

FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

MEIRE MIDORI HORI PEREIRA

**IMPACTO DAS NEGOCIAÇÕES ALGORÍTMICAS DE ALTA FREQUÊNCIA NO
MERCADO FUTURO DE DÓLAR**

SÃO PAULO

2014

MEIRE MIDORI HORI PEREIRA

**IMPACTO DAS NEGOCIAÇÕES ALGORÍTMICAS DE ALTA FREQUÊNCIA NO
MERCADO FUTURO DE DÓLAR**

**Dissertação apresentada ao Programa de
Mestrado Profissional em Economia da
Fundação Getúlio Vargas/EESP, como parte dos
requisitos para a obtenção do título de Mestre
em Economia, linha de Finanças.**

Orientador: Prof. Dr. Pedro Luiz Valls Pereira

SÃO PAULO

2014

Pereira, Meire Midori Hori.

Impacto das negociações algorítmicas de alta frequência no mercado futuro de dólar / Meire Midori Hori Pereira. - 2014. 46 f.

Orientador: Pedro Luiz Valls Pereira

Dissertação (MPFE) - Escola de Economia de São Paulo.

1. Mercado de câmbio. 2. Mercado financeiro. 3. Algoritmos. 4. Dólar. 5 Bolsa de Valores. I. Pereira, Pedro Luiz Valls. II. Dissertação (MPFE) - Escola de Economia de São Paulo. III. Título.

CDU 336.76

MEIRE MIDORI HORI PEREIRA

**IMPACTO DAS NEGOCIAÇÕES ALGORÍTMICAS DE ALTA FREQUÊNCIA NO
MERCADO FUTURO DE DÓLAR**

**Dissertação apresentada ao Programa de
Mestrado Profissional em Economia da
Fundação Getúlio Vargas/EESP, como parte dos
requisitos para a obtenção do título de Mestre
em Economia, linha de Finanças.**

Data da Aprovação:

____/____/____

Banca Examinadora:

Prof. Dr. Pedro Luiz Valls Pereira (Orientador)
FGV - EESP

Prof. Dr. Emerson Fernandes Marçal
FGV - EESP

Prof. Dr. Luiz Koodi Hotta
IMECC-UNICAMP

AGRADECIMENTOS

Aos meus pais, pelo apoio e incentivo em todos os momentos.

Ao professor Pedro, pela orientação e paciência.

Aos meus colegas do curso por trilharem este caminho ao meu lado, compartilhando seus conhecimentos e acima de tudo a amizade.

Aos meus colegas de trabalho pelo incentivo e apoio.

Aos professores pelos ensinamentos e apoio.

RESUMO

Este trabalho apresenta um estudo do impacto das negociações algorítmicas no processo de descoberta de preços no mercado de câmbio. Foram utilizados dados de negociação de alta frequência para contratos futuros de reais por dólar (DOL), negociados na Bolsa de Valores de São Paulo no período de janeiro a junho de 2013. No intuito de verificar se as estratégias algorítmicas de negociação são mais dependentes do que as negociações não algorítmicas, foi examinada a frequência em que algoritmos negociam entre si e comparou-se a um modelo *benchmark* que produz probabilidades teóricas para diferentes tipos de negociadores. Os resultados obtidos para as negociações minuto a minuto apresentam evidências de que as ações e estratégias de negociadores algorítmicos parecem ser menos diversas e mais dependentes do que aquelas realizadas por negociadores não algorítmicos. E para modelar a interação entre a autocorrelação serial dos retornos e negociações algorítmicas, foi estimado um vetor autorregressivo de alta frequência (VAR) em sua forma reduzida. As estimações mostram que as atividades dos algoritmos de negociação causam um aumento na autocorrelação dos retornos, indicando que eles podem contribuir para o aumento da volatilidade.

Palavras-chave: algoritmos de negociação; dados de alta frequência.

ABSTRACT

This work presents a study of the impact of algorithmic tradings in the process of price discovery in the foreign exchange market. It was used high-frequency data for the U.S. Dollar Futures Contract trade in the São Paulo Stock Exchange, from January 2013 to June 2013. In order to verify if algorithmic trading strategies are more dependent than those of non-algorithmic tradings, it was examined the frequency at which algorithmic traders negotiate with each other and compared it to a benchmark model that produces theoretical probabilities for different types of traders. The results obtained for these minute by minute tradings present evidence that actions and strategies for algorithmic traders appear to be less diverse and more dependent than those held by non-algorithmic traders. And it was estimated a reduced form of a high-frequency vector autoregressive (VAR) to model the interaction between serial autocorrelations of returns and algorithmic trading. The estimates show that algorithmic trading activities cause an increase in the autocorrelations of returns, indicating they may contribute to a rise in the volatility.

Keywords: algorithmic trading; high-frequency data.

SUMÁRIO

1 INTRODUÇÃO	8
2 REVISÃO BIBLIOGRÁFICA	12
3 FUNDAMENTAÇÃO TEÓRICA	14
3.1 Modelo benchmark para medição da correlação	14
3.2 Outras medidas das características das atividades de negociação algorítmica	17
3.3 Retornos	18
3.3.1 Retornos para dados de alta frequência	19
3.4 Modelos Autoregressivos Vetoriais	20
3.4.1 Construção de Modelos VAR	20
3.5 Causalidade de Granger	23
4 METODOLOGIA	24
4.1 Dados	24
4.2 Cálculo da medida R	25
4.3 Atividades de negociações algorítmicas	28
4.4 Modelo VAR	30
4.4.1 Estimação do modelo VAR em sua forma reduzida para as variáveis endógenas AC5_Retorno e AT (VAT, VCt, VCm, OFCt e $\ln(\hat{R})$)	32
5 CONSIDERAÇÕES FINAIS	36
REFERÊNCIAS	39

1 INTRODUÇÃO

Nos últimos anos, a utilização de algoritmos de negociação em ambientes onde computadores gerenciam processos de negociação de alta frequência, tem se tornado comum na maioria dos mercados financeiros. Isso tem conferido muitos benefícios aos participantes do mercado como uma variedade de estratégias de negociação, aumento da velocidade, precisão e ferramentas disponíveis para executar tais estratégias.

Com isso aumenta também o interesse dos participantes e, principalmente dos órgãos reguladores, em entender de que forma sua atuação pode impactar o mercado, já que alguns eventos relacionados às negociações automatizadas foram o centro das atenções.

Por exemplo, em um relatório da Commodity Futures Trading Commission (CFTC) de 2013, são apresentados alguns exemplos, como o episódio do *Flash Crash* de 6 de maio de 2010, quando os maiores índices de ações, tanto no mercado de ações quanto no mercado futuro, perderam mais de 5% de seus valores em questão de minutos. Foi causado por uma ordem automatizada que levou a um movimento extremo de queda de preços e uma crise de liquidez no Chicago Mercantile Exchange's (CME) em contratos futuros de E-mini.

Em 2012, uma empresa de serviços financeiros no mercado de ações, Knight Capital Group (mais tarde adquirida pela Getco LLC, formando a KCG Holdings) submeteu de forma equivocada, um número significativo de ordens proprietárias à New York Stock Exchange ("NYSE"), causando uma brusca oscilação em quase 150 ativos, que custou a empresa aproximadamente \$440 milhões no processo. O episódio mais recente ocorreu em Agosto de 2013, quando as negociações no mercado de ações da NASDAQ foram interrompidas por três horas devido ao mau funcionamento nos sistemas de disseminação das ofertas e potenciais questões de conectividade com outros sistemas de plataforma de negociação.

Exemplos como estes mostram que inovações tecnológicas complexas nos mercados financeiros requerem uma análise cuidadosa dos potenciais impactos e riscos que eles podem acarretar.

Existem opiniões controversas sobre esses impactos no mercado provocados por sua utilização. Enquanto uns destacam os benefícios proporcionados

como mais eficiência na descoberta de preços, redução *do bid-ask spread*, provimento de liquidez, fragmentação do mercado, outros se preocupam com questões como o aumento do custo de seleção adversa, excesso de volatilidade, afastamento dos preços de seus valores fundamentais, criação de momentum de preços.

Negociação de alta frequência ou *High Frequency Trading (HFT)* é uma forma de negociação automatizada baseada em algoritmos matemáticos. Não é uma estratégia de negociação em si, mas um meio de colocar certas estratégias em prática nas plataformas de negociação. A principal característica é a importância dos cálculos rápidos e a velocidade de execução dessas estratégias.

Pode ser classificado como uma subcategoria de algoritmos de negociação, que é o termo coletivo para todas as estratégias onde ordens são dadas de acordo com um conjunto de regras pré-programadas (algoritmos). Portanto, nem todos os tipos de negociações automatizadas podem ser classificados como *HFT*.

Segundo Aldridge (2010), a principal inovação que separa uma negociação de alta frequência de uma negociação de baixa frequência é uma alta rotatividade de capital em respostas rápidas direcionadas por computadores para mudanças nas condições do mercado. As estratégias de *HFT* são caracterizadas por um grande volume de negócios e margens de lucro muito pequenas por transação.

Ao contrário dos tradicionais participantes que mantêm suas posições por dias ou até mesmo meses, gerando pequenos pontos percentuais em retorno por negócio, *HFTs* executam vários negócios por dia, ganhando uma fração de porcentagem de retorno por negócio e suas posições são normalmente tomadas com a intenção de ser *market-neutral* (não direcional), hedgeadas (*delta hedge*) e serão fechadas no final do dia (posições são raramente mantidas de um dia para o outro). Os períodos médios em que essas posições são mantidas normalmente têm durações muito curtas, de segundos a vários minutos. Estratégias de *HFT* podem ser divididas em algumas categorias como *market making*, arbitragem estatística e estratégias de baixa latência.

Os volumes das posições e o tempo para o qual essas posições são mantidas são determinados pelos algoritmos de negociação e podem flutuar durante o dia. Explosões de grandes quantidades de ordens, inseridas de repente, são uma das principais características de *HFT*. Essas grandes quantidades normalmente são

alternadas com períodos de relativa calma na qual praticamente quase nenhum negócio ocorre, antes de uma nova oportunidade de negócio.

A maioria das ordens inseridas não é executada, sendo estas canceladas logo após sua entrada, por serem continuamente atualizadas de acordo com as constantes mudanças das condições de mercado.

Os maiores nomes internacionais de alta frequência incluem Millennium, DE Shaw, Worldquant e Renaissance Technologies. A maior parte das empresas de alta frequência são fundos ou outras empresas de investimentos proprietárias como ATD, Cisco, Citadel, Getco, Madison Tyler, Tradebot, All Options, Flow Traders, IMC, Optiver, mesas de negociação de grandes bancos (de investimento) tais como Goldman Sachs, JP Morgan, Morgan Stanley, BNP Paribas e Société Générale também utilizam estratégias *HFT* como parte de suas negociações proprietárias.

O volume de negociação de alta frequência representa uma porção significativa do mercado americano, chegando a quase 70% das operações. No Brasil, essa participação ainda está em crescimento e é mais modesta, com aproximadamente 13% no segmento de ações e 9% em derivativos, segundo dados divulgados no site da BM&FBOVESPA.

Esta dissertação está baseada no artigo de Chaboud et al. (2013), *Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market*, cujo objetivo é estudar o impacto dos algoritmos de negociação no processo de tornar a informação dos preços mais eficiente no mercado de câmbio, para contratos futuros de reais por dólar (DOL) ao analisar o comportamento de suas estratégias.

Uma das preocupações destacadas na literatura, como no trabalho de Jarrow e Protter (2012) é de que as estratégias de negociação utilizadas por computadores são mais correlacionadas do que aquelas utilizadas pelos humanos. Isso porque atuam coletivamente de forma similar, aumentando a movimentação do mercado e causando efeitos negativos sobre os preços e potencialmente criando volatilidade excessiva.

Para verificar se as estratégias algorítmicas de negociação são mais correlacionadas do que as negociações não algorítmicas, examinou-se a frequência em que algoritmos negociam entre si e comparou-se a um modelo *benchmark* que produz probabilidades teóricas para diferentes tipos de negociadores. A principal ideia

desta medida é que negociadores que seguem estratégias de negociação similares, irão negociar menos entre si comparado àqueles que seguem estratégias menos correlacionadas. Comparando as previsões do modelo com os valores observados nos dados, há evidências de que as estratégias utilizadas com os auxílio de algoritmos pelos computadores são mais correlacionadas e menos diversas do que aquelas utilizadas por humanos.

Em seguida, foi investigado o efeito que os algoritmos de negociação tem sobre uma medida mais geral de preços não ser informacionalmente eficiente, através da autocorrelação serial estimados de retornos de 5 segundos sobre intervalos de 5 minutos. Para ser eficiente, os preços precisam exibir correlação nula. Para esta análise, estimou-se um vetor autorregressivo (VAR) com dados de alta frequência.

O trabalho está organizado da seguinte forma. No capítulo 2 é discutida a literatura empírica relacionada sobre os efeitos dos algoritmos de negociação na qualidade do mercado. No capítulo 3 são apresentados conceitos e o modelo que será utilizado. O capítulo 4 apresenta a descrição dos dados e a aplicação empírica da modelagem apresentada. No capítulo 5 são apresentadas as considerações finais e sugestões para futuros estudos.

2 REVISÃO BIBLIOGRÁFICA

Brogaard (2010) analisa o impacto dos *High Frequency Traders (HFTs)* no mercado de ações americano. Ele estuda suas estratégias, rentabilidade e relação com características de todo o mercado, como liquidez, descoberta de preço e volatilidade. Dentre os resultados, o autor verificou que *HFTs* tendem a seguir uma estratégia reversa de preço causada por desequilíbrio nas ordens, lucram aproximadamente \$3 bilhões por ano, não fazem *front running* sistematicamente contra não *HFTs*, contam com um conjunto de estratégias menos diversificado, entre outros. Ele também analisa as consequências de *HFT* na funcionalidade do mercado de ações americano, realizando vários testes e algumas das conclusões são que *HFTs* contam com um conjunto de estratégias menos diversificada, contribuem substancialmente para o processo de descoberta de preços, fornecem a melhor oferta de compra, não há evidências de que aumentem a volatilidade e podem na verdade reduzi-la.

O comportamento de *High Frequency Traders* e outras categorias de negociadores no *Flash-Crash* de 06 de maio de 2010, no mercado Futuro de Índice *E-mini Standard&Poor's 500* foi analisado por Kirilenko et al. (2011). Eles concluíram que os *HFT* não acionaram o *crash* mas contribuíram por terem deixado o mercado assim que as condições se tornaram desfavoráveis. Acreditam que com as mudanças no mercado, proteções adequadas devem ser implementadas.

Biais e Woolley (2011) procuram mostrar os dois lados da atuação de *HFTs*, baseados em discussões com participantes do Mercado e estudos empíricos. São apresentadas as motivações para *HFT*, como por exemplo, redução dos custos operacionais de bancos e firmas financeiras, melhora no funcionamento dos mercados, colaboração com a fragmentação do mercado promovendo competição por facilidades de negociações efetivas e baratas.

As negociações de alta frequência também podem ser utilizadas para a manipulação de mercado, com a utilização de estratégias como *stuffing*, *smoking*, *spoofing*, entre outros. A utilização dessas estratégias pode provocar, por exemplo, seleção adversa, competições imperfeitas, risco sistêmico.

Bias, Foucault e Moinas (2011) analisam equilíbrios de negociação para um dado nível de investimento em tecnologias de negociação rápida. Concluem que

isso ajuda as instituições financeiras a lidarem com a fragmentação de mercado, melhorando suas habilidades em aproveitar as oportunidades das negociações, e aumentando os ganhos gerados pelo negócio. Por outro lado, instituições rápidas observam e processam informações antes dos negociadores mais lentos, criando seleção adversa. Portanto investimento em negociações rápidas gera externalidades negativas.

Uma outra relação importante é a relação empírica entre algoritmos de negociação e liquidez. Os resultados de Hendershott, Jones e Menkveld (2011) destacam esta importância.

Hendershott e Riordan (2011) examinam o papel dos *HFTs* na descoberta de preços. Concluem que eles desempenham um papel positivo na eficiência de preços ao negociarem na direção de mudanças de preço permanente e na direção oposta dos erros de apreçamento transitórios através de suas ordens negociáveis. Por outro lado, as ordens passivas não negociáveis dos *HFTs* são selecionadas adversamente tanto nos componentes transitórios como permanentes dos preços. Os resultados não fornecem evidências de que *HFT* contribuem para a instabilidade dos preços no mercado.

Jarrow e Protter (2012) constroem um modelo para mostrar que negociadores de alta frequência não aumentam a eficiência dos mercados eletrônicos e podem provocar aumento da volatilidade e desviar os preços de seus valores fundamentais por apresentarem comportamentos similares e atuarem como um “*big trader*”

Hasbrouck e Saar (2013) analisam certos tipos de negociadores algorítmicos e concluem que eles têm um efeito positivo nas medidas tradicionais de qualidade do mercado – diminuição dos *spreads*, aumento na profundidade observada no livro de ordens limite e diminuição da volatilidade de curto prazo.

Chaboud et al. (2013) estudam os negociadores algorítmicos nos mercados de câmbio e encontram evidências de que eles promovem uma melhora em duas medidas de eficiência de mercado: a frequência de oportunidades de arbitragem triangular e a autocorrelação dos retornos de alta frequência.

3 FUNDAMENTAÇÃO TEÓRICA

3.1 Modelo *benchmark* para medição da dependência

Os dados de transação fornecem informações sobre o volume e direção, compra (B) ou venda (S), dos negócios que podem ser atribuídos a negociadores algorítmicos (computadores (C)) e não algorítmicos (humanos (H)). Será utilizada a expressão “*maker*” para designar o negociador que coloca ofertas antes que o “*taker*” escolha negociar naquele preço. O *maker* fornece liquidez enquanto o *taker* é o “agressor” ou “*initiator*” do negócio. Existem quatro possíveis pares de humanos e computadores *makers* e *takers*, dos quais podem ser observados o volume de negociação e o fluxo de ordem a cada minuto: humano-*maker* / humano-*taker* (HH), computador-*maker* / humano-*taker* (CH), humano-*maker* / computador-*taker* (HC) e computador-*maker* / computador-*taker* (CC). Para cada par, por exemplo, CC , o elemento da direita representa o *taker* e o da esquerda o *maker*.

A soma dos volumes de compra e venda para cada par fornece o volume de negociação atribuível àquela particular combinação de *maker* e *taker*, denotado como $Vol(HH)$, $Vol(CH)$, $Vol(HC)$ e $Vol(CC)$. A soma dos quatro volumes, $Vol(HH+CH+HC+CC)$ representa o volume total de negócios no mercado.

A soma dos volumes de compra para cada par, fornece o volume de negócio daquela particular combinação de *maker* e *taker*, onde o *taker* está comprando, denotado por $Vol(HH^B)$, $Vol(CH^B)$, $Vol(HC^B)$ e $Vol(CC^B)$.

A soma dos volumes de venda para cada par, fornece o volume de negócio daquela particular combinação de *maker* e *taker*, onde o *taker* está vendendo, denotado por $Vol(HH^S)$, $Vol(CH^S)$, $Vol(HC^S)$ e $Vol(CC^S)$.

A diferença entre o volume de compra e venda para cada par fornece o fluxo de ordem atribuível àquela combinação de *maker* e *taker*, representado por $OF(HH)$, $OF(HC)$, $OF(CH)$ e $OF(CC)$. A soma dos quatro fluxos de ordem, $OF(HH) + OF(HC) + OF(CH) + OF(CC)$ representa o fluxo de ordem total.

No modelo *benchmark*, existem H_m potenciais humano-*makers*, que é o número de humanos que estão prontos para prover liquidez, H_t potenciais humano-*takers*, C_m potenciais computador-*makers* e C_t computador-*takers*.

A probabilidade de um computador prover liquidez para um negociador é igual

$$Prob(\text{computador} - \text{maker}) = \frac{C_m}{C_m + H_m} = \alpha_m$$

e a probabilidade de um computador tirar liquidez do mercado é

$$Prob(\text{computador} - \text{taker}) = \frac{C_t}{C_t + H_t} = \alpha_t.$$

Os *makers* e *takers* humanos são representados em proporção, respectivamente por $(1 - \alpha_m)$ e $(1 - \alpha_t)$.

Assumindo que os eventos são independentes, as probabilidades das quatro possíveis combinações de negócios são dadas por:

$$Prob(HH) = (1 - \alpha_m)(1 - \alpha_t) \quad (3.1)$$

$$Prob(HC) = (1 - \alpha_m)\alpha_t \quad (3.2)$$

$$Prob(CH) = \alpha_m(1 - \alpha_t) \quad (3.3)$$

$$Prob(CC) = \alpha_m\alpha_t. \quad (3.4)$$

E deriva-se a seguinte identidade

$$Prob(HH) \times Prob(CC) \equiv Prob(HC) \times Prob(CH),$$

$$\frac{Prob(HH)}{Prob(CH)} \equiv \frac{Prob(HC)}{Prob(CC)}.$$

O lado esquerdo da expressão representa a razão humano-*taker*, $RH \equiv \frac{Prob(HH)}{Prob(CH)}$, e o lado direito, a razão computador-*taker*, $RC \equiv \frac{Prob(HC)}{Prob(CC)}$. Considerando que existem mais negociadores *makers* e *takers* humanos, espera-se que estas razões sejam maiores do que um, pois computadores tomam mais liquidez de humanos do que de outros computadores, $Prob(HC) > Prob(CC)$, e humanos tomam mais liquidez de humanos do que de computadores, $Prob(HH) > Prob(CH)$. Então, sob a suposição de combinação aleatória dos pares do modelo, a identidade acima mostra que a razão das razões, será igual a um,

$$R \equiv \frac{RC}{RH} = 1. \quad (3.5)$$

Humanos tomarão liquidez de outros humanos na mesma proporção que

computadores tomam liquidez de humanos. Por exemplo, a probabilidade de aleatoriamente termos um humano-maker e um humano-taker $Prob(HH) = \frac{H_m \times H_t}{(H_m + C_m) \times (H_t + C_t)} = 1$, previsto pelo modelo.

Nos dados, se for observado $R > 1$, pode implicar que as estratégias algorítmicas são mais correlacionadas do que o modelo de combinações aleatórias prevê. Ou seja, pode-se observar que ou computadores negociam entre eles menos que o esperado, $Prob(CC) < \frac{C_m \times C_t}{(H_m + C_m) \times (H_t + C_t)}$, ou computadores negociam com humanos mais que o esperado, $Prob(CH) > \frac{C_m \times H_t}{(H_m + C_m) \times (H_t + C_t)}$ ou $Prob(HC) > \frac{H_m \times C_t}{(H_m + C_m) \times (H_t + C_t)}$.

Acrescentando agora o sinal do negócio ao modelo *benchmark*, assume-se que a probabilidade de um *taker* comprar um ativo é α_B e a probabilidade de um *taker* vender é $1 - \alpha_B$. Assumindo cada evento independente, derivam-se as seguintes probabilidades:

$$Prob(HH^B) = (1 - \alpha_m)(1 - \alpha_t)\alpha_B \quad (3.6)$$

$$Prob(HC^B) = (1 - \alpha_m)\alpha_t\alpha_B \quad (3.7)$$

$$Prob(CH^B) = \alpha_m(1 - \alpha_t)\alpha_B \quad (3.8)$$

$$Prob(CC^B) = \alpha_m\alpha_t\alpha_B \quad (3.9)$$

$$Prob(HH^S) = (1 - \alpha_m)(1 - \alpha_t)(1 - \alpha_B) \quad (3.10)$$

$$Prob(HC^S) = (1 - \alpha_m)\alpha_t(1 - \alpha_B) \quad (3.11)$$

$$Prob(CH^S) = \alpha_m(1 - \alpha_t)(1 - \alpha_B) \quad (3.12)$$

$$Prob(CC^S) = \alpha_m\alpha_t(1 - \alpha_B) \quad (3.13)$$

Dessas probabilidades, é possível derivar as seguintes razões:

$$\frac{Prob(HH^B)}{Prob(CH^B)} \equiv \frac{Prob(HC^B)}{Prob(CC^B)},$$

onde o lado esquerdo da expressão representa a razão humano-taker comprador, RH^B , e o lado direito, a razão computador-taker comprador, RC^B .

E similarmente:

$$\frac{Prob(HH^S)}{Prob(CH^S)} \equiv \frac{Prob(HC^S)}{Prob(CC^S)},$$

com a razão humano-*taker* vendedor, RH^S , e a razão computador-*taker* vendedor, RC^S .

E assim como no caso da razão R , considerando-se mais negociadores humanos (*makers* e *takers*) do que computadores e cada uma das razões será maior do que um. Portanto, sob as suposições de combinações aleatórias do modelo, a razão das razões

$$R^B \equiv \frac{RC^B}{RH^B} = 1 \quad (3.14)$$

$$R^S \equiv \frac{RC^S}{RH^S} = 1. \quad (3.15)$$

Então por exemplo, a probabilidade de um humano-*maker* vendedor/ humano-*taker* comprador, $Prob(HH^S) = (1 - \alpha_m)(1 - \alpha_t)(1 - \alpha_B)$ é igual a um, quando previsto pelo modelo, sob as suposições de que as atividades dos negociadores são aleatórias e independentes. Se $R^B > 1$, pode implicar que as estratégias algorítmicas de compradores são mais dependentes do que o modelo prevê, ou seja, ou computadores negociam menos entre eles do que o esperado quando estão comprando, $(Prob(CC^B) < \alpha_m \alpha_t \alpha_B)$ ou computadores negociam mais com humanos do que o esperado quando estão comprando, $Prob(CH^B) > \alpha_m(1 - \alpha_t)\alpha_B$ ou $Prob(HC^B) > (1 - \alpha_m)\alpha_t\alpha_B$. Simetricamente, as mesmas observações são feitas para o caso em que o *taker* está comprando e em que $R^S > 1$.

3.2 Outras medidas das características das atividades de negociação algorítmica

Além da medida R , utilizada para inferir sobre a dependência das estratégias algorítmicas, também foram calculadas outras medidas para estudar os mecanismos através dos quais as negociações algorítmicas afetam o mercado. Foram criadas algumas variáveis na frequência de um minuto e diária.

A primeira delas é a medida das atividades algorítmicas total, representada pela fração do volume total onde um computador está em pelo menos uma das contrapartes do negócio, representada por

$$VAT = 100 \times \frac{Vol(HC) + Vol(CH) + Vol(CC)}{Vol(HH) + Vol(HC) + Vol(CH) + Vol(CC)}. \quad (3.16)$$

Também é calculada a fração do volume de negociação total onde um computador é o agressor no negócio, medindo-se a atividade tomadora relativa dos computadores (*taker*). É definida como

$$VCt = 100 \times \frac{Vol(HC) + Vol(CC)}{Vol(HH) + Vol(HC) + Vol(CH) + Vol(CC)}. \quad (3.17)$$

Da mesma forma, calcula-se a fração do volume de negociação total onde um computador fornece liquidez, medindo-se a atividade provedora relativa dos computadores (*maker*). É definida como

$$VCm = 100 \times \frac{Vol(CH) + Vol(CC)}{Vol(HH) + Vol(HC) + Vol(CH) + Vol(CC)}. \quad (3.18)$$

Outra medida é a participação de fluxo de ordem no mercado provenientes de negociadores algorítmicos, em relação ao fluxo de ordem total do mercado. Esta variável leva em conta o sinal dos negócios iniciados pelos computadores, representando uma medida da intensidade direcional relativa de negócios iniciados por computadores no mercado. É representada por

$$OFct = 100 \times \frac{|OF(C - taker)|}{|OF(C - take)| + |OF(H - taker)|}. \quad (3.19)$$

3.3 Retornos

Segundo Morettin (2011, p. 7), o risco de uma carteira é normalmente avaliado em termos de variação de preços dos ativos. Seja P_t o preço de um ativo no instante t . O *retorno líquido simples* ou *taxa de retorno* deste ativo entre os instantes t e $t - 1$ é definido por

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}. \quad (3.20)$$

O *retorno bruto simples* é definido por

$$1 + R_t = \frac{P_t}{P_{t-1}}. \quad (3.21)$$

Sendo $p_t = \log P_t$, o *retorno composto continuamente* ou *log-retorno* é definido como

$$r_t = \log \frac{P_t}{P_{t-1}}. \quad (3.22)$$

Para dados intradiários, consideremos $p_{d,m}$ o log-preço do ativo no período m ($m=1, 2, \dots, M$) do dia d ($d=1, 2, \dots, D$), sendo o número total de observações $T = M \times D$. Neste caso, o *retorno intradiário* é dado por

$$r_{d,m} = p_{d,m} - p_{d,m-1}, \quad m = 2, \dots, M, \quad d = 1, \dots, D. \quad (3.23)$$

Para incorporar as informações relativas ao intervalo entre o fechamento do pregão do dia anterior e a abertura do pregão do dia atual, existe o chamado *retorno overnight* definido por

$$r_{d,1} = p_{d,1} - p_{d-1,M}, \quad (3.24)$$

onde $p_{d,1}$ é o log-preço do ativo na primeira cotação do dia d e $p_{d-1,M}$ o log-preço do ativo na última cotação do dia $d - 1$.

O retorno é normalmente mais apropriado que o preço para análises, por várias razões. É a variável de interesse para os investidores que a utilizam como uma medida direta do sucesso de um investimento. É livre de escalas e tem propriedades estatísticas como a estacionariedade e ergodicidade por exemplo.

3.3.1 Retornos para dados de alta frequência

Dados financeiros de alta frequência são importantes em estudos de diversos assuntos ligados ao processo de negociação, da volatilidade realizada e da microestrutura dos mercados financeiros, onde existem grandes movimentos de preços dos ativos em intervalos de tempo menores do que um dia. Podem ser usados para comparar a eficiência de diferentes sistemas de negociação na descoberta de preços, no estudo da dinâmica das ofertas de compra e venda de um particular ativo, na dinâmica de ordens.

Esses dados são normalmente não sincronizados, onde ativos diferentes têm frequências de negociações diferentes e até mesmo para um simples ativo, a intensidade de negociação varia de hora em hora e ao longo dos dias, sendo maior

na abertura e fechamento dos pregões e menor no horário do almoço. Dessa forma, os tempos entre as negociações apresentam um ciclo diário, na forma de “U”.

Para retornos diários de ativos, negociações não sincronizadas podem introduzir:

- (a) Correlação não nula na primeira defasagem dos retornos de uma ação;
- (b) Correlações e correlações cruzadas não nulas na primeira defasagem de uma carteira de ações;
- (c) Em algumas situações, autocorrelações seriais negativas para séries de retorno de uma ação.

3.4 Modelos Autoregressivos Vetoriais

Importante classe de modelos lineares multivariados, os modelos autoregressivos vetoriais de ordem p , serão denotados por VAR(p) (de *vector autoregression*).

Como definido em Morettin (2011, p. 265) e Tsay (2010, p. 399), o processo Y_t , de ordem $n \times 1$, segue um modelo VAR(p) se

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + a_t, \quad (3.25)$$

onde $a_t \sim RB(0, \Sigma)$, $\Phi_0 = (\phi_{10}, \dots, \phi_{n0})'$ é um vetor $n \times 1$ de constantes e Φ_k são matrizes $n \times n$ de coeficientes, com elementos $\phi_{ij}^{(k)}$, $i, j = 1, \dots, n$, $k = 1, \dots, p$. Este é o modelo VAR em sua forma reduzida.

Para o modelo VAR(p) (3.25) ser estacionário devemos ter que as soluções de

$$|I_n - \Phi_1 z - \dots - \Phi_p z^p| = 0$$

devem estar fora do círculo unitário .

3.4.1 Construção de Modelos VAR

A construção do modelo VAR segue os seguintes passos:

- i. Identificação

Para a identificação da ordem p de um modelo VAR(p), uma maneira é o ajuste sequencial de modelos autorregressivos vetoriais de ordem 1, 2, ..., k e o teste de significância dos coeficientes (matrizes). Sejam

$$\begin{aligned} Y_t &= \Phi_0^{(1)} + \Phi_1^{(1)} Y_{t-1} + a_t^{(1)}, \\ Y_t &= \Phi_0^{(2)} + \Phi_1^{(2)} Y_{t-1} + \Phi_2^{(2)} Y_{t-2} + a_t^{(2)}, \\ &\dots \quad \dots \\ Y_t &= \Phi_0^{(k)} + \Phi_1^{(k)} Y_{t-1} + \dots + \Phi_k^{(k)} Y_{t-k} + a_t^{(k)}. \end{aligned} \quad (3.29)$$

Caso a ordem utilizada incluir a ordem verdadeira, o método de MQO pode ser utilizado para a estimação dos parâmetros e fornece estimadores consistentes e eficientes. Testa-se então a significância dos coeficientes:

$$\begin{aligned} H_0: \Phi_k^{(k)} &= 0, \\ H_1: \Phi_k^{(k)} &\neq 0, k = 1, 2, \dots \end{aligned} \quad (3.26)$$

O teste de razão de verossimilhança é baseado nas estimativas das matrizes de covariâncias dos resíduos dos modelos ajustados. O resíduo da k -ésima equação é dado por:

$$\hat{a}_t^{(k)} = Y_t - \hat{\Phi}_0^{(k)} - \hat{\Phi}_1^{(k)} Y_{t-1} - \dots - \hat{\Phi}_k^{(k)} Y_{t-k}.$$

A matriz de covariância dos resíduos, que estima Σ é

$$\hat{\Sigma}_k = \frac{1}{T-k} \sum_{t=k+1}^T \hat{a}_t^{(k)} (\hat{a}_t^{(k)})', k \geq 0 \quad (3.27)$$

onde para $k=0$, $\hat{a}_t^{(0)} = Y_t - \bar{Y}$. A estatística do teste da razão de verossimilhanças para (3.26), que tem distribuição qui-quadrado com n^2 graus de liberdade, $\chi^2(n^2)$, é dada por:

$$RV(k) = (T - k) \ln \frac{|\hat{\Sigma}_{k-1}|}{|\hat{\Sigma}_k|}, \quad (3.28)$$

Uma forma alternativa para determinar a ordem do VAR e através da estimação sequencial dos VAR e escolher aquele modelo que tiver o menor critério de informação, sendo os mais utilizados:

$$AIC(k) = \ln(|\hat{\Sigma}_k|) + 2kn^2/T \quad (\text{Akaike}),$$

$$BIC(k) = \ln(|\hat{\Sigma}_k|) + kn^2 \ln(T)/T \quad (\text{Schwarz}),$$

$$HQC(k) = \ln(|\hat{\Sigma}_k|) + kn^2 \ln(\ln(T))/T \quad (\text{Hannan-Quinn}). \quad (3.29)$$

ii. Estimação

Supondo $a_t \sim N(0, \Sigma)$, os coeficientes podem ser estimados por máxima verossimilhança, e os estimadores de MQ são equivalentes a estimadores de MV condicionais.

Para um VAR(p), os EMV condicionais são obtidos por métodos de maximização numérica. No caso de um VAR(1), maximiza-se

$$l = -\frac{n(T+1)}{2} \ln(2\pi) + \frac{(T-1)}{2} \ln|\Sigma^{-1}| - \frac{1}{2} \sum_{t=2}^T (Y_t - \Phi Y_{t-1})' \Sigma^{-1} (Y_t - \Phi Y_{t-1}), \quad (3.30)$$

obtendo-se

$$\hat{\Phi} = [\sum_{t=2}^T Y_t Y_{t-1}'] [\sum_{t=2}^T Y_{t-1} Y_{t-1}']^{-1} \quad (3.31)$$

$$\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \hat{a}_t (\hat{a}_t)',$$

$$\hat{a}_t = Y_t - \hat{\Phi} Y_{t-1}.$$

iii. Diagnóstico

Os resíduos são utilizados para verificar a validade do modelo. É necessário verificar se a ordem do VAR foi escolhida de forma adequada e para testar esta hipótese usamos a versão multivariada da estatística de Box-Ljung-Pierce, para testar não autocorrelação serial dos resíduos e a estatística de teste é dada por:

$$Q(m) = T^2 \sum_{\tau=1}^m \frac{1}{T-\tau} \text{tr}(\hat{\Gamma}(\tau)' \hat{\Gamma}(0)^{-1} \hat{\Gamma}(\tau) \hat{\Gamma}(0)^{-1}), \quad (3.32)$$

onde esta estatística tem distribuição $\chi^2(n^2(m-p))$.

Deveremos também testar normalidade e homoscedasticidades dos resíduos e isto pode ser feito usando as versões multivariadas do teste de Jarque-Bera e de White respectivamente.

3.5 Causalidade de Granger

O conceito estabelecido por Granger (1969) para sistemas temporais define causalidade em termos de previsibilidade. A variável X causa a variável Y , se os valores passados de X tornam a previsão do presente valor de Y mais eficiente do que se não utilizado esse passado. Formalmente, Y não causa X no sentido de Granger se para todo $s > 0$ o Erro Médio Quadrático (EMQ) da previsão de X_{t+s} usando o conjunto de informação $\{X_t, X_{t-1}, \dots\}$ é o mesmo EMQ da previsão de X_{t+s} usando os conjuntos de informações $\{X_t, X_{t-1}, \dots\}$ e $\{Y_t, Y_{t-1}, \dots\}$.

No teste de causalidade de Granger, a hipótese nula a ser testada é a de que o conjunto dos coeficientes defasados da variável independente são estatisticamente iguais a zero. Se a hipótese nula for rejeitada para o nível de significância escolhido, conclui-se que os coeficientes são diferentes de zero, ou seja, a variável independente causa, no sentido de Granger, a variável dependente.

4 METODOLOGIA

Aplicação

Nesta seção, serão apresentados o conjunto de dados utilizado e a aplicação empírica do modelo *benchmark* proposto por Chaboud et al. (2013) para inferir sobre a correlação das estratégias algorítmicas nos dados de negociação com a utilização da medida R . Serão calculadas também algumas estatísticas que mostram a participação dos algoritmos no volume de negociação do dólar futuro no período. Em seguida, é feita a estimação da autocorrelação serial dos retornos de 5 segundos em intervalos de 5 minutos, para sua utilização na estimação do modelo VAR.

Por apresentarem estratégias pré-programadas e talvez menos diversas e mais correlacionada, os computadores podem reagir da mesma forma, ao mesmo tempo, a mesma informação, criando volatilidade excessiva no mercado, especialmente em alta frequência. Através dos dados de negociação de computadores e humanos, é possível obter algumas informações sobre a correlação das estratégias algorítmicas. A ideia intuitiva é a de que os negociadores que seguem estratégias de negociação e enviam instruções similares e ao mesmo tempo, irão negociar menos entre si, comparado àqueles que seguem estratégias menos correlacionadas.

Outro ponto a ser verificado é em relação a uma das preocupações do mercado de que negociações algorítmicas podem contribuir para o desvio temporário dos preços dos ativos de seus valores fundamentais, resultando em excesso de volatilidade, principalmente em alta frequência. Para tratar essa questão, propõem-se investigar os efeitos que os negociadores algorítmicos têm sobre a autocorrelação dos retornos de alta frequência, mais precisamente retornos de 5 segundos. Será estimado um modelo VAR para identificar a relação entre a autocorrelação dos retornos e as atividades algorítmicas, representadas pelas medidas anteriormente calculadas.

4.1 Dados

Serão utilizados dados de negociação de contratos futuros de reais por dólar (DOL) no mercado futuro da BM&FBOVESPA fornecidos pela Bovespa. Sua

cotação é em reais por US\$1.000,00, com até três casas decimais. Possui vencimentos em todos os meses, que ocorrem no primeiro dia útil do mês. Dentre os derivativos financeiros de taxas de câmbio, este é o contrato futuro mais negociado, com uma média mensal de aproximadamente 7.300.000 contratos negociados.

A base a ser analisada contém apenas o primeiro vencimento de cada mês, dos negócios realizados entre 03 de Janeiro a 27 de Junho de 2013, compreendendo um total de aproximadamente 1.700.000 observações ao longo de 120 dias. Contém informações como data, horário com precisão de segundos, símbolo do instrumento e série, número do negócio, preço, quantidade de contratos, indicador de negócio (1-Ativo / 2-Cancelado), condição da oferta de compra e de venda (0-Oferta Neutra, 1-Oferta Agressora, 2-Oferta Agredida).

Foram utilizados dados de negócio compreendidos no horário normal de negociação, entre 09h e 18h. Foi eliminado o pregão do dia 13 de fevereiro com horário reduzido de negociação (Quarta-feira de cinzas) e também o último dia de negociação de cada mês por apresentarem pouca negociação. Os dados foram organizados em intervalos de um minuto e diários.

Também foi incorporada a base original, informação sobre a direção do negócio, compra ou venda, e um indicador de negociação algorítmica em alta frequência. A BM&FBOVESPA utiliza vários critérios e parâmetros de acordo com o segmento de mercado e tipo de investidor para classificá-lo como *HFT*. Estas informações sobre negociadores que utilizam algoritmos para negociações em alta frequência não são públicas. Então para fins acadêmicos e para realizar o enquadramento do investidor como *HFT*, utilizou-se como critério o volume de contratos negociados no dia e volume de *daytrade*.

4.2 Cálculo da medida R

Para detectar se existe dependência entre as estratégias algorítmicas comparadas àquelas realizadas por não algoritmos, verificam-se as medidas provenientes dos dados ao modelo *benchmark* utilizado em Chaboud et al. (2013), descrito no capítulo 3, que assume que os negociadores atuam de forma aleatória e independente. Essas suposições são adotadas porque os negociadores não são classificados em algorítmicos ou não algorítmicos no processo de combinação dos

compradores e vendedores. Também, os negociadores não conhecem a identidade ou tipo de negociador que está na contraparte até que o negócio seja totalmente realizado. Então o modelo permite o cálculo das probabilidades teóricas dos quatro possíveis tipos de negociadores: (HH) , (CH) , (HC) e (CC) .

Sob as suposições do modelo, espera-se que ocorram mais atividades de negociação realizadas por humanos do que por computadores. Dessa forma calculam-se as razões computador-*taker*, RC e humano-*taker*, RH , e espera-se que sejam maiores do que um. O modelo também prevê que a medida $R = RC/RH = 1$, com computadores tomando liquidez de humanos na mesma proporção que humanos tomando liquidez de outros humanos. Se observarmos a medida $R > 1$ nos dados, indicará que os computadores estão negociando menos entre eles e mais com os humanos do que o modelo *benchmark* prevê. Portanto, uma razão maior do que um pode ser visto como uma evidência de que computadores possuem estratégias de negociação que são mais dependentes do que as estratégias realizadas por não algoritmos.

No exemplo acima o valor de R acaba sendo alto porque, ao atuarem de forma similar, algoritmos enviam ao mesmo tempo instruções para comprar, por exemplo, e dessa forma não irão se encontrar no livro de ordens, por atuarem na mesma direção, resultando em uma baixa frequência de CC . Por outro lado, se metade dos negociadores algorítmicos está comprando e metade está vendendo, terão uma alta probabilidade de se encontrarem e será observado uma alta frequência de CC e um valor baixo de R .

Também é calculada a razão computador-*taker* comprando, RC^B , a razão computador-*taker* vendendo, RC^S , a razão humano-*taker* comprando, RH^B e a razão humano-*taker* vendendo, RH^S . Assim como para a medida R , o modelo prevê que cada uma das quatro razões será maior do que um e as razões das razões $R^B \equiv \frac{RC^B}{RH^B}$ e $R^S \equiv \frac{RC^S}{RH^S}$ serão ambas iguais a um.

Utilizando-se o conjunto de dados, calcularam-se as medidas R , R^S e R^B nas frequências de 1 minuto e diário. O valor realizado de RH é dado por $\widehat{RH} = \frac{Vol(HH)}{Vol(CH)}$ e de RC é $\widehat{RC} = \frac{Vol(HC)}{Vol(CC)}$. De forma similar, tem-se $\widehat{RH^S} = \frac{Vol(HH^S)}{Vol(CH^S)}$, $\widehat{RH^B} = \frac{Vol(HH^B)}{Vol(CH^B)}$,

$$\widehat{RC^S} = \frac{Vol(HC^S)}{Vol(CC^S)} \text{ e } \widehat{RC^B} = \frac{Vol(HC^B)}{Vol(CC^B)}.$$

Por fim calculou-se o logaritmo natural das razões, $\ln \hat{R} = \ln(\frac{\widehat{RC}}{\widehat{RH}})$, $\ln \widehat{R^S} = \ln(\frac{\widehat{RC^S}}{\widehat{RH^S}})$ e $\ln \widehat{R^B} = \ln(\frac{\widehat{RC^B}}{\widehat{RH^B}})$ para cada frequência. Dessa forma, as previsões do modelo *benchmark* serão equivalentemente iguais a $\ln \hat{R} \equiv 0$, $\ln \widehat{R^B} \equiv 0$ e $\ln \widehat{R^S} \equiv 0$.

A hipótese a ser testada é se os negociadores algorítmicos utilizam um conjunto de estratégias menos diversa do que os negociadores não algorítmicos.

A Tabela 1 apresenta as médias do log natural das razões R para as frequências de 1 minuto e diária. As significâncias estatísticas para os níveis de 0,01, 0,05 e 0,1 foram representadas por ***, ** e * respectivamente.

Tabela 1 – Dependência entre estratégias de negociação algorítmica utilizando a medida R .

	Frequência 1 minuto			Frequência diária		
	$\ln(\hat{R})$	$\ln(\widehat{R^S})$	$\ln(\widehat{R^B})$	$\ln(\hat{R})$	$\ln(\widehat{R^S})$	$\ln(\widehat{R^B})$
Média	0,2364***	-0,0155	0,0209	0,0489**	0,0619*	0,0390***
Erro padrão	0,0116	0,0171	0,0143	0,0229	0,0338	0,0147
Desvio padrão	1,9706	2,1411	1,7168	0,2453	0,3611	0,1573
Fração de obs. > 0	0,5561	0,4947	0,5018	0,5789	0,5964	0,6403
Nº de obs. não faltantes	28.674	15.645	14.388	114	114	114
Total de obs.	60.722	60.722	60.722	114	114	114

Observa-se que para a frequência de 1 minuto, $\ln \hat{R}$ é significativamente maior do que zero para todos os níveis de significância. Para a frequência diária, $\ln(\hat{R})$ é significativo ao nível de 0,05 e 0,1, $\ln(\widehat{R^S})$ ao nível de 0,1 e $\ln(\widehat{R^B})$ é significativo para todos os níveis de significância. Então, os resultados mostram que computadores não negociam entre eles de forma tão aleatória quanto o modelo propõe. Isto é uma evidência de que as estratégias utilizadas pelos negociadores algorítmicos são menos diversas do que as estratégias de negociação utilizadas pelos negociadores não algorítmicos.

A Tabela 1 também apresenta o número de períodos em que as estatísticas são maiores do que zero. A proporção de valores faltantes é maior em frequências menores porque o número de observações em que $Vol(CC) = 0$ é maior. Isso

provavelmente ocorrer porque existe pouca negociação algorítmica nesse intervalo de tempo. Outra possibilidade seria uma atividade algorítmica abundante, mas com estratégias tão similares que a combinação CC não ocorre.

4.3 Atividades de negociações algorítmicas.

Foram calculadas algumas variáveis utilizando os dados de negociação minuto a minuto, onde são fornecidas informações como volume de negociação, direção, qual negociador iniciou o negócio atribuído a algoritmos e não algoritmos. Dessa forma é possível estudar não apenas o impacto geral da presença de algoritmos de negociação no mercado, mas também o mecanismo pelo qual as negociações algorítmicas afetam o mercado.

A participação geral dos algoritmos no mercado é representada pela estatística *VAT*, que é fração do volume em que eles atuam em pelo menos uma das partes do negócio. A fração do volume em que computadores negociam agredindo uma oferta existente é representada pela medida *VCt*. Para a situação oposta, em que a oferta colocada pelo computador é agredida, é o *VCm*. A variável que representa a participação algorítmica no fluxo de ordem do mercado é *OFCt*.

A Tabela 2 apresenta as estatísticas descritivas para a participação dos algoritmos de negociação no mercado para a frequência de 1 minuto.

Tabela 2 – Estatísticas da participação dos negociadores algorítmicos no volume negociado no mercado futuro e dólar, minuto a minuto.

Estatísticas	Tipo de Atividade		
	<i>VAT</i>	<i>VCt</i>	<i>VCm</i>
Média	67,1601	39,4661	46,7689
Desvio Padrão	31,8135	33,9048	31,7100
Erro Padrão	0,1319	0,1405	0,1314
Mínimo	0	0	0
5%	0	0	0
25%	45,0000	5,2631	19,6428
50%	76,8116	33,7588	45,1613
75%	95,8647	68,9922	73,0159
95%	100	98,0392	100
Máximo	100	100	100

Os resultados mostram que os negociadores algorítmicos estão envolvidos em 67,2% de todo o volume negociado no mercado futuro de dólar para a amostra no

período (VAT). Eles demandam liquidez (VCt) em 39,47% de todo o volume negociado e fornecem liquidez (VCm) em 46,77%.

Essas participações estão representadas na Figura 1.

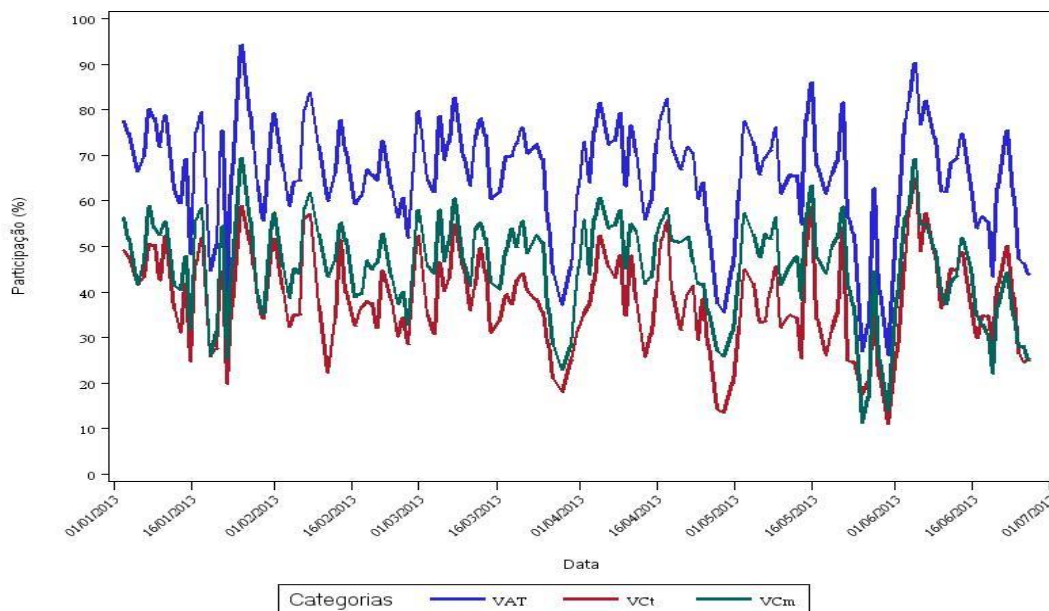


Figura 1 – Participação dos negociadores algorítmicos no volume de negociação por minuto no mercado futuro de dólar.

A participação média dos algoritmos no fluxo de ordem do mercado (OFCt), que é a medida da intensidade direcional relativa dos negócios iniciados por computadores, é de 39,63%.

A Tabela 3 apresenta as estatísticas descritivas para a participação dos algoritmos de negociação no mercado para a frequência diária.

Tabela 3 – Estatísticas da participação dos negociadores algorítmicos no volume diário negociado no mercado futuro de dólar.

Estatísticas	Tipo de Atividade		
	VAT	VCt	VCm
Média	68,7093	43,8297	45,7919
Desvio Padrão	13,3228	11,3132	11,4550
Erro Padrão	1,2423	1,0549	1,0682
Mínimo	30,0989	10,3704	18,6460
5%	40,5772	22,5775	21,9193
25%	61,0474	36,5126	38,2619
50%	71,5519	48,1559	48,1559
75%	78,4043	52,3083	54,1949
95%	85,6220	61,2018	62,2250
Máximo	92,0064	64,6128	67,5840

Observa-se que as participações médias dos negociadores algorítmicos é maior na frequência diária para o *VAT* e *VCt*. Eles participam de quase 69% de todo o volume negociado no mercado futuro de dólar. Demandam liquidez em 43,83% de todo o volume de dólar futuro negociado e fornecem liquidez em 45,80%.

As participações diárias estão representadas na Figura 2.

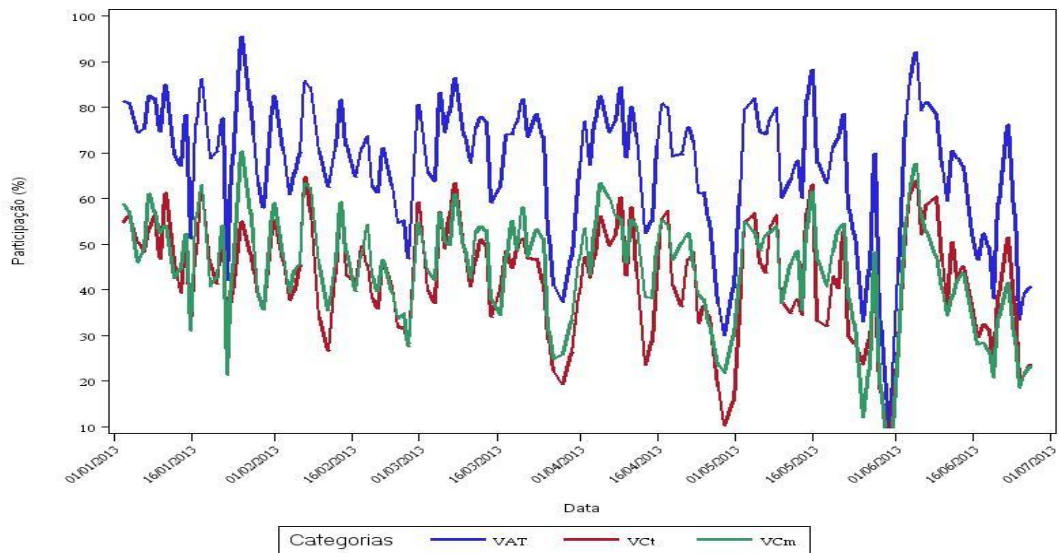


Figura 2 – Participação dos negociadores algorítmicos no volume de negociação diária minuto no mercado futuro de dólar.

A participação média diária dos algoritmos no fluxo de ordem do mercado (*OFCt*) é de 36,43%.

4.4 Modelo VAR

Para a estimação do modelo VAR, além das estatísticas de negociação algorítmicas, também foi estimada a autocorrelação serial de primeira ordem dos retornos de 5 segundos calculada a cada intervalo de 5 minutos (*AC5_Retorno*). A escolha para o cálculo em intervalos de 5 minutos foi baseado no fato de que esta janela é grande o bastante com observações suficientes para estimar a autocorrelação e pequena o bastante para que o teste de causalidade de Granger seja sensível.

A negociação algorítmica é medida em frequência de um minuto nas cinco formas distintas apresentadas nas sessões anteriores: a fração do volume total de negócios que envolvem um computador em pelo menos um lado do negócio (*VAT*), a atividade tomadora relativa de computadores (*VCt*), a atividade provedora relativa de

computadores (VCM), a atividade tomadora relativa na mesma direção de computadores OFC_t e o grau de dependência na atividade de negociação de computadores ($\ln(\hat{R})$).

Para capturar as tendências de tempo de longo e curto prazo, foram criadas variáveis *dummies* mês-ano para cada mês de negociação ($M1, M2, M3, M4, M5$ e $M6$), e variáveis *dummies* meia-hora para cada dia de negociação ($MIN1_{30}, MIN2_{30}, MIN3_{30}, MIN4_{30}, MIN5_{30}, MIN6_{30}, MIN7_{30}, MIN8_{30}, MIN9_{30}, MIN10_{30}, MIN11_{30}, MIN12_{30}, MIN13_{30}, MIN14_{30}, MIN15_{30}, MIN16_{30}, MIN17_{30}$ e $MIN18_{30}$).

Seja $|SAC_t|$ a medida de cinco minutos do valor absoluto da autocorrelação serial sobre retornos de 5 segundos expressa em porcentagem, e AT representando as cinco medidas de atividade de negociação algorítmica. Definindo-se $Y_t = (|SAC_t|, AT_t)$, a forma reduzida é definida como

$$Y_t = \Phi(L)Y_t + \Lambda X_{t-1:t:20} + \psi G_t + \varepsilon_t \quad (4.1)$$

onde $X_{t-1:t:20}$ inclui a soma do volume de negócios nos últimos 20 minutos ($X20_Qtd$) e a volatilidade dos últimos 20 minutos, calculado como a somados retornos absolutos desses 20 minutos ($X20_Retorno$). G_t representa um conjunto de funções determinísticas de tempo t , capturando tendências individuais e padrões intradiários nas variáveis em AY_t . Em particular,

$$G_t = (M_1, \dots, M_6, MIN1_{30}, \dots, MIN18_{30})$$

capturando tendências seculares de longo prazo nos dados pelas variáveis *dummies* de mês, e também captura os padrões intradiários pelas variáveis *dummies* de meia-hora. É assumido que os *shocks* estruturais ε_t são independentes entre si e não correlacionados serialmente em todos as defasagens.

O VAR na sua forma reduzida é estimado equação por equação utilizando mínimos quadrados ordinários e o teste de causalidade de Granger é utilizado para avaliar o papel dos algoritmos de negociação. Para isso é verificado se todos os coeficientes das defasagens da variável causadora são iguais a zero.

Foi estimado um modelo VAR em sua forma reduzida conforme a equação (4.1), utilizando uma frequência de 5 minutos. Para verificar se as atividades de negociação algorítmicas tem um impacto causal sobre a autocorrelação e vice-versa,

a estimação foi feita para cada uma das variáveis AT (VAT , VCt , VCm , $OFCt$ e $\ln(\hat{R})$). A seguir são apresentados os resultados para os modelos estimados.

4.4.1 Estimação do modelo VAR em sua forma reduzida para as variáveis endógenas $AC5_Retorno$ e AT (VAT , VCt , VCm , $OFCt$ e $\ln(\hat{R})$).

Quatro modelos foram estimado para cada uma das variáveis AT (VAT , VCt , VCm , $OFCt$ e $\ln(\hat{R})$):

O modelo 1 considera as variáveis endógenas $AC5_Retorno$, AT e com variável exógena sendo somente a constante.

O modelo 2, as variáveis endógenas $AC5_Retorno$, AT e com variáveis exógenas sendo as *dummies* dos meses ($M1$ a $M6$).

O modelo 3, as variáveis endógenas $AC5_Retorno$, AT e com variáveis exógenas sendo as *dummies* dos meses ($M1$ a $M6$), as variáveis $X20_Retorno$ e $X20_Qtd$.

O modelo 4, as variáveis endógenas $AC5_Retorno$, AT e as variáveis exógenas sendo as *dummies* dos meses ($M1$ a $M6$) e a variável $X20_Retorno$.

As variáveis *dummies* para cada meia-hora de negociação ($MIN1_30$, $MIN2_30$, ..., $MIN18_30$) não foram utilizadas pois os dados estão em frequência de 5 minutos e portanto não seriam adequadas.

Na Tabela 4 são apresentados os resultados para os testes de causalidade de Granger e exogeneidade em blocos de Wald para verificar se as atividades de negociações algorítmicas tem um impacto causal sobre a autocorrelação dos retornos, no lado esquerdo da tabela, e se a autocorrelação tem um impacto causal sobre as atividades de negociações algorítmicas, no lado direito da tabela. Os resultados estão dispostos separadamente para cada uma das medidas das atividades de negociações algorítmicas AT (VAT , VCt , VCm , $OFCt$ e $\ln(\hat{R})$).

Tabela 4 – Testes de causalidade de Granger para a autocorrelação dos retornos e as atividades algorítmicas.

Modelo	Variável dependente: AC5_RETORNO					Variável dependente: LN_R				
	Excluído	Soma coef.	Chi-sq	df	Prob.	Excluído	Soma coef.	Chi-sq	df	Prob.
1	LN_R	-0,04472	3,371163	7	0,8487	AC5_RETORNO	0,00027	1,819004	7	0,9692
	All		3,371163	7	0,8487	All		1,819004	7	0,9692
2	LN_R	-0,03503	3,358889	7	0,8499	AC5_RETORNO	0,00022	1,821531	7	0,9691
	All		3,358889	7	0,8499	All		1,821531	7	0,9691
3	LN_R	-0,07585	3,495567	7	0,8357	AC5_RETORNO	0,00022	1,844870	7	0,9679
	All		3,495567	7	0,8357	All		1,844870	7	0,9679
4	LN_R	-0,06095	3,428147	7	0,8428	AC5_RETORNO	0,00010	1,805419	7	0,9698
	All		3,428147	7	0,8428	All		1,805419	7	0,9698
Modelo	Variável dependente: AC5_RETORNO					Variável dependente: VAT				
	Excluído	Soma coef.	Chi-sq	df	Prob.	Excluído	Soma coef.	Chi-sq	df	Prob.
1	VAT	0,01998	30,21743	8	0,0002	AC5_RETORNO	0,12391	22,53114	8	0,0040
	All		30,21743	8	0,0002	All		22,53114	8	0,0040
2	VAT	0,02170	30,99463	8	0,0001	AC5_RETORNO	0,12689	22,50115	8	0,0041
	All		30,99463	8	0,0001	All		22,50115	8	0,0041
3	VAT	0,01698	25,75939	8	0,0012	AC5_RETORNO	0,10696	21,57712	8	0,0058
	All		25,75939	8	0,0012	All		21,57712	8	0,0058
4	VAT	0,01991	29,47785	8	0,0003	AC5_RETORNO	0,12136	22,14931	8	0,0046
	All		29,47785	8	0,0003	All		22,14931	8	0,0046
Modelo	Variável dependente: AC5_RETORNO					Variável dependente: VCt				
	Excluído	Soma coef.	Chi-sq	df	Prob.	Excluído	Soma coef.	Chi-sq	df	Prob.
1	VCt	0,03190	39,82110	8	0,0000	AC5_RETORNO	0,14185	23,05763	8	0,0033
	All		39,82110	8	0,0000	All		23,05763	8	0,0033
2	VCt	0,03036	37,83570	8	0,0000	AC5_RETORNO	0,13051	22,03670	8	0,0048
	All		37,83570	8	0,0000	All		22,03670	8	0,0048
3	VCt	0,02336	29,18364	8	0,0003	AC5_RETORNO	0,09767	18,99051	8	0,0149
	All		29,18364	8	0,0003	All		18,99051	8	0,0149
4	VCt	0,00028	34,68190	8	0,0000	AC5_RETORNO	0,11953	20,92831	8	0,0073
	All		34,68190	8	0,0000	All		20,92831	8	0,0073

Tabela 4 – Testes de causalidade de Granger para a autocorrelação dos retornos e as atividades algorítmicas.

(conclusão)

Modelo	Variável dependente: AC5_RETORNO					Variável dependente: VCm				
	Excluído	Soma coef.	Chi-sq	df	Prob.	Excluído	Soma coef.	Chi-sq	df	Prob.
1	VCm	0,01290	16,20044	7	0,0233	AC5_RETORNO	0,08886	14,01907	7	0,0508
	All		16,20044	7	0,0233	All		14,01907	7	0,0508
2	VCm	0,01601	17,32234	7	0,0154	AC5_RETORNO	0,10301	14,86026	7	0,0378
	All		17,32234	7	0,0154	All		14,86026	7	0,0378
3	VCm	0,01426	16,53286	7	0,0207	AC5_RETORNO	0,09568	14,55361	7	0,0422
	All		16,53286	7	0,0207	All		14,55361	7	0,0422
4	VCm	0,01489	17,11966	7	0,0166	AC5_RETORNO	0,10264	14,83899	7	0,0381
	All		17,11966	7	0,0166	All		14,83899	7	0,0381
Modelo	Variável dependente: AC5_RETORNO					Variável dependente: OFCt				
	Excluído	Soma coef.	Chi-sq	df	Prob.	Excluído	Soma coef.	Chi-sq	df	Prob.
1	OFCt	0,02584	29,11491	4	0,0000	AC5_RETORNO	0,19296	17,83734	4	0,0013
	All		29,11491	4	0,0000	All		17,83734	4	0,0013
2	OFCt	0,02366	25,68058	4	0,0000	AC5_RETORNO	0,17708	15,27923	4	0,0042
	All		25,68058	4	0,0000	All		15,27923	4	0,0042
3	OFCt	0,01755	17,58736	4	0,0015	AC5_RETORNO	0,14460	11,48186	4	0,0217
	All		17,58736	4	0,0015	All		11,48186	4	0,0217
4	OFCt	0,02180	22,96008	4	0,0001	AC5_RETORNO	0,16502	13,69806	4	0,0083
	All		22,96008	4	0,0001	All		13,69806	4	0,0083

Fonte: Elaboração própria com base nos cálculos efetuados no pacote econométrico EvIEWS 6.0.

Considerando um nível de significância de 5%, observa-se na Tabela 4 que no modelo em que a variável $\ln(\hat{R})$ é endógena, a hipótese nula de que $\ln(\hat{R})$ não causa AC5_Retorno no sentido de Granger não é rejeitada. Também é verificado o mesmo resultado para o caso da avaliação conjunta dos coeficientes. Quando a variável dependente é $\ln(\hat{R})$, AC5_Retorno também não causa $\ln(\hat{R})$ no sentido de Granger.

Portanto, nos modelos estimados para verificar se o grau de correlação na atividade de negociação de computadores ($\ln(\hat{R})$) tem um impacto causal sobre a autocorrelação dos retornos (AC5_Retorno) e vice-versa, não foram apresentadas evidências estatísticas de causalidade unidirecional e bidirecional.

Para as variáveis *AT* (*VAT*, *VCt* e *OFCt*), a hipótese nula é rejeitada e, portanto a variável *AT* causa *AC5_Retorno* no sentido de Granger. É verificado o mesmo resultado para o caso da avaliação conjunta dos coeficientes. Quando a variável dependente é uma das variáveis *AT*, *AC5_Retorno* também causa *AT* no sentido de Granger.

Portanto, nos modelos estimados para verificar se a variável *AT* (*VAT*, *VCt*, e *OFCt*), tem um impacto causal sobre a autocorrelação dos retornos (*AC5_Retorno*), e vice-versa, foram apresentadas evidências estatísticas de que *AT* causa *AC5_Retorno* e também *AC5_Retorno* causa *AT*, para as especificações e nos dois sentidos.

Para a variável *AT* (*VCm*), a hipótese nula de que a variável *AT* não causa *AC5_Retorno* no sentido de Granger tem p-valor de aproximadamente 2%. Já a hipótese nula de que a variável *AC5_Retorno* não causa *AT* no sentido de Granger tem p-valor de varia entre 3,8% e 5%. Nestes casos os testes não são conclusivos.

5 CONSIDERAÇÕES FINAIS

Neste trabalho foi apresentada uma metodologia para estudar o impacto das negociações algorítmicas no mercado de dólar futuro, no que se refere a diversidade de suas estratégias.

Foram calculadas as razões para os negociadores algorítmicos e não algorítmicos, quando estes estão tirando liquidez do mercado. A partir dessas razões, foram feitas inferências tendo por base um modelo *benchmark* proposto por Chaboud et al. (2013), que estabelece a existência de mais negociadores humanos que computadores, e as combinações entre compradores e vendedores é independente e aleatória e, portanto, a medida R que é a razão dessas razões, é igual a 1.

Os resultados obtidos para as negociações minuto a minuto para a razão R apresentam evidências de que as ações e estratégias de negociadores algorítmicos parecem ser menos diversas e mais dependentes e similares do que aquelas realizadas por negociadores não algorítmicos. Isso pode indicar que computadores estão negociando menos entre eles e mais com os humanos comparado ao que o modelo *benchmark* prevê. Este resultado é consistente com aqueles obtidos por Chaboud et al. (2013) que calculou a medida R para o mercado global de câmbio representada pelas três taxas mais negociadas, dólar, euro e iene, e Brogaard (2010) que estudou os impactos dos negociadores de alta frequência no mercado de ações americano.

Dependendo das condições do mercado, essa característica apresentada pelos algoritmos pode causar efeitos negativos sobre os preços e potencialmente volatilidade excessiva, por atuarem coletivamente de forma similar, aumentando a movimentação do mercado. Segundo Kirilenko et al. (2011), esta pode ter sido uma das razões que agravaram o episódio do *Flash Crash*, quando os algoritmos deixaram o mercado ao mesmo tempo.

Também são apresentadas as participações dos algoritmos de negociação no volume de dólar futuro negociado no período. A participação geral média chega a quase 70% do volume total. Sua participação como provedor de liquidez representa em média 46% do volume total, enquanto sua atuação demandando liquidez alcança aproximadamente 40%. Essas participações também podem ter sido de alguma forma influenciadas pela alta volatilidade nos preços do dólar no período de análise.

No artigo de Chaboud et al. (2013), no final do período analisado, de setembro de 2003 a dezembro de 2007, a participação geral média para as negociações de euro-dólar e dólar-iene chegou a quase 60% enquanto para euro-iene alcançou aproximadamente 80%. A participação dos algoritmos como provedores de liquidez representa em média 30% para as taxas de câmbio USD/EUR e JPY/USD e 45% para JPY/EUR. Já sua atuação demandando liquidez alcança aproximadamente 35% para USD/EUR e JPY/USD e 65% para JPY/EUR.

Brogaard (2010) também analisou a participação dos algoritmos no mercado de ações e verificou que eles estão envolvidos em 68,5% de todo o volume diário negociado na amostra. Demandaram liquidez em 42,7% de todo o volume negociado e fornecem liquidez em 41,1%.

Como descrito no artigo de Jarrow e Protter (2012), uma das preocupações do mercado é que as negociações algorítmicas podem contribuir para o desvio temporário dos preços dos ativos de seus valores fundamentais, resultando em excesso de volatilidade, principalmente em alta frequência. Para tratar essa questão, investigou-se os efeitos que os negociadores algorítmicos têm sobre a autocorrelação dos retornos de alta frequência, mais precisamente retornos de 5 segundos. Em seguida foi estimado um modelo VAR para identificar a relação entre a autocorrelação dos retornos e as atividades algorítmicas, representadas pelas medidas anteriormente calculadas. Foi estudado o impacto geral da participação das atividades algorítmicas no mercado, VAT , a atividade relativa como tomador de liquidez, VCt e também como provedor de liquidez ao mercado, VCm , a presença relativa de fluxo de ordem de computadores tomadores de liquidez, $OFCt$, e o $\ln(R)$.

Com a estimação do modelo VAR, observou-se que, com exceção do grau de correlação na atividade de negociação de computadores ($\ln(\hat{R})$) e a participação dos algoritmos de negociação como provedores de liquidez ao mercado (VCm), existe uma relação de causalidade entre a autocorrelação dos retornos e as atividades algorítmicas nos dois sentidos. Um aumento nas atividades algorítmicas tende a aumentar o valor absoluto da autocorrelação, mostrando um indício de que as atividades algorítmicas podem contribuir para o excesso de volatilidade no mercado.

Este resultado acabou contrastando com aquele encontrado no artigo de Chaboud et al. (2013), onde foram encontradas evidências de que negociações

algorítmicas melhoram a eficiência informacional ao acelerarem a descoberta de preços. Não foram encontrados padrões claros para o impacto das autocorrelações absolutas mais altas sobre as atividades de negociações algorítmicas.

Alguns pontos como o período de análise menor, tendência de alta na cotação do dólar ao longo do período e mercados diferentes podem ter sido responsáveis pelas diferenças nos resultados dos trabalhos.

Existem muitos trabalhos desenvolvidos no âmbito internacional para estudar os algoritmos de negociação e sua atuação. No Brasil, não foram identificados trabalhos que tratem especificamente das negociações algorítmicas no mercado brasileiro. Este é um assunto amplo e muitas frentes de pesquisa podem ser desenvolvidas para melhor entender o comportamento da negociação algorítmica e sua influência no mercado.

No mercado brasileiro, por exemplo, sua atuação é bastante frequente nas negociações de contrato futuro de dólar padrão e mini, buscando a arbitragem entre eles. O mesmo ocorre para contratos futuros de índice padrão e o mini. Seria um assunto interessante a ser estudado. Outro tópico bastante importante, principalmente para a fiscalização das negociações, é tentar entender as estratégias utilizadas pelos algoritmos, que nem sempre são vistas como boas práticas no mercado.

REFERÊNCIAS

ALDRIDGE, Irene. *High frequency trading: a practical guide to algorithmic strategies and trading systems*. 2nd ed. New Jersey, Wiley-Interscience, 2010.

BIAIS, Bruno; FOUCAULT, Thierry; MOINAS Sophie. *Equilibrium Algorithmic Trading*. Working Paper, TSE and HEC, 2011.

BIAIS, Bruno; WOOLLEY, Paul. *High Frequency Trading*. Manuscript, Toulouse University, IDEI, 2011.

BROGAARD, Jonathan. *High Frequency Trading and its Impact on Market Quality*. Northwestern University Kellogg School of Management Working Paper, 2010.

BROGAARD, Jonathan. *High frequency trading and volatility*. University of Washington Working paper, 2011.

COMMODITY FUTURES TRADING COMMISSION (CFTC). *Concept Release on Risk Controls and System Safeguards for Automated Trading Environments*, 2013. Disponível em: <http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/tac103012_reference.pdf>. Acesso em: 10 Set. 2013.

CHABOUD, Alain, et al. *Rise of the machines: Algorithmic trading in the foreign exchange market*. Journal of Finance, Forthcoming (2013).

DACOROGNA, Michael M. et al. *An Introduction to High-Frequency Finance*, Academic Press, California, 2001.

EASLEY, David; PRADO, Marcos Lopez de; O'HARA, Maureen. *The microstructure of the 'flash crash': flow toxicity, liquidity crashes and the probability of informed trading*. Journal of Portfolio Management 37(2):118-128, 2011.

GRANGER, Clive WJ. *Investigating causal relations by econometric models and cross-spectral methods*. Econometrica: Journal of the Econometric Society, 37, 424-438, 1969.

HASBROUCK, Joel; SAAR, Gideon. *Low-latency trading*. Journal of Financial Markets, V. 16, N. 4, P. 646-679, 2013.

HENDERSHOTT, Terrence; JONES, Charles M.; MENKVELD, Albert J. *Does Algorithmic Trading Improve Liquidity?*, The Journal of Finance, v. 66, n. 1, p. 1-33,

2011.

HENDERSHOTT, Terrence; RIORDAN, Ryan. *Algorithmic Trading and Information*, Manuscript, University of California, Berkeley, 2011.

HENDERSHOTT, Terrence; RIORDAN, Ryan. *High frequency trading and price discovery*. Manuscript, University of California, Berkeley, v.3, 2011.

JARROW, Robert A.; PROTTER, Philip. *A dysfunctional role of high frequency trading in electronic markets*. International Journal of Theoretical and Applied Finance, v. 15, n. 03, 2012.

KIRILENKO, Andrei et al. *The flash crash: The impact of high frequency trading on an electronic market*. Manuscript, University of Maryland, 2011.

MORETTIN, Pedro A., *Econometria Financeira*. Segunda ed., Editora Blucher, São Paulo, 2011.

TSAY, Ruey S. *Analysis of financial time series*. John Wiley & Sons, 2005.

ZIVOT, Eric. (2005). *Analysis of High Frequency Financial Data: Methods, Models and Software*. 11th School of Time Series and Econometrics. Vila Velha, August 2005. Brazilian Statistical Association.

APÊNDICE A – Análise descritiva dos log-retornos intradiários.

Tabela A1 – Estatística descritiva dos log-retornos intradiários.

Estatísticas Descritivas	
Média	0,0000051
Erro padrão	0,000006
1º Quartil	-0,000249
3º Quartil	0,000250
Mediana	0,00000
Moda	0,00000
Desvio padrão	0,000638
Variância	0,00000
Curtose	54,258146
Assimetria	-1,031645
Mínimo	-0.015789
Máximo	0,008067
Soma	0,063622
Contagem	12.426

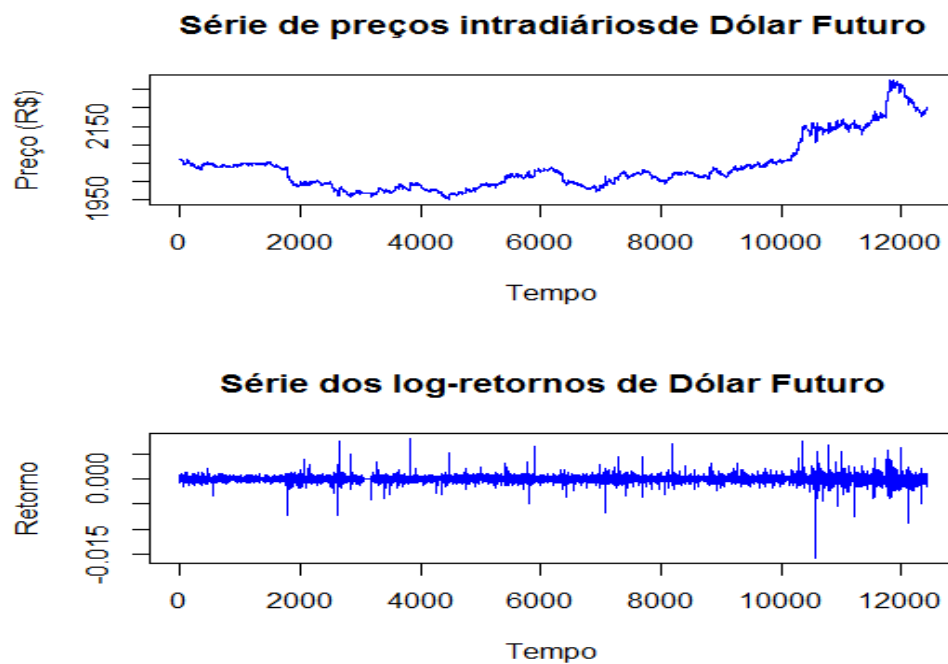


Figura A1- Série de preços e log-retornos intradiários de Dólar Futuro.

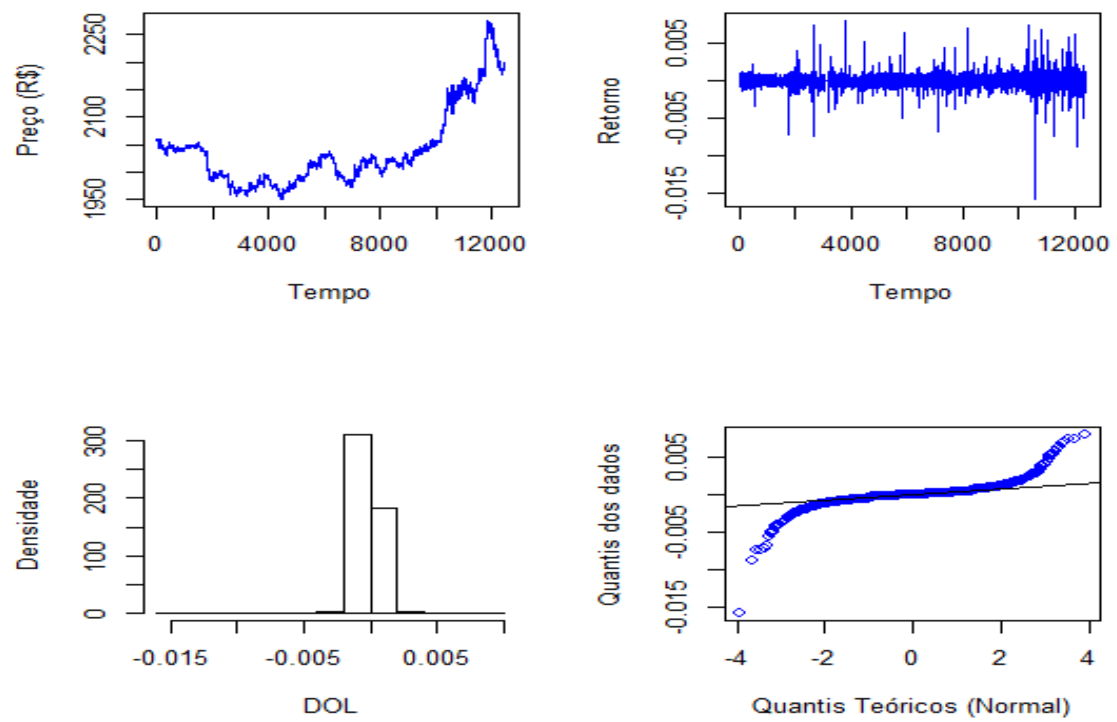


Figura A2- Série de preços e log-retornos intradiários, histograma e Q-Q plot dos log-retornos de Dólar Futuro.

APÊNDICE B – Modelo VAR em sua forma reduzida para as variáveis endógenas *AC5_Retorno* e *AT* (*VAT*, *VCt*, *VCm*, *OFCt* e $\ln(\hat{R})$) e as variáveis exógenas representadas pelas *dummies* dos meses (*M1* a *M6*), as variáveis *X20_Retorno* e *X20_Qtd*.

Vamos apresentar o VAR estimado para o primeiro modelo na especificação 3, isto é as variáveis endógenas são *AC5_Retorno* e $\ln(\hat{R})$ e as variáveis exógenas as *dummies* dos meses, *X20_Retorno* e *X20_Qtd*.

Vamos apresentar também os testes de diagnósticos para os resíduos, a saber, autocorrelação, normalidade e homoscedasticidade.

Tabela B1 - Estimação do VAR Modelo 1 especificação 3

Estimação do VAR		
	AC5_Retorno	LN_R
AC5_Retorno(-1)	0.050834 (0.00975) [5.21194]	0.048045 (0.10925) [0.43977]
AC5_Retorno(-2)	0.044171 (0.00976) [4.52606]	0.054478 (0.10932) [0.49835]
AC5_Retorno(-3)	0.045544 (0.00975) [4.67089]	-0.104108 (0.10922) [-0.95321]
AC5_Retorno(-4)	0.050265 (0.00977) [5.14446]	-0.012999 (0.10944) [-0.11878]
AC5_Retorno(-5)	0.038503 (0.00979) [3.93327]	-0.013630 (0.10965) [-0.12431]
AC5_Retorno(-6)	0.034008 (0.00980) [3.47073]	0.077100 (0.10975) [0.70248]
AC5_Retorno(-7)	0.031321 (0.00982) [3.19030]	-0.027250 (0.10997) [-0.24780]
LN_R(-1)	0.000596 (0.00087) [0.68888]	0.000907 (0.00970) [0.09354]
LN_R(-2)	-0.000825 (0.00087) [-0.95138]	0.003511 (0.00972) [0.36126]
LN_R(-3)	-0.000459 (0.00087) [-0.52959]	0.023712 (0.00971) [2.44160]

LN_R(-4)	-0.000624 (0.00087) [-0.71912]	-0.002347 (0.00972) [-0.24139]
LN_R(-5)	-0.000401 (0.00087) [-0.46398]	0.014046 (0.00969) [1.44944]
LN_R(-6)	4.59E - 05 (0.00086) [0.05311]	0.033570 (0.00967) [3.47102]
LN_R(-7)	0.000910 (0.00086) [1.05374]	0.002562 (0.00967) [0.26494]
M1	0.092723 (0.00441) [21.0313]	0.070272 (0.04938) [1.42298]
M2	0.084396 (0.00452) [18.6811]	0.091458 (0.05060) [1.80734]
M3	0.078360 (0.00421) [18.6153]	0.047803 (0.04715) [1.01382]
M4	0.072828 (0.00407) [17.9062]	0.052915 (0.04556) [1.16150]
M5	0.068488 (0.00409) [16.7457]	0.087681 (0.04581) [1.91395]
M6	0.074879 (0.00481) [15.5651]	0.097535 (0.05389) [1.81005]
X20_RETORNO	-0.501572 (1.29771) [-0.38651]	14.66898 (14.5359) [1.00915]
X20_Qtd	4.37E - 06 (5.6E - 07) [7.85695]	-3.25E - 06 (6.2E - 06) [-0.52177]
R^2	0.032930	0.002595
\bar{R}^2	0.030998	0.000603
Soma Quad. Res.	157.3105	19737.32
<i>S.E. equation</i>	0.122313	1.370060
<i>F-statistic</i>	17.04977	1.302491
Log Verossim.	7199.666	-18257.96
AIC	-1.362374	3.469671
SC	-1.347210	3.484835
Média Var Dep.	0.125834	0.098110
D.P. Var. Dep	0.124254	1.370473
Det. Cov Res. (ajust. g.l.)		0.02881
Det. Cov. Res.		0.027963
AIC		2.107246

SC	2.137573
Notas: Amostra 1/03/2103 9:00 até 6/27/2013 17:55	
Número de Observações: 10.537	
Erro Padrão Entre ()	
Estatísticas t Entre []	
SC – Critério de Schwarz	
AIC - Critério de Akaike	
\bar{R}^2 em o R^2 ajustado por g.l.	

A seguir vamos apresentar os testes de diagnóstico para os resíduos

Primeiro o teste Ljung-Box de Autocorrelação onde a hipótese nula é de não autocorrelação nos resíduos até a defasagem h.

Tabela B2 – Teste de Autocorrelação

Defasagens	Estat. Q	p-valor	Estat. Q adj	p-valor	g.l
1	0.051112	NA*	0.051117	NA*	NA*
2	0.150555	NA*	0.150579	NA*	NA*
3	0.459181	NA*	0.459293	NA*	NA*
4	0.838198	NA*	0.838453	NA*	NA*
5	1.023210	NA*	1.023553	NA*	NA*
6	2.234950	NA*	2.235984	NA*	NA*
7	2.688451	NA*	2.689786	NA*	NA*
8	8.236891	0.0803	8.332511	0.0801	4
9	12.01028	0.1507	12.01905	0.1504	8
10	21.19515	0.0476	21.21264	0.0474	12
Notas: O teste é válido para defasagens superiores a ordem do VAR					
g.l. graus de liberdade (aproximados) da distribuição qui-quadrada					
g.l.e p-valor podem não ser válidos se modelo tiver variáveis exógenas					

Observe que a hipótese nula não é rejeitada para h = 8,9 e 10 a um p-valor superior a 4,8%.

A seguir é apresentado o teste de normalidade dos resíduos, onde a hipótese nula é que os resíduos tem distribuição normal multivariada.

Tabela B3 – Teste de Normalidade

Componente	Assimetria	Qui-quad	g.l.	p-valor
1	0.915344	1471.413	1	0.0000
2	0.110406	21.40695	1	0.0000
Conjunto		1492.820	2	0.0000
Componente	Curtosis	Qui-quad	g.l.	p-valor
1	3.361006	57.21837	1	0.0000
2	8.098425	11412.42	1	0.0000
Conjunto		11469.64	2	0.0000

Componente	Jarque-Bera	g.l.	p-valor
1	1528.631	2	0.0000
2	11433.83	2	0.0000
Conjunto	12962.46	4	0.0000

Os testes individuais de assimetria e curtosis para cada equação e para as duas ao mesmo tempo rejeita a hipótese nula. O mesmo acontece com o teste conjunto de Jarque-Bera que rejeita normalidade em cada uma das equações e no sistema.

A seguir é apresentado o teste de heteroscedasticidade sem os termos cruzados, isto é a regressão auxiliar só usa nível e quadrado.

Tabela B4 – Teste de Heteroscedasticidade

Teste Conjunto					
Qui-quad	g.l.	p-valor			
560.2478	111	0.0000			
Componentes Individuais					
Var. Dep	R^2	F(37,10499)	p-valor	Qui-quad(37)	p-valor
res1*res1	0.024606	7.158198	0.0000	259.2714	0.0000
res2*res2	0.024291	7.064286	0.0000	255.9525	0.0000
res2*res1	0.004514	1.286686	0.1141	47.56402	0.1144

O teste conjunto rejeita a nula de homoscedasticidade.

Devido aos resultados dos testes de normalidade e heteroscedasticidade os resultados das estimativas do VAR e o teste de causalidade tem p-valores que estão subestimados.