

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



Ana Beatriz Vieira de Mattos

**Impacto dos Investidores HFTs na Formação
de Preço no Mercado Cambial Brasileiro**

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da PUC-Rio como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Prof. Álvaro de Lima Veiga Filho

Rio de Janeiro

Julho de 2014



Ana Beatriz Vieira de Mattos

Impacto dos Investidores HFTs na Formação de Preço no Mercado Cambial Brasileiro

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Álvaro de Lima Veiga Filho
Orientador

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Antonio Carlos Figueiredo Pinto
Departamento de Administração – PUC-Rio

Prof. Jorge Passamani Zubelli
IMPA

Prof. José Eugenio Leal
Coordenador Setorial do Centro
Técnico Científico

Rio de Janeiro, 21 de julho de 2014

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, da autora e do orientador.

Ana Beatriz Vieira de Mattos

Graduou-se em Ciências Econômicas na Universidade de Brasília.

Ficha Catalográfica

Mattos, Ana Beatriz Vieira de

Impacto dos investidores HFTs na formação de preço no mercado cambial brasileiro / Ana Beatriz Vieira de Mattos; orientador: Álvaro de Lima Veiga Filho – 2014.

46 f.; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2014.

Inclui bibliografia

1. Engenharia elétrica – Teses. 2. High frequency trading. 3. Modelos de microestrutura. 4. Formação de preço. 5. Mercado cambial brasileiro. I. Veiga Filho, Álvaro de Lima. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

Agradecimentos

Aos meus pais, Wellington e Dilma, e ao meu irmão Jorge pelo apoio incondicional e por todo o carinho em mais esta etapa da minha vida.

Ao Wagner pela paciência e companheirismo em todos os momentos de trabalho e ansiedade. Obrigada pelo carinho e compreensão em todos os momentos!

Em especial, agradeço ao Guilherme Lamacie, Thalita Franklin e Yurie Yassunaga que além de grandes motivadores e companheiros, foram quem tornou possível este trabalho. Seus conhecimentos técnicos e dedicação foram imprescindíveis para o tratamento da base de dados, classificação dos investidores e longas discussões sobre o tema.

À todos os meus amigos, por tudo que representam na minha vida. Enumerá-los seria injusto.

Por fim, agradeço ao CNPq pelo suporte financeiro e a PUC pela oportunidade da excelente formação, ao meu orientador Professor Álvaro Veiga pela paciência, disponibilidade e incentivo para a realização deste trabalho e a todos os professores e funcionários do Departamento de Engenharia Elétrica da PUC.

À todos, meu sincero obrigado!

Resumo

Mattos, Ana Beatriz Vieira de; Veiga Filho, Álvaro de Lima (Orientador). **Impacto dos investidores HFTs na formação de preço no mercado cambial brasileiro.** Rio de Janeiro, 2014. 46p. Dissertação de Mestrado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

As mudanças tecnológicas e regulatórias foram facilitadores para o surgimento dos investidores de alta frequência, HFTs, no mix de participantes do mercado financeiro. Como classe, estes investidores não constituem uma entidade coerente e seu impacto e contribuição para a formação do preço não é clara. Esse trabalho analisou 10 categorias de investidores, que se diferenciam por suas características de latência, a partir de uma base de dados composta por todas as ordens enviadas para o *book* de dólar futuro com vencimento em 1º de agosto de 2013, da Bolsa de Valores e Mercadorias e Futuros (BM&FBovespa). Dentre toda as categorias de investidores testadas, a categoria HFT 1, a mais rápida de todas, foi a que apresentou o maior coeficiente de impacto no preço, 20%, e a maior medida de contribuição relativa para a volatilidade fundamental, 10%.

Palavras-chave

High frequency trading; Modelos de microestrutura; Formação de preço; Mercado cambial brasileiro.

Abstract

Mattos, Ana Beatriz Vieira de; Veiga Filho, Álvaro de Lima (Advisor). **HFT investor's impact on price formation in the Brazilian foreign exchange market.** Rio de Janeiro, 2014. 46p. MSc Dissertation - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Technological and regulatory changes were facilitators for the emergence of high frequency traders, HFTs, in the mix of financial market participants. As a class, these investors do not constitute a coherent entity and its impact and contribution to the price formation are not clear. This study analyzed 10 categories of investors, who are distinguished by their latency characteristics from a database comprised of all orders sent to the book of future dollar maturing on August 1, 2013, in the Brazilian Stock Exchange and Commodities and Futures Exchange (BM&FBovespa). Among all the categories of investors tested, the HFT 1, the fastest of all, was the one that had the highest coefficient of impact on price, 20%, and a larger measure of relative contribution to fundamental volatility, 10%.

Keywords

High frequency trading; Microstructure model; Price formation; Brazilian foreign exchange market.

Sumário

1 INTRODUÇÃO	10
2 MODELOS DE MICROESTRUTURA	13
2.1 Modelos de impacto sobre os preços	14
2.2 Modelo de decomposição do preço de negociação	15
2.3 Classificação dos diferentes investidores	17
3 METODOLOGIA EMPÍRICA	18
3.1 Modelo de impacto sobre os preços	18
3.1.1 Variáveis	18
3.1.2 Especificação do modelo	19
3.2 Modelo de decomposição do preço de negociação	20
4 BASE DE DADOS E ESTATÍSTICA DESCRITIVAS	24
4.1 Order book estilizado	24
4.2 Base de dados	26
4.3 Categorias de investidores	28
4.4 Métricas Descritivas dos Grupos de Investidores	32
5 RESULTADOS	35
5.1 Testes de robustez do modelo de impacto sobre os preços	35
5.2 Impacto no preço por categoria de investidores	38
5.3 Contribuição de cada categoria para a volatilidade fundamental	40
6 CONCLUSÃO	42
7 REFERÊNCIAS	44

Lista de Figuras

Figura 1: <i>Book</i> estilizado (estado 1)	24
Figura 2: <i>Book</i> estilizado (estado 2)	25
Figura 3: <i>Book</i> estilizado (estado 3)	25
Figura 4: <i>Book</i> estilizado (estado 4)	26
Figura 5: <i>Book</i> estilizado (estado 5)	26
Figura 6: Distribuição dos investidores a partir do volume negociado e posição acumulada máxima	29
Figura 7: Retornos e desequilíbrios do fluxo de ordens (OFI)	37
Figura 8: Comportamento intradiário dos coeficientes de impacto sobre o preço e profundidade do <i>book</i> de ordens.....	37
Figura 9: Coeficiente de impacto sobre preço e profundidade do mercado.	38

Lista de Tabelas

Tabela 1: Base de dados disponibilizadas pela BM&FBovespa	27
Tabela 2: Estatísticas diárias médias para cada grupo de investidor	34
Tabela 3: Resultado da regressão multivariada com desequilíbrios no fluxo de ordens dos dez níveis do <i>book</i> – média de todas as subamostras T	36
Tabela 4: Resultado da regressão univariada com desequilíbrios no fluxo de ordens dos dez níveis do <i>book</i> – média de todas as subamostras T	36
Tabela 5: Resultado da regressão com desequilíbrios do fluxo de ordens de cada categoria de investidor	40
Tabela 6: Resultado da regressão com desequilíbrios do fluxo de ordens de todas as categorias de investidor	40
Tabela 7: Contribuição de cada categoria de investidor para a volatilidade fundamental	41

INTRODUÇÃO

O novo ambiente de negociação é caracterizado pelo crescente ritmo da coleta de informação e a tomada de decisão baseada nesta informação. Velocidade é importante para os investidores por dois motivos. Primeiro, a volatilidade fundamental inerente aos instrumentos financeiros leva ao consentimento de que o rebalanceamento rápido das posições pode resultar em maior utilidade. Segundo, independente da velocidade absoluta, ser mais rápido na aquisição e compreensão da informação que outros investidores pode gerar oportunidade de lucro.

A última consideração parece conduzir a uma corrida armamentista, na qual os investidores utilizam tecnologia de ponta e colocam seus computadores o mais próximo possível do local de negociação a fim de reduzir a latência de suas ordens e ganhar alguma vantagem. Como resultado, os mercados de hoje experimentam intensa atividade em intervalos de milissegundos, algoritmos de computador respondem uns aos outros numa velocidade 100 vezes maior que a necessária para um operador humano piscar.

As bolsas estão investindo fortemente na modernização de seus sistemas para reduzir o tempo necessário para enviar informação para os investidores e para receber e manipular as ordens que estes enviam para o ambiente de negociação, além de disponibilizarem a possibilidade dos investidores colocarem seus computadores perto do sistema de negociação. Simultaneamente, os investidores estão investindo em tecnologia para processar a informação mais rápido, reduzindo o tempo que as mensagens levam para chegar ao sistema e deste ao investidor para pouco mais que um milissegundo.

Uma questão importante desta discussão é quem se beneficia deste novo ambiente de negociação. Este ambiente criou um novo tipo de investidor, o investidor de alta frequência (HFT, da sigla em inglês). Este investidor implementa estratégias de alta frequência, as quais respondem aos eventos de mercado em milissegundos, e são os responsáveis pela maioria das mensagens enviadas, além de participar de grande parte dos negócios executados. Embora pareça que sua principal atividade seja de intermediação, fornecendo liquidez aos

outros participantes do mercado, ainda não está claro se a atividade de alta frequência contribui com a qualidade do mercado.

No Brasil a atividade em alta frequência ainda não é majoritária e enfrenta bastante ceticismo por parte dos investidores e reguladores e as pesquisas acadêmicas sobre o impacto deste tipo de negociação no mercado brasileiro são ainda menores que mundo afora. O objetivo desta dissertação é analisar a influência dos investidores de alta frequência no mercado cambial brasileiro.

O mercado cambial no Brasil possui uma configuração absolutamente atípica, quanto à relação entre os mercados spot e futuro. Ventura [1] demonstra que a cotação da taxa de câmbio se forma primeiro no mercado futuro, sendo então transmitida por arbitragem para o mercado à vista. A questão que essa dissertação se propõe a responder é qual a influência dos investidores HFTs na formação do preço no mercado de dólar futuro negociado na Bolsa de Valores Mercadorias e Futuros (BM&FBovespa).

A disponibilidade de base de dados de ordens e negócios em alta frequência estimulou uma literatura de microestrutura, teórica e empírica, sobre a relação entre o fluxo de ordens e os movimentos dos preços nos mercados cuja negociação é feita através do envio de ordens. Apesar de vários modelos de impacto de preços e volatilidade terem sido propostos na literatura de microestrutura, há pouco consenso quanto a melhor forma de modelar o movimento dos preços neste tipo de mercado [2].

Na literatura empírica, o impacto de preços tem sido descrito por vários autores como linear, não-linear, quadrático, virtual, mecânico, temporário, imediato, permanente ou transitório. O único consenso parece ser a noção intuitiva de que o desequilíbrio entre a demanda e a oferta do fluxo de ordens movimentam os preços.

Utilizou-se o desequilíbrio do fluxo de ordens como única variável para estimar o impacto da atividade dos investidores de alta frequência na formação do preço do dólar com vencimento em 1º de agosto de 2013, como proposto por Cont et al [3].

Para avaliar a contribuição dos investidores de alta frequência na formação do preço fundamental estimou-se o modelo de decomposição da variância proposto por Hasbrouck [4], o qual decompõe a variância do preço de negociação

em variância fundamental e transitória através de um vetor autoregressivo (VAR) dos retornos do preço observado e variáveis relativas aos negócios executados.

A organização do trabalho é como segue. O próximo capítulo explica o estado da arte da literatura empírica de microestrutura e discute os principais pontos críticos da modelagem de dados para estes modelos. No Capítulo 3 a estratégia empírica é delineada, descrevendo a metodologia utilizada no trabalho. O Capítulo 4 explica o funcionamento do *book*, a partir de um *book* estilizado, explicita como a base de dados utilizada nesta dissertação foi produzida e como foi feita a classificação dos investidores em diferentes categorias. O Capítulo 5 apresenta os resultados para o impacto da atividade dos investidores HFT no preço de negociação e a contribuição destes investidores para a formação do preço fundamental. E, por fim, o Capítulo 6 apresenta as conclusões a partir dos resultados encontrados.

2

MODELOS DE MICROESTRUTURA

As pesquisas em microestrutura estão direcionadas a responder dois tipos de questões. A primeira se refere ao estudo da estrutura de mercado: como os custos de transação devem ser estimados, quais são as estratégias de negociação ótimas e como os mercados devem ser organizados. O segundo conjunto de questões se refere à influência que o mercado exerce na formação do preço (a incorporação de novas informações no preço).

A microestrutura empírica, por sua vez, é aplicada a três áreas do conhecimento. A primeira compreende os modelos econômicos formais de comportamento individual que oferecem previsões substantivas sobre como as variáveis observáveis deveriam se comportar. A segunda área é de análise estatística de séries temporais. A terceira área diz respeito à realidade institucional: os processos atuais através dos quais indivíduos e algoritmos computacionais se comportam no ambiente de negociação.

Os estudos empíricos atuais em microestrutura são caracterizados por uma grande diversidade de técnicas. Os dados de mercado são extremamente volumosos e apresentam características hostis à modelagem estatística: dinâmica complexa, não linearidade, não estacionariedade, periodicidade irregular e etc. A impossibilidade de modelar todas essas características conjuntamente, através de uma especificação que possa também resolver hipóteses econômicas alternativas, leva a uma variedade de modelos mais simples que tentam capturar um ou dois fenômenos relevantes para o problema em questão.

Para responder a questão proposta recorreu-se à segunda área da literatura de microestrutura empírica. Utilizou-se modelos lineares multivariados de análise de séries temporais.

2.1

Modelos de impacto sobre os preços

A literatura empírica sobre o impacto dos preços se concentrou principalmente nas séries de negócios executados. Uma abordagem é estudar o impacto das *parent orders* gradualmente executadas ao longo do tempo a partir de uma base de dados privada (Engle et. Al [5], Almgren et. Al [6]). Alternativamente, os estudos empíricos a partir de dados públicos ([7], [8], [9], [10], [11], [12], [13], [14]) investigaram a relação entre a direção e tamanhos de negócios e mudanças de preço e, normalmente, concluíram que o impacto sobre os preços dos negócios executados é uma função crescente e côncava do seu tamanho.

O foco em negociações deixa de fora as informações contidas nas apregoações, que proporcionam uma imagem mais detalhada da formação de preços [15]. Tomado como um todo, a literatura mostra consistentemente que a informação de apregoação tornou-se tão relevante quanto a informação de negociação. Os resultados de Cont et al. [3] demonstram que o desequilíbrio do fluxo de ordens é um *driver* mais forte da evolução dos preços de alta frequência do que as medidas padrão de negócios executados.

Embora estudos anteriores deem uma descrição relativamente complexa do impacto dos eventos do *book* sobre os preços dos ativos, Cont et al. [3] argumenta que este impacto pode ser modelado utilizando uma única variável, o desequilíbrio do fluxo de ordens (OFI), a qual representa o fluxo líquido de ordens de compra e de venda e acompanha as mudanças no tamanho das filas de compra e venda:

- aumentando toda vez que a quantidade de ordens de compra aumenta, a quantidade de ordens de venda diminui ou o preço de compra/preço de venda aumenta; e
- diminuindo toda vez que a quantidade de ordens de compra diminui, a quantidade de ordens de venda aumenta ou o preço de compra/ preço de venda diminui.

A variável OFI trata uma venda a mercado¹ e um cancelamento de compra do mesmo tamanho da mesma forma, uma vez que eles têm o mesmo efeito sobre o tamanho da fila de oferta. O modelo de impacto sobre os preços resultante desta variável relaciona preços, negócios executados, ordens e cancelamentos de uma maneira simples: linear, requer a estimação de um único parâmetro e mostrou-se robusto para diferentes instrumentos e intervalos de tempo.

O coeficiente de impacto sobre os preços resultante do modelo apresenta sazonalidade intradiária em consonância com os padrões observados nos *spreads*, profundidade de mercado e volatilidade dos preços ([16], [17], [18], [19]), que foram explicadas em termos de mudanças intradiárias na assimetria de informação [20] ou quantidade de informação nos negócios executados [21].

Neste modelo, o coeficiente de impacto sobre os preços e as variações na profundidade de mercado são inversamente proporcionais. Isto permite explicar os padrões intradiários do impacto sobre os preços e a volatilidade dos preços utilizando apenas o desequilíbrio do fluxo de ordens e a profundidade do mercado, ao invés de parâmetros não observáveis anteriormente invocados na literatura, como assimetria de informação ou quantidade de informação nos negócios executados.

2.2

Modelo de decomposição do preço de negociação

A contribuição dos investidores HFTs na formação do preço fundamental é apreendida na literatura através de modelos de decomposição do preço de negociação. A partir da perspectiva econômica, o preço do instrumento em muitos modelos de microestrutura pode ser interpretado como um preço “informacional fundamental” idealizado, corrompido por perturbações atribuídas às fricções do processo de negociação, de modo que a variação do preço do instrumento se divide entre componentes fundamentais, (permanente ou informacional) e transitórios (fricções atribuídas ao mercado).

¹ Operação de venda resultante de uma ordem de venda a mercado. Este tipo de ordem não contém informação de preço, informa apenas a quantidade a ser vendida ao melhor preço no momento de seu envio.

No contexto dessa dissertação, as fricções de mercados estão relacionadas a atividade de negociação dos investidores HFTs. Esta distinção é mais sutil do que a primeira, porque, enquanto a diferença entre o componente fundamental e o transitório surge frequentemente em análises econômicas, a preocupação com o papel da atividade *per se* na formação do preço é largamente peculiar aos estudos de microestrutura.

O modelo de Hasbrouck [4] supõe que os preços de transação são determinados pela soma de dois componentes não observáveis: um referente ao preço fundamental, que segue um passeio aleatório, e outro componente transitório e estacionário referente a diferença entre o preço fundamental e o preço de transação, denominado erro de preço.

A decomposição é feita a partir de uma representação de média móvel (MA) para as mudanças de preços, e desta representação são feitas as inferências desejadas. Esta é uma abordagem importante porque os modelos estruturais apresentados anteriormente na literatura empírica de microestrutura não são abrangentes e realistas, dado a complexidade dos processos de negociação. Os modelos estatísticos de séries temporais impõe menos estrutura aos dados, e podem, portanto, ser mais robustos.

Para a decomposição do preço, o modelo considera além deste outras variáveis relativas aos negócios executados. Porque muitas considerações econômicas sugerem que ordens grandes devem ter um alto conteúdo informacional, faz sentido incluir como variável do modelo de decomposição o volume direcional do negócio executado e/ou transformações monotônicas destes volumes. No entanto, estas variáveis geralmente apresentam pouco poder explicativo incremental em conjunto com a variável direção do negócio, compra ou venda, ([21]). Uma explicação para o baixo poder explicativo das variáveis de volume direcional é a mudança de liquidez presente no mercado em diferentes períodos de tempo. Se os investidores negociam grandes volumes quando há liquidez no *book* ou negociam pequenos volumes quando há pouca liquidez, o impacto sobre o preço será pouco afetado pelo tamanho da ordem.

2.3

Classificação dos diferentes investidores

Finalmente, precisa-se separar os diferentes tipos de investidores. A atividade de negociação em alta frequência abrange uma grande variedade de estratégias e agressividade para diferentes ativos ou instrumentos.

Para distinguir os investidores de alta frequência de outros investidores e entre eles, os investidores devem ser separados em grupos homogêneos que exibem características semelhantes. Este é um desafio, devido à qualidade dos dados, bem como uma visão limitada das estratégias utilizadas, dada a sua natureza proprietária. A literatura disponível sugere que os HFTs podem ser classificados em geral, embora não necessariamente ou exclusivamente, por tipos de estratégia ([22], [23], [24] e [25]), agressividade na negociação ([26], [27], [28] e [29]) ou latência ([30]).

Escolheu-se classificar os investidores a partir de suas características de latência, uma vez que esta classificação permite uma divisão mais detalhada dos investidores e porque dessa forma as variáveis poderiam ser ordenadas, com relação a latência, no modelo de decomposição da variância conforme requerido pela decomposição de Cholesky .

Os investidores indentificados na base de dados foram classificados de acordo com uma série de *proxies* da sua atividade diária. As *proxies* e as regras de classificação dos investidores foram escolhidas a partir do reconhecimento dos padrões de atividade de cada grupo de investidor.

Esta abordagem não classifica os investidores com exatidão. Em vez disso, ela infere a classificação a partir dos dados com consequências relativamente desconhecidas para a qualidade da inferência do modelo. Mas a vantagem é que ela permite a construção de uma base de dados fina o suficiente para permitir inferências bastante gerais.

Os dados foram disponibilizados pela Diretoria de Controles Internos, *Compliance* e Risco Corporativo - Superintendência de Riscos Financeiros e Modelagem. A classificação dos investidores foi feita pela Superintendência de Riscos Financeiros e Modelagem antes da base de dados ser disponibilizada, por questões de segurança da informação.

3

METODOLOGIA EMPÍRICA

3.1

Modelo de impacto sobre os preços

Para avaliar a contribuição relativa nas diferentes categorias de investidores para o processo de formação de preço, calibrou-se o modelo de impacto sobre o preço de Cont et al [3]. O modelo relaciona as mudanças de preços com os desequilíbrios no fluxo de ordens e profundidade do *book*, proporcionando uma ligação entre a volatilidade do fluxo de ordens e a liquidez do mercado.

3.1.1

Variáveis

Calculou-se as variáveis a partir da emulação do *book*, de modo que para cada nova ordem calculou-se o melhor preço de compra, p^c , a profundidade do primeiro nível do *book* de compra, d^c , o melhor preço de venda, p^v e a profundidade do primeiro nível do *book* de venda, d^v , depois que cada ordem é processada². Enumerou-se as observações por n e calculou-se as seguintes diferenças entre observações consecutivas:

$$e_n = d_n^c 1_{\{p_n^c \geq p_{n-1}^c\}} - d_{n-1}^c 1_{\{p_n^c \leq p_{n-1}^c\}} - d_n^v 1_{\{p_n^v \leq p_{n-1}^v\}} + d_{n-1}^v 1_{\{p_n^v \geq p_{n-1}^v\}} \cdot \quad (1)$$

A variável e_n é a contribuição dos eventos do *book* para o desequilíbrio entre oferta e demanda. Utiliza-se duas frequências temporais: $\{T_0 \dots T_1\}$ e $\{t_{0,0} \dots t_{1,1}\}$ para controlar a não estacionariedade intradiária na variação dos preços e do fluxo de ordens.

² O melhor preço de compra se refere ao maior preço presente no *book* de compra, sendo este o seu primeiro nível. O melhor preço de venda se refere ao menor preço presente no *book* de venda, sendo este o seu primeiro nível.

As frequências são tais que $T_i - T_{i-1} \cong 16$ minutos e $t_{j,i} - t_{j-1,i} = \Delta t = 5$ segundos. Para cada intervalo longo $[T_{i-1}, T_i]$ calcula-se a profundidade média do primeiro nível do *book*, AD_i , e 200 variações de preços, $\Delta p_{j,i}$, e desequilíbrio no fluxo de ordens, $OFl_{j,i}$, indexados por j .

$$\Delta p_{j,i} = \frac{p_{j,i}^c + p_{j,i}^v}{2\delta} - \frac{p_{j-1,i}^c + p_{j-1,i}^v}{2\delta}, \quad (2)$$

$$OFl_{j,i} = \sum_{n=N(t_{j-1}+1)}^{N(t_{j-1})} e_n, \quad (3)$$

$$AD_i = \frac{\sum_{n=N(t_{j-1}+1)}^{N(t_{j-1})} d_n^c + d_n^v}{2}. \quad (4)$$

Onde $N(t_{j-1} + 1)$ e $N(t_{j-1})$ são índices da primeira e última observação pertencente ao intervalo $[t_{j-1,i}, t_{j,i}]$. O tick size δ é de cinquenta centavos para o mercado de dólar futuro.

O intervalo de tempo da frequência mais rápida é da mesma ordem de magnitude (ou maior) que o tempo médio entre mudanças de preços consecutivas.

Antes de calcular as variáveis eliminaram-se as observações com os maiores *spreads*³, pois variações no *spread* conduzem a *outliers* na variável dependente $\Delta p_{t,T}$. Então, estimaram-se os coeficientes do modelo através de regressões MQO.

3.1.2

Especificação do modelo

O modelo consiste de duas equações – a Equação (5) relaciona as mudanças de preços e o fluxo de ordens em uma escala de tempo rápida, e a Equação (6) relaciona o coeficiente de impacto de preço da Equação (5) e a profundidade do mercado em uma frequência temporal menor. A Equação (7) apresenta uma generalização do modelo original, ou seja, uma versão detalhada da Equação (5) que distingue ordens de K categorias de investidores.

³ Diferença entre o melhor preço de compra e o melhor preço de venda (*spread*) maior que 20 pontos bases (da média entre estes melhores preços), mas no máximo 5% de todas as observações com os maiores valores de *spread*.

$$\Delta p_{j,i} = \alpha_i + \beta_i OFI_{j,i} + \varepsilon_{j,i} , \quad (5)$$

$$\log(\beta_i) = \gamma + \eta \log(AD_i) + \zeta_i , \quad (6)$$

$$\Delta p_{j,i} = \sum_{k=1}^K \beta_{k,i} OFI_{k,j,i} + \xi_{j,i} . \quad (7)$$

Onde, β_i é o coeficiente de impacto sobre o preço para a subamostra i , e γ, η são parâmetros.

A Equação (7) é uma versão da Equação (5), na qual a variável $OFI_{j,i}$ é desagregada em K variáveis de desequilíbrio do fluxo de ordens, uma para cada categoria de investidor. Finalmente, $\varepsilon_{j,i}$, $\xi_{j,i}$ e ζ_i representam os erros do modelo, aleatórios e não correlacionados com as outras variáveis.

Primeiramente, calibrou-se as Equações (5) e (6) para estabelecer a relação entre mudanças de preços, fluxo de ordens e profundidade do *book*. Em seguida a variável $OFI_{j,i}$ foi desagregada em parcelas referentes a cada categoria de investidor k e estimou-se a Equação (7) a fim de comparar os coeficientes de impacto sobre o preço entre as diferentes categorias.

3.2

Modelo de decomposição do preço de negociação

A formação de preço é essencialmente a volatilidade do preço fundamental, sendo este interpretado como o valor que o preço de transação assume quando todos os fundamentos necessários para sua valoração são conhecidos por todo mercado e todos os investidores negociam baseados nestes fundamentos. Isto, no entanto, é improvável, já que a atividade de negociação pode gerar um excesso de volatilidade (ou ruído) que não está relacionada com a informação sobre os fundamentos. Assim o preço transacionado é o preço fundamental perturbado por algum ruído.

Supondo que este ruído está relacionado com a atividade de negociação dos investidores, analisou-se a contribuição de investidores de diferentes categorias na volatilidade fundamental através do método de decomposição da variância de Hasbrouck [4].

O método de Hasbrouck [4] supõe que os preços de transação são determinados pela soma de dois componentes não observáveis: p_n^* , o preço

fundamental, que segue um passeio aleatório e s_n , um componente transitório e estacionário referente a diferença entre o preço fundamental e o preço de transação, denominado erro de preço, tal que:

$$p_n = p_n^* + s_n, \quad (8)$$

$$p_n = \frac{p_n^c + p_n^v}{2}, \quad (9)$$

$$p_n^* = p_{n-1}^* + v_n, \quad v_n \sim iid(0, \sigma_v^2). \quad (10)$$

As inovações v_n refletem atualizações da informação pública, incluindo, quando relevante, o último negócio realizado. O termo s_n é um processo estocástico não determinístico de média zero, com a covariância entre s_n e v_n constante no tempo. Intuitivamente, s_n incorpora todas as imperfeições transitórias de microestrutura que fazem o preço observado se desviar do fundamental. s_n pode ser serialmente correlacionado e parcialmente correlacionado com v_n .

A suposição de estacionariedade na covariância implica que $\lim_{h \rightarrow \infty} E[s_{n+h}] = 0$, o que reflete a transitoriedade das imperfeições de microestrutura. Outra implicação desta suposição é homocedasticidade.

A variância fundamental pode ser estimada através de um vetor autoregressivo (VAR) dos retornos de p_n , r_n , e a variável $q_{k,n}$, que assume os valores -1 se a categoria de investidor k enviou uma ordem de venda que gerou negócio, 0 se a categoria de investidor k enviou uma ordem que não gerou negócio e 1 se a categoria de investidor k enviou uma ordem de compra que gerou negócio em n .

As observações n foram filtradas de modo que considerou-se para estimação do modelo apenas as observações em que houve negócios ou o melhor preço, de compra ou de venda, se altera. Atualizando o tempo dessa maneira, sempre que ocorre um negócio há uma atualização do retorno, mas não o contrário. Isto implica que o sentido da causalidade contemporânea é apenas do negócio realizado para a revisão de ordens enviadas, o que é verdade no mercado eletrônico, dado que uma ordem a mercado suficientemente grande pode consumir toda a profundidade do melhor nível disponível e alterá-lo. Adicionalmente, o modelo não permite que os negócios de várias categorias influenciem o retorno

contemporaneamente, dado que analisam-se as transações resultantes de ordens submetidas para um único *book*.

A especificação do modelo VAR, na forma matricial, é como segue:

$$\mathbf{y}_n = \sum_{z=1}^p A_z \mathbf{y}_{n-z} + \mathbf{v}_n, \quad (11)$$

$$\mathbf{y}_n = [r_n \quad q_{1,n} \quad \cdots \quad q_{k,n}]^T, \quad (12)$$

$$\mathbf{v}_n = [v_{r,n} \quad v_{1,n} \quad \cdots \quad v_{k,n}]^T. \quad (13)$$

Onde, A_z são matrizes de coeficientes autoregressivos, uma para cada defasagem z e \mathbf{v}_n é vetor de inovações de média zero serialmente não correlacionados com $\text{Var}(v_{r,n}) = \sigma_r^2$, $\text{Var}(v_{k,n}) = \sigma_k^2$ e $E[v_{r,n}, v_{k,n}] = 0$.

Assumindo invertibilidade, os retornos e a variável de negócio podem ser expressos como uma função linear das inovações correntes e passadas. A representação VMA correspondente ao VAR representado na Equação (11) é como descrito na Equação (14):

$$\mathbf{y}_n = \sum_{z=1}^q B_z \mathbf{v}_{n-z}. \quad (14)$$

Para o cálculo da variância fundamental apenas a equação referente aos retornos dos modelos VAR e VMA, Equação (15) e (16), respectivamente, são necessárias:

$$r_n = \sum_{z=1}^p a_z^* r_{n-z} + \sum_{k=1}^K \sum_{z=0}^{p-1} a_{k,z} q_{n-z} + v_{r,n}, \quad (15)$$

$$r_n = \sum_{z=0}^{q-1} b_z^* v_{r,n-z} + \sum_{k=1}^K \sum_{z=0}^{q-1} b_{k,z} v_{k,n-z}. \quad (16)$$

O modelo considera o impacto contemporâneo (imediato) das transações no nível de preços. A ordenação causal das transações para os preços é feita com base na observação do comportamento das operações no *book*. Sempre que há uma transação no mercado, o topo do *book* é atualizado. A inclusão de variáveis de negócios contemporâneas na equação de retornos garante que a estrutura recursiva da matriz de relações contemporâneas das variáveis, requerida pela decomposição de Cholesky, seja atendida e que a matriz de covariância do modelo VAR se torne diagonal por blocos, permitindo a decomposição das volatilidades em componentes relacionados aos negócios de cada categoria de investidor.

Dessa forma, a variância fundamental é calculada da seguinte maneira:

$$\sigma_w^2 = (\sum_{z=0}^{\infty} b_z^*)^2 \sigma_r^2 + \sum_{k=1}^K (\sum_{z=0}^{\infty} b_{k,z})^2 \sigma_k^2. \quad (17)$$

Onde b_z^* e $b_{k,z}$ são coeficientes do modelo VMA. O termo $(\sum_{z=0}^{\infty} b_z^*)^2 \sigma_r^2$, na Equação (17), se refere a contribuição da informação pública na volatilidade fundamental. A contribuição relativa da infomação pública é dada por:

$$\frac{(\sum_{z=0}^{\infty} b_z^*)^2 \sigma_r^2}{\sigma_w^2}. \quad (18)$$

A contribuição relativa de cada categoria k de investidor na volatilidade fundamental é dada por:

$$\frac{(\sum_{z=0}^{\infty} b_{k,z})^2 \sigma_k^2}{\sigma_w^2}. \quad (19)$$

Uma interpretação intuitiva para o cálculo da variância fundamental pode ser dado através da Equação (16). Assuma que em $n = 0$ o sistema esteja inicialmente estável com todas as inovações defasadas iguais a zero. Suponha um choque de negócio dado por $u_{k,0}$. A primeira atualização de ordens será $r_0 = b_{k,0}u_{k,0}$, a segunda será $r_1 = b_{k,1}u_{k,0}$ e assim por diante. Acumulando todas as revisões, o efeito do negócio inicial no preço pode ser calculado como $(\sum_{j=0}^{\infty} b_{k,n})u_{k,0}$.

A contrinuição para a volatilidade fundamental de cada categoria é uma medida da informação contida nos negócios feitos por uma categoria relativa às outras categorias, mas não permite avaliar a contribuição total e absoluta de cada categoria para a volatilidade fundamental.

4

BASE DE DADOS E ESTATÍSTICA DESCRITIVAS

4.1

Order book estilizado

As medidas de atividade no mercado e as variáveis para estimação dos modelos apresentados nesse trabalho é calculado a partir da emulação do *book*. Para exemplificar a emulação segue um exemplo de um *book* estilizado, onde o efeito instantâneo de eventos no *book* podem ser explicitamente observados.

Inicialmente o *book* apresenta o melhor preço de venda igual a R\$ 20,10, sendo este o seu primeiro nível, com profundidade igual a 213. Do lado oposto, o primeiro nível de compra é a R\$ 20,00 com profundidade igual a 205.

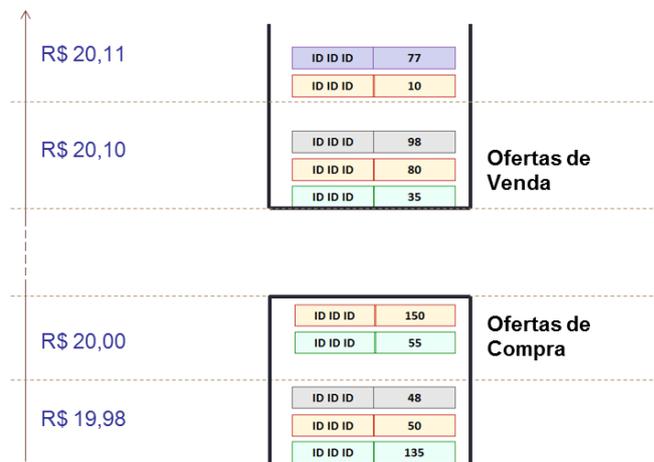


Figura 1: *Book* estilizado (estado 1)

Para cada nova ordem que chega ao *book* há três possibilidades:

- 1) entra no final da fila no nível de preço igual ao seu preço;
- 2) cria um novo nível, caso não exista um nível com o seu preço; ou
- 3) realiza um negócio com uma ordem do lado oposto do *book*, quando o seu preço é igual ou melhor que o primeiro nível do *book* do lado oposto. Os negócios são feitos respeitando a fila de ordens do *book* e caso o tamanho da

ordem seja maior que a profundidade disponível, a tamanho restante da ordem após a execução dos negócios é inserido no *book* do seu lado.

Por exemplo, se chegasse uma ordem de venda a R\$ 20,11 e tamanho 50, ela entraria no final da fila do segundo nível de venda e o estado final do *book* seria como na Figura 2.

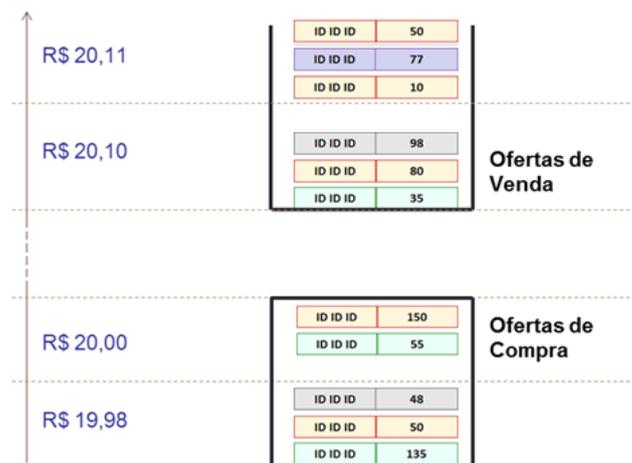


Figura 2: *Book* estilizado (estado 2)

A chegada de uma ordem de venda a R\$ 20,05 e tamanho 5 criaria um novo primeiro nível de venda com este preço e profundidade e o estado final do *book* seria como na Figura 3.

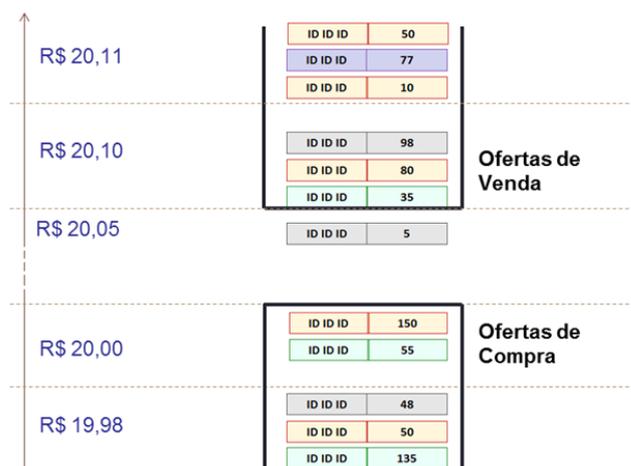


Figura 3: *Book* estilizado (estado 3)

Por fim, a chegada de uma ordem de compra a R\$ 20,10 de tamanho 228 exerceria quatro negócios e criaria um novo primeiro nível de compra a R\$ 20,10

de tamanho 10. A Figura 4 mostra a entrada da ordem de compra e a Figura 5 o estado final do *book*.

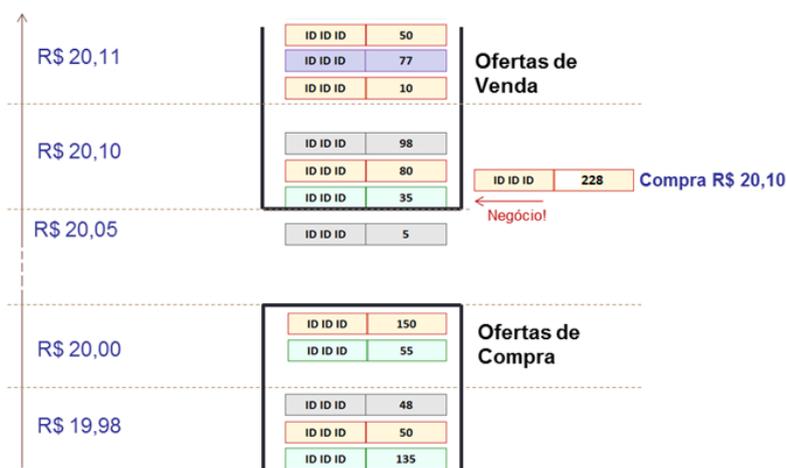


Figura 4: *Book* estilizado (estado 4)

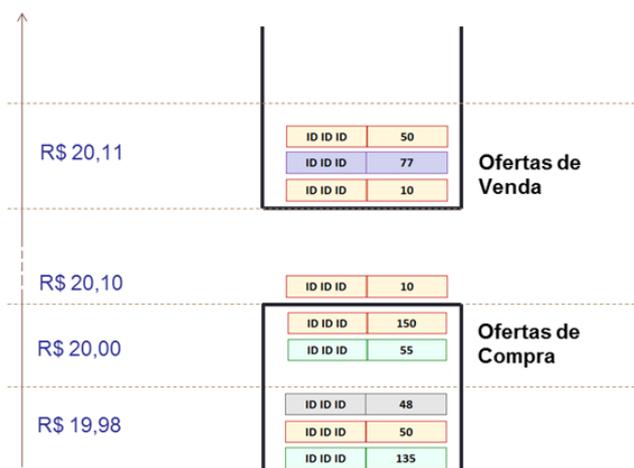


Figura 5: *Book* estilizado (estado 5)

4.2

Base de dados

A base de dados objeto desse trabalho compreende as ordens enviadas para o *book* de DOLQ13, futuro de dólar com vencimento em 1º de agosto de 2013 (ordens novas, cancelamentos e modificações) da Bolsa de Valores Mercadorias e Futuros (BM&FBovespa). Foram consideradas as ordens enviadas no mês anterior ao vencimento, uma vez que o primeiro vencimento é o mais

líquido deste contrato. O dia 31 de julho foi excluído da base por ser um dia de negociação atípico - último dia de negociação do vencimento.

A base traz para cada ordem enviada os seguintes campos: categoria do investidor, horário na precisão de milissegundos, tipo da ordem (limite ou mercado), lado do *book* para o qual a ordem foi endereçada (compra ou venda), ação da ordem (nova, modificação ou cancelamento), preço da ordem, quantidade da ordem, além do melhor preço de compra e de venda no momento em que a ordem foi enviada. São consideradas apenas as ordens enviadas no período de negociação, excluindo-se as ordens enviadas no período de leilão. O formato da base de dados pode ser visualizado na Tabela 1.

A BM&FBovespa permite negociação através de outros tipo de ordens que limite e mercado. No entanto, a grande maioria das ordens são deste tipo, de modo que decidiu-se considerar apenas estas ordens para a emulação do *book*, dado o custo computacional para o tratamento de outros tipos de ordem e pouca representatividade. A ordem limite determina o preço e a quantidade a ser executado no *book*, enquanto a ordem a mercado determina apenas a quantidade a ser executada ao melhor preço do *book* no momento de seu envio.

Tabela 1: Base de dados disponibilizadas pela BM&FBovespa

ateg.	hh	mm	ss	ms	Tipo	Lado	Ação	Preço	Qty	Best.Bid	Best.Ask
1	10	0	1	1	1	0	0	7.50	7700	11.80	12,20
1	10	0	1	2	1	0	0	15.00	200	11.80	12,20
1	10	0	1	16	1	1	0	12.20	5000	11,85	12,20
4	10	0	16	20	1	0	0	12.30	5000	11,85	12.15
5	10	0	35	40	2	0	0	11.51	2000	11,85	12,20
2	10	0	35	65	1	0	0	16.10	1000	11.80	12.15
2	10	0	36	12	2	1	0	12.15	1000	11.80	12.15
8	10	0	49	42	1	0	0	11.85	2000	11,85	12.15
1	10	1	0	15	1	0	0	11.80	2000	11.80	12,20
1	10	1	0	16	1	0	0	11.85	2400	11,85	12,20

Para o propósito desse estudo, o conjunto de dados está baseado na sequência de pontos discretos que representam a entrada das ordens no *book*, indexados por n . Usando este conjunto de campos sobre as ordens enviadas, reconstruiu-se, dinamicamente, o *book* de cada dia. Na emulação do *book*, para cada atualização é guardado o preço e a quantidade da ordem que gera a

atualização, além da quantidade total de ordens e os preços dos dez primeiros níveis do *book* (tanto de compra como de venda) depois da atualização. Para minimizar possíveis erros do emulador (gerados por causa de leilões intradiários, por exemplo, que não podem ser observados nos dados), é sincronizado, dinamicamente, o melhor preço de compra e de venda do emulador com o melhor preço de compra e de venda da base de dados.

O resultado do emulador é uma tabela com todas as ordens enviadas em um dia e o estado do *book* depois que cada ordem é processada. A partir desta tabela calculou-se várias medidas de atividade no mercado e as variáveis para estimação dos modelos.

4.3

Categorias de investidores

Uma maneira simplista de classificar os investidores de um mercado é separá-los entre investidores de alta frequência, posicionais, arbitradores e formadores de mercado. A atividade dos dois primeiros é mais facilmente apreendida através de métricas calculadas a partir dos dados de sua atividade diária. Os investidores de alta frequência, em geral, negociam grandes volumes, mas não apresentam uma posição intradiária grande. Os investidores posicionais, ao contrário, apresentam posições intradiárias próximas ao volume negociado.

Os investidores presentes na base de dados foram classificados a partir de uma série de *proxies* da sua atividade diária no mercado durante diferentes dias de negociação. Duas dessas *proxies* são suficientes para classificar os investidores de alta frequência e os investidores posicionais: posição máxima acumulada durante o dia e volume negociado diário.

A disposição dos diferentes investidores em um gráfico destas duas variáveis mostra um agrupamento que pode ser usado como regra de classificação. Observam-se claramente os investidores HFTs no centro superior do gráfico e os investidores posicionais nas laterais. No centro inferior do gráfico estão os formadores de mercado e um grupo de investidores que apresentam comportamentos diversos. A Figura 6 é uma representação do gráfico gerado a partir dos dados.

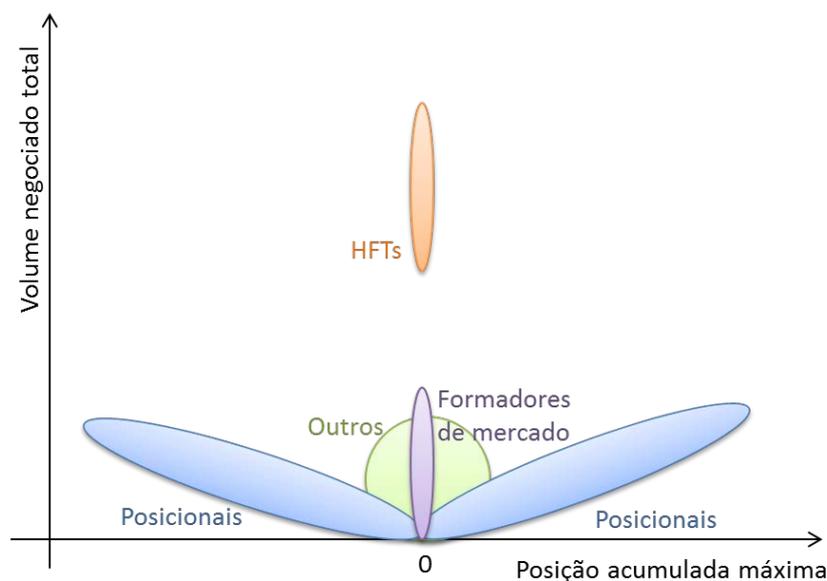


Figura 6: Distribuição dos investidores a partir do volume negociado e posição acumulada máxima

Esta classificação seria suficiente para os objetivos desse trabalho, mas utilizou-se uma segunda regra de classificação. Esta regra considera além do volume negociado a velocidade da atividade dos investidores. O objetivo é diferenciar os investidores HFTs dos investidores algoritmos, sendo que estes podem estar atuando como fundamentalista ou qualquer outro tipo de investidor, mas em velocidade acima da considerada possível para humanos. Definiu-se dez categorias de investidores, que se diferenciam por suas características de latência, e cada investidor foi classificado em uma destas categorias. É importante notar que a categoria HFT 1 da segunda regra de classificação é, por construção, coincidente com a primeira regra de classificação para HFT apresentada e será o foco desse trabalho.

O primeiro passo da segunda classificação é separar os investidores inativos, que negociam um volume diário muito pequeno (como grupo, eles negociam os menores volumes do dia analisado, que representam 10% do volume total). Os investidores remanescentes são separados em nove grupos. Quatro destes grupos são subclassificações de HFTs, outras quatro são subclassificações de algoritmos e a última classe são os investidores que apresentam características de negociação compatíveis com a capacidade humana.

É importante ressaltar que dependendo da atividade do investidor no mercado, ele pode trocar de categorias nos diferentes dias, mas a maioria dos investidores pertence a uma única categoria por todo o período analisado.

HFTs: Investidores de alta frequência são caracterizados por um alto número de mensagens, intervalos curtos entre mensagens e participação constante nas seções de negociação. Estas características podem ser capturadas a partir da relação entre sua posição acumulada máxima e o volume total negociado no dia. Especificamente, eles têm as seguintes características:

- **HFT 1:**
 - Ser investidor ativo;
 - Apresentar posição acumulada máxima dividida pelo volume total negociado menor que 0,02.
- **HFT 2:**
 - Ser investidor ativo;
 - Não ter sido classificado como HFT 1;
 - Apresentar posição acumulada máxima dividida pelo volume total negociado entre 0,02 e 0,15;
 - A distribuição do tempo entre mensagens ter mediana < 5 ms e percentil 25 < 1 ms.
- **HFT 3:**
 - Ser investidor ativo;
 - Não ter sido classificado como HFT 1 ou 2;
 - Apresentar posição acumulada máxima dividida pelo volume total negociado menor que 0,25;
 - Apresentar posição final menor que 5% do volume total negociado.
 - Enviar mais de 100 mensagens a cada 30 minutos, em média.
 - A distribuição do tempo entre mensagens ter mediana < 5 ms e percentil 25 < 1 ms.
- **HFT 4:**
 - Ser investidor ativo;
 - Não ter sido classificado como HFT 1, 2 ou 3;
 - Pertencer a lista de investidores marcados como HFT independente do seu comportamento no dia. Estes investidores não se comportam como HFT todos os dias, mas apresentam capacidade de impactar o mercado como tal.

Algoritmos: Estes são investidores cuja atividade de envio de mensagens apresenta velocidade maior que a capacidade humana, podendo ou não apresentar características cíclicas. Para identificar a existência de ciclos calcula-se o número de mensagens enviadas em cada intervalo de 50 milissegundos de cada segundo. As características para ser classificado como algoritmo, daqui para frente denominado algo, são:

- **Algo 1:**
 - Ser investidor ativo;
 - Não ter sido classificado como HFT;
 - Enviar mais de 100 mensagens a cada 30 minutos, em média.
 - A distribuição do tempo entre mensagens ter mediana < 5 ms e percentil 25 < 1 ms.
- **Algo 2:**
 - Ser investidor ativo;
 - Não ter sido classificado como HFT ou Algo 1;
 - Enviar mais de 60 mensagens a cada 30 minutos, em média.
 - A distribuição do tempo entre mensagens ter mediana < 15 ms e percentil 25 < 5 ms.
- **Algo 3:**
 - Ser investidor ativo;
 - Não ter sido classificado como HFT ou Algo 1 ou 2;
 - A distribuição do número de mensagens por intervalos de 50 milissegundos dentro do segundo apresentar um intervalo com quantidade de mensagens significativamente maior.
- **Algo 4:**
 - Ser investidor ativo;
 - Não ter sido classificado como HFT ou Algos 1, 2 ou 3;
 - Pertencer a lista de investidores marcados como Algoritmos independente do seu comportamento no dia. Estes investidores não se comportam como Algoritmos todos os dias, mas apresentam capacidade de impactar o mercado como tal.

Algo lento ou humano: todos os investidores que não apresentaram nenhuma das características necessárias para serem classificados como HFT ou algoritmos.

4.4

Métricas Descritivas dos Grupos de Investidores

Calculou-se, para cada investidor, uma série de métricas descritivas para analisar a relevância dos investidores HFTs no mercado cambial brasileiro. Depois, calculou-se a média destas métricas dentro de cada categoria para cada dia e finalmente a média de todos os dias para cada categoria. A definição de cada métrica é como segue:

- Participação no volume negociado – volume negociado, como percentagem do volume total negociado no mercado.
- Participação no número de mensagens – número de mensagens enviadas, como percentual do total de mensagens enviadas para o *book*.
- Volume de negócios passivos – volume de negócios gerados com ordens passivas como percentual do volume negociado total para cada investidor.
- Posição máxima percentual – Posição máxima acumulada por um investidor durante o dia dividido pelo seu volume negociado total no mesmo dia.
- Tempo de *double-sided quote* – percentual do tempo que o investidor manteve ordens nos dois lados do *book*.
- Tempo de *single-sided quote* – percentual do tempo que o investidor manteve ordens em pelo menos um lado do *book*.
- Duração entre mensagens (mediana) – mediana da distribuição do tempo entre mensagens.
- Duração entre mensagens (perc. 25) – percentil 25 da distribuição do tempo entre mensagens.
- $|\text{Posição no final do dia}| / \text{Volume}$ – valor absoluto da posição no final do dia como percentagem do volume total negociado.

- Proporção de negócios – volume de ordens executadas, incluindo execuções parciais, como percentual do volume total de ordens novas e de modificação.
- Proporção de cancelamentos – volume de ordens de cancelamento, incluindo cancelamentos parciais, como percentual do volume total de ordens novas e de modificação.
- Proporção de modificações – volume de ordens de modificação, como percentual do volume total de ordens novas e de modificação.

As métricas descritivas para cada categoria de investidor estão apresentadas na Tabela 2. Através delas observa-se o domínio dos investidores HFT no mercado de dólar futuro. Os investidores HFTs, conjuntamente, negociam 20% do volume negociado total e enviam metade do total de mensagens enviadas. Os HFTs 1, apenas, são responsáveis por 37% do total de mensagens enviadas e operam 9% do volume negociado total.

Tabela 2: Estatísticas diárias médias para cada grupo de investidor.

	Inativos	HFT1	HFT2	HFT3	HFT4	Algo1	Algo2	Algo3	Algo4	Outros
Volume (% do volume total)	10,02%	9,18%	8,81%	1,61%	0,47%	2,34%	1,35%	18,57%	10,57%	40,33%
N° de mensagens (% do N° de mensagens total)	37,09%	37,33%	7,36%	6,02%	0,53%	9,20%	1,49%	5,76%	1,03%	3,11%
Volume Passivo (% do volume total)	49,60%	59,84%	40,18%	69,89%	4,08%	81,80%	65,96%	56,81%	46,69%	50,78%
Posição máxima acumulada (% do volume total)	62,81%	1,04%	6,11%	21,00%	9,65%	59,96%	61,93%	30,21%	51,06%	49,18%
Tempo de <i>double-sided quote</i> (% do tempo c/ ordens no book)	17,62%	43,76%	30,16%	98,66%	24,72%	78,48%	63,56%	33,00%	12,04%	13,56%
Tempo de <i>single-sided quote</i> (% do tempo c/ ordens no book)	22,70%	70,51%	49,35%	99,53%	23,21%	80,81%	43,96%	63,07%	35,57%	39,96%
Dururação entre mensagens (mediana)	565,37	3,52	1,04	0,01	51,69	0,21	2,33	22,18	210,10	158,31
Dururação entre mensagens (percentil 25)	38,01	0,87	0,07	0,00	4,30	0,11	0,61	4,78	24,18	19,84
Posição no final do dia (% do volume negociado)	49,37%	0,15%	0,98%	6,44%	1,65%	51,60%	54,61%	22,54%	41,21%	34,98%
Proporção de negócios	67,32%	37,31%	48,07%	4,34%	78,68%	18,21%	42,32%	52,31%	62,46%	68,77%
Proporção de cancelamentos	38,05%	58,27%	52,60%	46,46%	20,33%	39,68%	21,70%	25,70%	19,45%	17,00%
Proporção de modificação	35,45%	19,33%	27,35%	49,50%	24,91%	46,41%	43,44%	23,91%	25,44%	21,82%

5

RESULTADOS

5.1

Testes de robustez do modelo de impacto sobre os preços

Seguindo o artigo de Cont et al. [3] fez-se alguns testes de robustez para o modelo de impacto sobre os preços. A Equação 3 foi estimada adicionando o desequilíbrio do fluxo de ordens de dez níveis do *book* de ordens um a um, até que a equação tivesse 10 variáveis dependentes. Estimou-se também a Equação 3 dez vezes, cada uma delas com o desequilíbrio do fluxo de ordens somando os níveis do *book* um a um, até o décimo. Os resultados são apresentados nas Tabelas 3 e 4, respectivamente.

As variáveis adicionais no primeiro teste não forneceram uma melhoria significativa em termos de ajuste e seus coeficientes são estatisticamente insignificante. No segundo teste, houve melhora no ajuste para a soma dos OFIs até o terceiro nível. Assim, na análise de impacto sobre os preços de cada categoria considerou-se o desequilíbrio do fluxo de ordens agregado dos três primeiros níveis do *book* de ordens.

Cont et al [3] estimou um R^2 médio de 65% para uma grande amostra de ações do mercado americano, um valor bem acima do valor encontrado para o mercado brasileiro de futuro de dólar. Esta diferença se dá, possivelmente, pelas particularidades do mercado cambial brasileiro, no qual os bancos realizam suas operações no mercado futuro e transferem ao longo do dia suas posições para o mercado à vista, dada a maior liquidez daquele.

Tabela 3: Resultado da regressão multivariada com desequilíbrios no fluxo de ordens dos dez níveis do *book* – média de todas as subamostras T

Nível	R ²	Incremento no R ²	t(β) médio
1	25,19%		3,41
2	29,40%	0,04%	1,25
3	34,19%	0,05%	0,48
4	37,54%	0,03%	-0,57
5	39,67%	0,02%	-0,83
6	41,19%	0,02%	-0,62
7	42,79%	0,02%	-0,35
8	44,27%	0,01%	-0,43
9	46,17%	0,02%	-0,53
10	48,13%	0,02%	-0,41

Tabela 4: Resultado da regressão univariada com desequilíbrios no fluxo de ordens dos dez níveis do *book* – média de todas as subamostras T

Nível	R ²	Incremento no R ²	t(β) médio
1	25,19%		3,98
1-2	25,47%	0,28%	4,33
1-3	26,77%	1,30%	4,64
1-4	24,22%	-2,55%	4,28
1-5	21,11%	-3,10%	3,81
1-6	19,53%	-1,58%	3,65
1-7	18,89%	-0,64%	3,55
1-8	17,90%	-1,01%	3,45
1-9	18,30%	0,37%	3,56
1-10	17,65%	-0,67%	3,49

As estatísticas-t dos coeficientes foram estimadas com Newey-West, dado que os resíduos das regressões apresentaram heterocedasticidade. Para verificar a existência de dependência não-linear no modelo, adicionou-se um termo quadrático $OFl_{t,T} * |OFl_{t,T}|$ como variável adicional na Equação 1. Esta variável apresentou-se estatisticamente insignificante a 5% de significância na maioria das subamostras, confirmando a linearidade na relação entre mudança nos preços e desequilíbrio do fluxo de ordens. Esta linearidade é uma característica importante do modelo, pois permite sintetizar facilmente o impacto das ordens sobre o preço com um único coeficiente β_T , e comparar o impacto de diferentes grupos de investidores através de um único conjunto de coeficientes de inclinação linear. A alta e significativa correlação entre mudanças nos preços e desequilíbrio no fluxo de ordens em alta frequência pode ser visualizada na Figura 7.

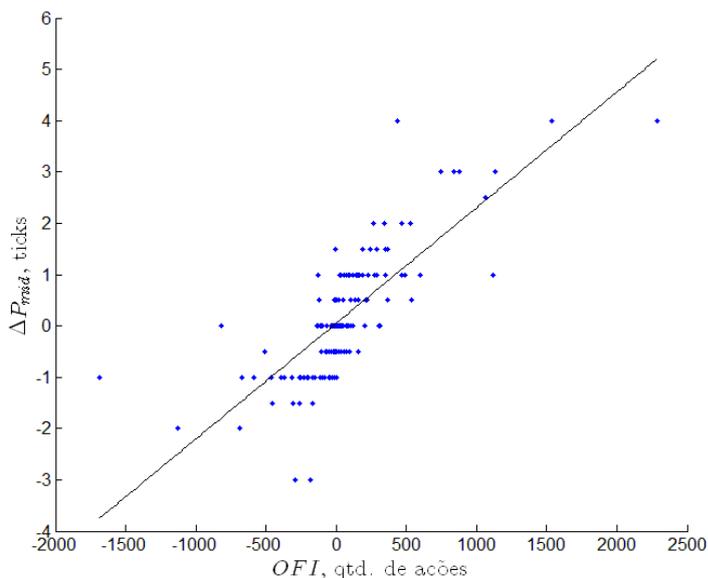


Figura 7: Retornos e desequilíbrios do fluxo de ordens (OFI)

A profundidade do *book* apresenta um padrão de comportamento intradiário contrário ao comportamento dos coeficientes de impacto sobre os preços, em linha com as previsões do modelo. Para demonstrar este comportamento, calculamos a média dos coeficientes de impacto e profundidade do *book* para cada 16 minutos de negociação de todos os dias amostrados. Os resultados são apresentados na Figura 8.

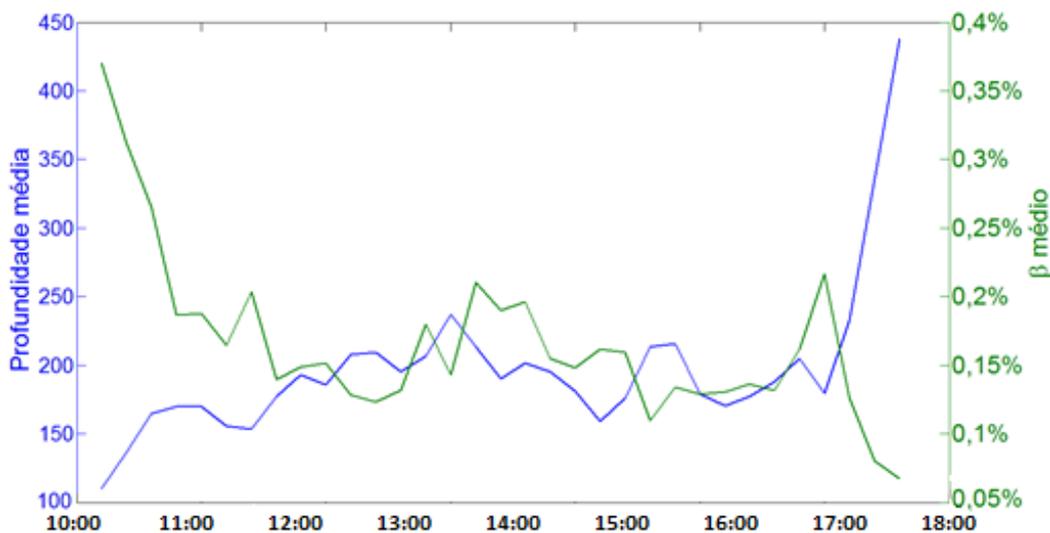


Figura 8: Comportamento intradiário dos coeficientes de impacto sobre o preço e profundidade do *book* de ordens.

Esta relação inversa implica que quando a profundidade é baixa, o impacto das ordens no preço é grande, e até mesmo um pequeno desequilíbrio no fluxo de ordens pode resultar em uma flutuação de preço significativa.

A mesma relação inversa pode ser observada através dos resultados médios da estimação da Equação 2 para todas as subamostras: $\eta = -1,14$, $t(\eta) = -17,30$ e $R^2 = 28\%$. A relação entre impacto nos preço e profundidade do *book* é mostrada na Figura 9.

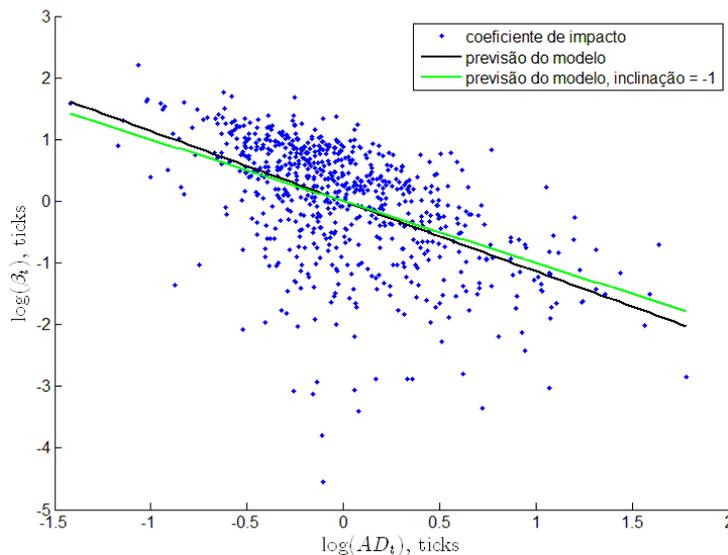


Figura 9: Coeficiente de impacto sobre preço e profundidade do mercado.

5.2

Impacto no preço por categoria de investidores

O modelo linear de impacto sobre os preços de Cont et al. [3] permite comparar o impacto das diferentes categorias de investidores no preço. Como o impacto no preço neste modelo é caracterizado por um único coeficiente de inclinação, pode-se reestimar a Equação 1 substituindo $OFI_{t,T}$ por $OFI_{k,t,T}$ (desequilíbrio do fluxo de ordens da categoria de investidor k no intervalo t) e comparar o coeficiente de impacto estimado entre as diferentes categorias.

No entanto, existem duas questões metodológicas que fazem esta comparação direta dos coeficientes da regressão inadequada. Primeiro, nota-se que o coeficiente angular de uma regressão MQO univariada depende por construção da $\text{var}[OFI_{k,t,T}]$, que é diferente para as diferentes categorias de investidores. Para que os coeficientes sejam comparáveis, as variáveis de desequilíbrio das diferentes categorias foram padronizadas. Após a padronização, o coeficiente angular de uma regressão univariada se torna a correlação de

Pearson entre $\Delta P_{t,T}$ and $OFI_{k,t,T}$. Segundo, a estimação da correlação entre $\Delta P_{t,T}$ e o desequilíbrio do fluxo de ordens de uma dada categoria de investidor pode estar enviesada pela influência de uma terceira variável, por exemplo, o desequilíbrio no fluxo de ordens de outra categoria. Para controlar esta influência, estimou-se uma regressão multivariada com o desequilíbrio do fluxo de ordens de todas as categorias de investidores.

Como todas as variáveis são padronizadas, os coeficientes estimados na regressão multivariada podem ser interpretados como correlações parciais entre as mudanças de preços e os desequilíbrios do fluxo de ordens de cada categoria de investidor sem a influência das outras categorias.

Os resultados da estimação das duas regressões, utilizando como variável independente o desequilíbrio de ordens nos três primeiros níveis do *book*, são apresentadas nas Tabelas 5 e 6. Observa-se que as correlações parciais são menores que as correlações simples, sugerindo a existência de correlações não triviais entre diferentes $OFI_{k,t,T}$.

Os investidores HFT 1 apresentaram o maior impacto na variação de preço, em ambas as regressões. Seus coeficientes de impacto no preço são maiores e significantes em mais da metade das subamostras. Este resultado evidencia a alta capacidade destes investidores de impactar o preço, mas ainda não temos evidência da natureza deste impacto. Este impacto pode ser benéfico para o mercado, ou seja, ele contribui para que o preço de negociação se aproxime do preço fundamental, ou ser apenas um ruído do padrão de negociação que estes investidores apresentam.

Tabela 5: Resultado da regressão com desequilíbrios do fluxo de ordens de cada categoria de investidor

Grupo de Investidor	β_T	Estatística-t (média das subamostras)	subamostras significantes (%)
Inativos	31%	2.87	63,70%
HFT 1	35%	3.40	68,29%
HFT 2	20%	1.94	40,17%
HFT 3	4%	1.93	01,58%
HFT 4	4%	3.16	12,34%
Algos 1	26%	3.17	35,29%
Algos 2	7%	1.19	14,20%
Algos 3	19%	2.38	48,92%
Algos 4	18%	2.58	52,08%
Algos lento e humanos	19%	2.15	42,90%

Tabela 6: Resultado da regressão com desequilíbrios do fluxo de ordens de todas as categorias de investidor

Grupo de Investidor	β_T	Estatística-t (média das subamostras)	subamostras significantes (%)
Inativos	19%	2.23	50,07%
FT 1	20%	2.25	53,95%
FT 2	11%	1.24	31,13%
FT 3	2%	0.61	01,29%
FT 4	1%	0.11	09,61%
lgos 1	19%	2.14	32,14%
lgos 2	4%	0.70	13,34%
lgos 3	11%	1.64	36,59%
lgos 4	12%	1.83	39,74%
lgos lento e humanos	13%	1.67	35,29%

5.3

Contribuição de cada categoria para a volatilidade fundamental

A contribuição de cada categoria de investidor para a formação do preço fundamental foi apreendida através da volatilidade fundamental calculada a partir do modelo de decomposição da variância de Hasbrouk [4].

O modelo foi estimado para 3 intervalos a fim de controlar a não estacionariedade intradiária na variação dos preços. Os intervalos considerados foram 1 hora, 3 horas e todo o dia de negociação. Como os resultados são robustos para todos os intervalos considerados, apresentamos apenas os resultados para os intervalos de 1 hora.

Dividindo os 21 dias de negociação em subamostras de 1 hora, temos um total de 189 subamostras. O modelo VAR foi estimado com lags de 1 a 15 para

cada sub-amostra, e então o melhor lag foi selecionado através do critério AIC. Para uniformizar o modelo VAR a ser estimado para todas as sub-amostras escolhemos 10 defasagens ($p = 10$) como padrão. Com esta quantidade de defasagens o modelo VAR estimado é estável e aceitamos a hipótese de não correlação dos erros para 137 subamostras. A volatilidade fundamental é calculada a partir de 100 defasagens do modelo VMA.

Os resultados encontrados para a contribuição de cada categoria de investidor para a volatilidade fundamental por sub-amostra, apresentados na Tabela 7, mostram a relevância dos investidores HFT 1 para a formação do preço fundamental.

As outras categorias que mais contribuem para a formação do preço fundamental são os Inativos, HFTs 2, Algos 3 e 4 e Algos lentos e humanos. Este resultado evidencia a função de intermediação dos investidores mais rápidos. Pode-se inferir que como os investidores mais rápidos facilitam a negociação dos investidores mais lentos, na maioria das vezes investidores fundamentalistas, eles contribuem para a formação do preço fundamental.

Tabela 7: Contribuição de cada categoria de investidor para a volatilidade fundamental

Grupo de Investidor	Contribuição Relativa para a Volatilidade Fundamental	
	média das sub-amostras	mediana das sub-amostras
Informação Pública	67%	69%
Inativos	5%	2%
HFT 1	10%	6%
HFT 2	5%	2%
HFT 3	0%	0%
HFT 4	1%	0%
Algos 1	0%	0%
Algos 2	1%	0%
Algos 3	4%	2%
Algos 4	3%	1%
Algos lentos e humanos	4%	1%

CONCLUSÃO

Os investidores HFTs não constituem uma entidade coerente. Suas estratégias podem variar com o estilo de negociação, nível de latência, opções de segurança, comércio regional e nuances de regulamentação e o estado do mercado. Por isso, é importante abordar a sua interação com outros participantes e sua contribuição para o impacto sobre a qualidade e a eficiência do mercado com essas diferenças em mente. Este é um exercício desafiador para os pesquisadores, dado que os parâmetros de controle não podem ser facilmente isolados por causa dos efeitos significativos da interação entre eles.

Utilizam-se as ordens enviadas para o book de DOLQ13, futuro de dólar com vencimento em 1º de agosto de 2013, da Bolsa de Valores e Mercadorias e Futuros (BM&FBovespa) para analisar o impacto dos investidores HFTs no processo de formação de preço deste mercado. Foram consideradas as ordens enviadas no mês anterior ao vencimento, uma vez que o primeiro vencimento é o mais líquido deste contrato.

Esse trabalho analisou 10 categorias de investidores, que se diferenciam por suas características de latência. O objetivo é diferenciar os investidores HFTs dos algoritmos, sendo que estes podem estar atuando como fundamentalista ou qualquer outro tipo de investidor, mas em velocidade acima da considerada possível para humanos. Cada uma destas 2 grandes classes são compostas de 4 subclasses, sendo os HFTs 1 os investidores mais rápidos. As classes HFTs 4 e Algos 4 são investidores conhecidamente HFTs ou algoritmos, respectivamente, mas que não apresentam comportamento padrão.

Dentre as categorias de investidores HFTs, a categoria HFT 1 foi a categoria que mais impactou o preço de negociação do dólar futuro e foi também a categoria com maior contribuição para volatilidade fundamental. Assim como a categoria HFT 2, mas em menor intensidade. Os investidores mais lentos, apesar de não terem apresentado coeficientes de impacto sobre o preço relevantes, apresentaram contribuição para a volatilidade fundamental comparável à categoria HFT 2.

Os resultados evidenciam um impacto positivo da interação dos investidores de alta frequência no ambiente de negociação. Em geral, os investidores mais lentos são investidores fundamentalistas, cujas demandas de compra e venda não estão perfeitamente sincronizadas no tempo. A presença de investidores mais rápidos permite que este descasamento entre demanda e oferta no tempo seja resolvido. Ou seja, os investidores de alta frequência tem o papel de intermediadores dos investidores fundamentalista, facilitando o processo de formação de preço.

Apesar deste resultado evidenciar um aspecto positivo da incorporação de investidores de alta frequência no mix de participantes do mercado futuro de dólar, para determinar que sua presença é benéfica é necessário analisar como seu comportamento contribui para a volatilidade ruidosa, tema para pesquisa futura.

A literatura teórica prevê que, sob diferentes pressupostos, os investidores HFTs tanto podem contribuir para a descoberta de preços (por exemplo, reagir mais rapidamente às notícias como em Martinez et al. [31]) como podem contribuir para o ruído (por exemplo, reagindo simultaneamente a sinais comuns e criando pressão sobre preços, como em Jarrow et al. [32]).

As implicações da atividade destes investidores no mercado dependem do benefício marginal da informação em alguns momentos, comparado a desutilidade marginal do excesso de volatilidade em outros momentos. Dependerá também dos tipos de HFTs que a estrutura de mercado vai privilegiar e de como a atividade de negociação e seus negócios impactam no processo de formação de preço.

Os investidores HFTs não surgiram espontaneamente. As mudanças tecnológicas e regulatórias foram facilitadores, além da mudança no mix de participantes do mercado, que criaram novos nichos de oportunidade de lucro que estes investidores podem explorar. Os mercados vão continuar a evoluir. Pesquisadores e participantes de mercado devem continuar a desenvolver sua capacidade de investigação e participar em debates sobre a estrutura de mercado adequada.

REFERÊNCIAS

- [1] VENTURA, A. e GARCIA, M. (2012): “Mercados futuro e à vista de câmbio no Brasil: o rabo abana o cachorro,” *Revista Brasileira de Economia*. 66(1), 21-48.
- [2] BOUCHAUD, J. P. (2010). Price impact. In.: ____Encyclopedia of quantitative finance.
- [3] CONT, R.; KUKANOV, A. and STOIKOV, S. (2013): “The price impact of order book events,” *Journal of financial econometrics*. 12(1), 47-88.
- [4] HASBROUCK, J. (1991b): “The summary informativeness of stock trades: An econometric analysis,” *Review of Financial Studies*. 4(3), 571-595.
- [5] ENGLE, R.; FERSTENBERG, R. and RUSSELL, J. (2006): Measuring and modeling execution cost and risk,” *NYU Working Paper*. N° FIN-06-044.
- [6] ALMGREN, R.; THUM, C.; HAUPTMANN, E. e LI, H. (2005): “Direct estimation of equity market impact,” *Risk*. 18, 5752.
- [7] EVANS, M.D. and LYONS, R.K. (1999): “Order flow and exchange rate dynamics,” *National bureau of economic research*. No. w7317.
- [8] GABAIX, X.; GOPIKRISHNAN, P.; PLEROU, V. and STANLEY, H.E. (2003): “A theory of power-law distributions in financial market fluctuations,” *Nature*. 423(6937), 267-270.
- [9] HASBROUCK, J. (1991a): “Measuring the information content of stock trades,” *The Journal of Finance*. 46(1), 179-207.
- [10] KEIM, D.B. and MADHAVAN, A. (1996): “The upstairs market for large-block transactions: Analysis and measurement of price effects,” *Review of Financial Studies*. 9(1), 1-36.
- [11] KEMPF, A. and KORN, O. (1999), “Market depth and order size,” *Journal of Financial Markets*. 2, 29-48.
- [12] TORRE, N.G. and FERRARI, M. (1997): Market impact model handbook. BARRA Inc, Berkeley.
- [13] PLEROU, V.; GOPIKRISHNAN, P.; GABAIX, X. and STANLEY, H.E. (2002): “Quantifying stock-price response to demand fluctuations,” *Physical Review E*. 66(2), 027104.

- [14] POTTERS, M. and BOUCHAUD, J.P. (2003): “More statistical properties of order books and price impact,” *Physica A: Statistical Mechanics and its Applications*. 324(1), 133-140.
- [15] ENGLE, R.F. and LUNDE, A. (2003): “Trades and quotes: a bivariate point process.” *Journal of Financial Econometrics*. 1(2), 159-188.
- [16] AHN, H.J.; BAE, K.H. and CHAN, K. (2001): “Limit orders, depth, and volatility: evidence from the stock exchange of Hong Kong,” *The Journal of Finance*. 56(2), 767-788.
- [17] ANDERSEN, T.G. and BOLLERSLEV, T. (1998): “Deutsche mark–dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies,” *Journal of Finance*. 53(1), 219-265.
- [18] LEE, C.M.; MUCKLOW, B. and READY, M.J. (1993): “Spreads, depths, and the impact of earnings information: An intraday analysis,” *Review of Financial Studie*. 6(2), 345-374.
- [19] McINISH, T.H. and WOOD, R.A. (1992): “An analysis of intraday patterns in bid/ask spreads for NYSE stocks,” *Journal of Finance*. 47(2), 753-764.
- [20] MADHAVAN, A.; RICHARDSON, M. and ROOMANS, M. (1997): “Why do security prices change? A transaction-level analysis of NYSE stocks,” *Review of Financial Studies*. 10, 1035-1064.
- [21] JONES, C.M.; KAUL, G. and LIPSON, M.L. (1994): “Transactions, volume, and volatility,” *Review of Financial Studies*. 7(4), 631-651.
- [22] Securities and Exchange Commission. (2010): “Concept release on equity market structure,” *Federal Register*. 75(13), 3594-3614.
- [23] HENDERSHOTT, T.; JONES, C.M. and MENKVELD, A.J. (2011): “Does algorithmic trading improve liquidity?,” *The Journal of Finance*. 66(1), 1-33.
- [24] BOEHMER, E.; FONG, K. and WU, J. (2012): “International evidence on algorithmic trading,” *Available at SSRN 2022034*.
- [25] KITTIKARASAKUN, J.; TSE, Y. and WANG, G. (2012): “The impact of trades by traders on asymmetric volatility for Nasdaq-100 index futures,” *Managerial Finance*. 38(2), 752-767.
- [26] HAGSTRÖMER, B. and NORDÉN, L. (2013): “The diversity of high-frequency traders,” *Journal of Financial Markets*. 16(4), 741-770.
- [27] BENOS, E. and SAGADE, S. (2012): “High-frequency trading behaviour and its impact on market quality: evidence from the UK equity market,” *Bank of England*. N° 469.

- [28] BARON, M.; BROGAARD, J. and KIRILENKO, A. (2012): “The trading profits of high frequency traders,” *Working paper*.
- [29] KEARNS, M.; KULESZA, A. and NEVMYVAKA, Y. (2010): “Empirical limitations on high frequency trading profitability,” *arXiv preprint arXiv:1007.2593*.
- [30] HASBROUCK, J. and SAAR, G. (2013): “Low-latency trading,” *Journal of Financial Markets*. 16(4), 646-679.
- [31] MARTINEZ, V.H. and ROSU, I. (2011): “High frequency traders, news and volatility,” *SSRN eLibrary*.
- [32] JARROW, R.A. and PROTTER, P. (2012): “A dysfunctional role of high frequency trading in electronic markets,” *International Journal of Theoretical and Applied Finance*, 15(03).