passivo
ARIMA	919,63 (-0,20%)	1.043,26 (0,10%)	1.043,26 (0,10%)	1028,64 (0,07%)	1190,37 (0,48%)	1.041,78 (0,11%)	1.041,78 (0,11%)	1.081,21 (0,22%)
ARFIMA	919,63 (-0,20%)	926,76 (-0,18%)	926,76 (-0,18%)	801,54 (-0,53%)	1190,37 (0,48%)	1.404,38 (0,94%)	1.404,38 (0,94%)	1.052,31 (0,14%)
VAR	919,63 (-0,20%)	<b>1051,47</b> (0,12%)	<b>1051,47</b> (0,12%)	1314,49 (0,65%)	1190,37 (5,71%)	<b>7.453,53</b> (5,71%)	<b>7.453,53</b> (5,17%)	6.175,54
VECM	919,63 (-0,20%)	932,33 (-0,20%)	932,33 (-0,20%)	803,63 (-0,52%)	1190,37 (0,48%)	7.237,83 (5,63%)	7.237,83 (5,63%)	4.961,88 (4,53%)

Fonte: elaborado pelo autor.

O investimento passivo (Invest. Passivo) consiste em duas negociações, em que se compra o índice no instante inicial e vende o índice no instante final. O investimento inicial é de \$1000 unidade monetária. Este é utilizado como *benchmark* para as outras estratégias. Pela Tabela 4, as melhores estratégias sem os custos de negociação são as estratégias LTS e BHS para todos os modelos. Estas apresentam o mesmo valor porque a diferença está no fato de que a estratégia BHS tem como objetivo realizar um número menor de negociações. Porém, como neste cenário não há custos de negociação, as estratégias vão ser iguais. Pela Tabela 4, o modelo VAR apresentou um melhor desempenho em relação aos outros, embora nos critérios estatísticos este modelo não tenha sido o de melhor desempenho.

Novamente, pela Tabela 4, pode-se perceber que o retorno médio mensal do modelo VAR nas estratégias LTS e BHS é de 0,12% para o primeiro período e de 5,71% para o segundo período, contra um retorno de -0,20% pela estratégia passiva para o primeiro período e de 0,48% para o segundo período. Por estes resultados é possível responder parcialmente o objetivo do presente trabalho, em que é possível obter retornos anormais, retornos acima de um *benchmark*, utilizando modelos previsão que utilizam as relações de liderança e defasagem entre o índice Bovespa e o índice Bovespa Futuro. A próxima etapa consiste em testar os modelos em um cenário onde existem custos de negociação e verificar se estes conseguem superar o *benchmark*. A Tabela 5 apresenta os retornos das estratégias

quando são considerados os custos de transação.

**Tabela 5**

Estratégias de negociação com base nos retornos preditos dos modelos de previsão considerando os custos de transação investimento final; retorno percentual mensal (21 dias)

Modelo	Período 2 de janeiro de 2009 a 16 de outubro de 2009				Período 01 de setembro de 2008 a 16 de outubro de 2009			
	Invest. Passivo	LTS	BHS	MFS	Invest. Passivo	LTS	BHS	MFS
ARIMA	919,13 (-0,20%)	206,11 (-3,69%)	675,56 (-0,93%)	665,10 (-0,97%)	1189,87 (0,48%)	99,72 (-6,18%)	575,72 (-1,52%)	638,60 (-1,23%)
ARFIMA	919,13 (-0,20%)	141,59 (-4,55%)	713,66 (-0,80%)	555,01 (-1,39%)	1189,87 (0,48%)	121,21 (-5,67%)	822,04 (-0,54%)	546,30 (-1,66%)
VAR	919,13 (-0,20%)	202,30 (-3,73%)	729,17 (-0,75%)	945,42 (-0,13%)	1189,87 (0,48%)	696,55 (-1,00%)	4.487,85 (4,24%)	4.110,51 (3,99%)
VECM	919,13 (-0,20%)	172,60 (-4,10%)	601,93 (-1,20%)	500,49 (-1,63%)	1189,87 (0,48%)	672,01 (-1,09%)	3.903,00 (3,84%)	2.742,07 (2,83%)

Fonte: elaborado pelo autor.

Quando os custos são inseridos, a estratégia que possui pior retorno para os dois períodos é a estratégia LTS, que envolve um número maior de negociações. Esta estratégia apresenta retorno negativo para todos os modelos quando os custos de negociação são considerados. A estratégia BHS foi à melhor para os dois períodos, considerando os custos, já que em todos os modelos esta apresenta retornos maiores que as outras estratégias.

A estratégia MFS apresenta um número pequeno de negociações, comparado às estratégias LTS e BHS, por ser mais seletiva e conservadora. A estratégia MFS, por ser mais conservadora, não apresenta um alto número de negociações, porém não consegue obter todo o ganho com a volatilidade do índice no período.

O resultado mais importante observado na Tabela 5 é que para o modelo VAR a estratégia BHS apresenta retorno maior que a estratégia passiva no segundo período de análise, o que viabiliza a estratégia. Este resultado contraria a literatura –, Silva (2006) para o Brasil e Brooks *et al.* (2001) para o mercado britânico – que mostra que as estratégias de negociação não seriam lucrativas em relação a uma estratégia passiva quando se consideram os custos de negociação. Porém, não foi considerado nenhum custo de corretagem. E, se este for uma porcentagem acima de 0,025% sobre cada movimentação, nenhum modelo e nenhuma estratégia de negociação será lucrativa, e com isso não se obtém retornos anormais.

Este resultado corrobora ainda mais a objetivo deste artigo: É possível obter estratégias de negociação lucrativas sobre o efeito de liderança e defasagem entre o índice Bovespa à vista e o índice Bovespa Futuro e é possível com isso auferir ganhos anormais utilizando modelos de previsão com base neste efeito e considerando os custos de negociação para o segundo período de análise. Porém, para o primeiro período de análise nenhuma estratégia consegue obter retornos positivos. Cabe ressaltar que o primeiro período utiliza as observações do período de maior intensidade da crise financeira para estimar os modelos e as previsões para o período de recuperação da crise.



Dante disso, nenhum modelo consegue ter um bom desempenho para o primeiro período, conforme pode ser visto nas Tabelas 4 e 5. Isto pode ser devido à alta volatilidade do índice no período da crise financeira, dessa maneira os modelos não conseguem ter um desempenho preditivo que consiga gerar estratégias lucrativas.

## 6. Considerações Finais

O objetivo deste trabalho foi identificar estratégias de negociação lucrativas com base nos efeitos de liderança e defasagens entre os mercados acionários à vista e futuro no Brasil, mediante a utilização de modelos econôméticos de previsão. A melhor estratégia de negociação encontrada para todos os modelos e para os dois períodos de análise foi à estratégia Buy and Hold (BHS).

O ajuste de um modelo ARIMA para a série Ibovespa gerou um percentual de acerto dos movimentos futuros do índice de 51,12% das vezes no primeiro período e 72,68% no segundo período. A estratégia BHS usando este modelo consegue obter um retorno maior do que a estratégia passiva apenas no primeiro período, considerando o cenário sem custo de transação. Para os outros cenários, a estratégia BHS não consegue obter um retorno anormal usando o modelo ARIMA.

Já as estimativas utilizando o modelo ARFIMA geram um percentual de acerto de somente 44,04% da direção do índice no primeiro período e 50,59% das vezes no segundo período. Apenas para o segundo período e em um cenário sem custos de negociação, a estratégia BHS usando este modelo, consegue ter um retorno anormal de 0,94%, superando a estratégia passiva (0,48%). Nos outro cenário para o segundo período e para os dois cenários no primeiro período a estratégia BHS usando este modelo não consegue superar a estratégia passiva.

Por meio do modelo VAR bivariado, que utiliza as séries Ibovespa e Ibovespa Futuro, foi possível detectar uma causalidade bidirecional entre as duas séries para os dois períodos de análise. Porém, este modelo também tem um desempenho insatisfatório, ao acertar 49,00% das vezes a direção do índice no primeiro período e 50,20% das vezes no segundo período. A estratégia BHS usando este modelo para o segundo período supera a estratégia passiva nos dois cenários de negociação. Para o primeiro período, apenas no cenário sem custos de negociação a estratégia BHS consegue superar a estratégia passiva.

Por fim, de modo a verificar a possível relação de longo prazo entre as séries, um modelo VECM foi ajustado de modo a captar a co-integração entre as séries Ibovespa e Ibovespa Futuro. Os ajustes para este modelo geraram um percentual de acerto de 50,20% na direção do índice para o primeiro período e 51,43% para o segundo período. A estratégia BHS usando este modelo para o segundo período consegue superar a estratégia passiva nos dois cenários. Porém, para o primeiro período, a estratégia BHS não consegue superar a estratégia em nenhum dos dois cenários.

Com base nos resultados de todos os modelos para os dois períodos de análise e utilizando as estratégias de negociação, algumas considerações podem ser feitas.

Destaca-se a obtenção de resultados contrários aos da literatura relacionada, que mostrava que as estratégias de negociação não eram lucrativas em relação a uma estratégia passiva quando considerava os custos de negociação. Isto porque, pelos resultados encontrados neste estudo, é possível obter retornos anormais com a utilização de estratégias de negociação com o modelo VAR sobre os efeitos de liderança e defasagem entre o índice Bovespa e o índice Bovespa futuro, sendo que esse resultado é válido apenas para o segundo período de análise. Adicionalmente, este resultado implica que, quando incorporadas às observações da amostra o período de maior volatilidade da crise de 2008, os modelos não conseguem obter boas previsões.

Porém, cabe destacar que nestes resultados não foram incorporados quaisquer custos de corretagem. Se os mesmos representarem uma porcentagem superior a 0,025% sobre cada movimentação, nenhum modelo e nenhuma estratégia de negociação obtém retornos anormais. Para a realização de novas pesquisas, sugere-se a investigação do impacto do viés do Bid&Ask spread nas estratégias de negociação. A suposição de um preço único na estratégia de negociação é uma limitação deste trabalho, que poderia ser investigada com mais detalhes em outros trabalhos. Outra sugestão é a incorporação da equação da variância em todos os modelos investigados, já que o impacto da volatilidade nos resultados das estratégias de negociação ainda é uma questão em aberto.

## Referências

- Bachelier, Louis. 1964. Theory of Speculation. In: Cootner, P. (ed), *The Random Character of Stock Market Prices*. Cambridge, MA: MIT Press. (Originalmente publicado em 1900).
- Box, George E. P., & Pierce, David. 1970. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of the American Statistical Association*, **65**, 1509–1526.
- Brooks, Chris, Rew, Alistair G., & Ritson, Stuart A. 2001. Trading Strategy Based on the Lead-Lag Relationship Between the Spot Index and Futures Contract for the FTSE 100. *International Journal of Forecasting*, **17**, 31–44.
- Cowles, Alfred. 1933. Can Stock Market Forecasters Forecast? *Econometrica*, **1**, 309–324.
- Elton, Edwin J., Gruber, Martin J., Brown, Stephen J., & Goetzmann, William N. 2004. *Moderna Teoria de Carteiras e Análise de Investimentos*. São Paulo: Atlas.
- Enders, Walter. 2004. *Applied Econometric Time Series*. Alabama: Wiley.
- Engle, Robert F., & Granger, Clive W. J. 1987. Cointegration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, **55**, 251–276.



- Fama, Eugene F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, **25**, 383–417.
- Galvão, Ana B. C., Portugal, Marcelo S., & Ribeiro, Eduardo P. 2000. Volatilidade e Causalidade: Evidências Para O Mercado À Vista e Futuro de Índice de Ações No Brasil. *Revista Brasileira de Economia*, **54**, 1–15.
- Geweke, John, & Porter-Hudak, Susan. 1983. The Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis*, **4**, 221–238.
- Granger, Clive W. J. 1969. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, **37**, 424–438.
- Grant, Jeremy. 2010. *High-Frequency Trading: Up Against a Bandsaw*. Financial Times. <http://www.ft.com/cms/s/0/b2373a36-b6c2-11df-b3dd-00144feabdc0.html>.
- Haugen, Robert A. 1997. *Modern Investment Theory*. New Jersey: Prentice-Hall.
- Kawaller, Ira G., Koch, Paul D., & Koch, Timothy W. 1987. The Temporal Price Relationship Between S&P 500 Futures and the S&P 500 Index. *Journal of Finance*, **42**, 1309–1329.
- Lien, Donald, & Tse, Yiu K. 1999. Forecasting the Nikkei Spot Index with Fractional Cointegration. *Journal of Forecasting*, **18**, 259–273.
- Oliveira, Gustavo R. 2008. *Testando a Existência de Efeito Lead-Lag Entre Os Mercados Acionários Norte-Americano e Brasileiro*. Dissertação (Mestrado em Administração) – Departamento de Administração, Universidade de Brasília, Brasilia.
- Peters, Edgard E. 1994. *Fractal Market Analysis*. New York: John Wiley.
- Silva, Daphnis T., Jr. 2006. *O Conteúdo Informacional Dos Contratos Futuros de IBOVESPA*. Tese (Doutorado em Contabilidade) – Departamento de Contabilidade e Atuária, Universidade de São Paulo, São Paulo, 2006.
- Stoll, Hans R., & Whaley, Robert E. 1990. The Dynamics of Stock Index and Stock Index Futures Returns. *Journal of Financial and Quantitative Analysis*, **25**, 441–468.
- Tsay, Ruey S. 2005. *Analysis of Financial Time Series*. Chicago: Wiley Series in Probability and Statistics.
- Tse, Yiu K. 1995. Lead-Lag Relationships Between Spot Index and Futures Price of the Nikkei Stock Average. *Journal of Forecasting*, **14**, 553–563.

Tse, Yiu K., & Chan, W. S. 2010. The Lead-Lag Relation Between the S&P500 Spot and Futures Markets: An Intraday-Data Analysis Using a Threshold Regression Model. *Japanese Economic Review*, **61**, 133–144.

Wahab, Mahmoud, & Lashgari, Malek. 1993. Price Dynamics and Error Correction in Stock Index and Stock Index Futures Markets: A Cointegration Approach. *Journal of Futures Markets*, **13**, 711–742.

