



Instituto Nacional de Matemática Pura e Aplicada

Impacto de Preço no Mercado Brasileiro

Autor: **Antônio Augusto Cardoso Correia**

Orientador: **Luca Mertens**

Rio de Janeiro
Março de 2014

Dedico este trabalho a minha família e aos meus amigos.

Agradecimentos

Agradeço a todos aqueles que me ajudaram e contribuíram para o desenvolvimento deste trabalho. Faço menção especial ao professor Jorge Zubelli responsável pelo mestrado em finanças do IMPA e ao meu orientador Luca Mertens pelo apoio e as orientações que ajudaram a engrandecer este trabalho. Também, gostaria de agradecer a minha família por estarem sempre me apoiando nos momentos mais difíceis.

Resumo

O impacto de mercado é a alteração no processo de formação do preço gerada pela submissão de uma ordem de compra ou venda. A execução de um grande lote é uma operação que tipicamente movimenta o preço na direção desfavorável à operação, baixando o preço na execução de uma ordem de venda e aumentando o preço na execução de uma ordem de compra. Ter um modelo sobre o impacto de mercado é, portanto fundamental para a calibração de uma estratégia de execução que busque minimizar o custo de execução. Neste trabalho, foram estudados modelos econométricos de impacto de preço que relacionam a variação do preço com a variação da demanda e da oferta no livro de ordens. Mais especificamente, foi aplicado o modelo proposto por Cont *et al.* ao mercado brasileiro. Os resultados das análises, de acordo com a literatura, mostram que a variável dada pelo fluxo líquido de ordens é uma boa variável explicativa para as mudanças de preço em alta frequência. O coeficiente de impacto de preço foi estimado em diferentes períodos do dia, com isso foi observado uma grande variabilidade intra-diária. Também, foi feita uma análise econométrica na série temporal deste parâmetro, o que mostrou evidências de autocorrelação e sazonalidade. Um modelo SARIMA foi calibrado na série deste coeficiente mostrando-se coerente para capturar a variabilidade intra-diária do impacto no preço. O resultado deste trabalho sugere que assumir o impacto de preço constante durante o dia pode afetar negativamente a performance de estratégias de execução. Além disso, foi apresentada uma estratégia de execução ótima que incorpora esta dinâmica intra-diária.

Palavras-chave: Impacto de mercado, execução ótima, fluxo de ordens, livro de ofertas, microestrutura de mercado, liquidez, negociação em alta frequência.

Sumário

Índice	vii
Introdução	1
1 Revisão bibliográfica	3
1.1 Modelos de impacto de mercado	3
1.1.1 O impacto linear e permanente: o modelo de Kyle	4
1.1.2 O impacto não é tão simples	4
1.1.3 Fluxo líquido de ordens e o impacto de preço	6
1.2 Estratégia de execução ótima	10
1.2.1 Definição da estratégia	10
1.2.2 Fronteira eficiente para execução ótima	13
2 Estimação mercado brasileiro	17
2.1 Dados	17
2.2 Definição das variáveis	17
2.3 Resultados empíricos	20
3 Análise temporal	23
3.1 Autocorrelação no impacto de preço	24
3.2 Aplicações	28
4 Conclusão	31
A Provas	33
B Códigos	35
C Tabelas	43
Referências Bibliográficas	43

Introdução

Recentemente houve uma digitalização dos principais mercados mundiais que se tornaram eletrônicos. Essa evolução possibilitou o estudo de dados em alta frequência, sendo possível obter dados históricos de cada evento do livro de ofertas: novas ordens, cancelamentos e negócios. Isto estimulou a surgir uma extensa literatura empírica e teórica relacionada ao fluxo de ordens, liquidez e impacto nos preços. Uma questão prática importante é o impacto das ordens nos preços, que surge em problemas de execução de grandes blocos de ordem em um horizonte de tempo fixo. Entender o impacto nos preços também é importante no sentido teórico, já que é um mecanismo fundamental de formação do preço.

O impacto de mercado refere-se ao fato da execução de uma ordem influenciar o preço do ativo subjacente. Normalmente, esta influência resulta em um efeito adverso criando custos de execução adicionais. O entendimento do impacto de mercado e a otimização de estratégias de execução para minimizar o impacto tem sido um importante objetivo de investidores institucionais. Há tipicamente liquidez insuficiente que permita executar grandes ordens sem entrar em vários níveis de preço do livro de ordens. Por causa disso, para minimizar os custos de negociação, grandes ordens são divididas em uma sequência de pequenas execuções, que são espalhadas ao longo de um certo intervalo de tempo. Outras questões importantes estão relacionadas a dependência do impacto com o volume (volumes maiores impactam mais o preço?), e o comportamento temporal do impacto, se o impacto é apenas imediato e permanente, ou se existe alguma dependência anterior do impacto [9].

O trabalho de Cont *et al.* [12] fez um estudo empírico do impacto de mercado utilizando os eventos do livro de ofertas no mercado de ações americano. Foi utilizada uma única variável, que é calculada através do fluxo líquido de ordens das melhores ofertas de compra e venda para estimar o impacto de preço. Os resultados obtidos mostraram que esta variável agregada explica variações no preço sob intervalos curtos de tempo para uma grande amostra de ações com uma média de R^2 de 65% [12]. A maior variabilidade do impacto de preço, para diferente escalas de tempo e ativos, estavam relacionadas as variações na profundidade do livro de ofertas. De fato, foi estabelecida uma relação inversa entre o coeficiente de impacto no preço e a profundidade do livro de ofertas ou liquidez implícita do mercado.

Tendo em vista os altos custos relacionados as execuções e os grandes volumes negociados no mercado brasileiro, o propósito deste trabalho é realizar um estudo do impacto de preço no mercado local. Foi aplicado o modelo proposto em [12], realizando estimações do coeficiente de impacto de preço nos principais mercados brasileiros: futuro de juros, futuro de dolar, futuro do índice Ibovespa e principais ações do Ibovespa. Além disso, também foi feita uma análise ecométrica do comportamento intra-diário do impacto de mercado, esta análise mostrou evidências de autocorrelação e sazonalidade deste parâmetro. O resultado deste trabalho sugere que assumir o impacto de preço constante durante o dia pode afetar a performance de estratégias de execução. Uma aplicação prática apresentada foi uma estratégia de execução ótima que considera a variação do impacto de preço durante o dia.

A estrutura deste trabalho segue a seguinte ordem. No capítulo 2 é feita uma revisão

bibliográfica introduzindo os principais conceitos relativos aos modelos de impacto de mercado e a relação deste com estratégias de execução. No capítulo 3 são apresentados os resultados obtidos das estimações no mercado brasileiro, a análise econométrica realizada na série deste parâmetro e uma extensão da estratégia de execução ótima para o caso de liquidez (impacto) estocástico. No capítulo 4 é apresentada uma conclusão. Todos as provas estão no apêndice [A](#), os códigos utilizados nas estimações estão no apêndice [B](#) e as tabelas dos resultados no apêndice [C](#).

Capítulo 1

Revisão bibliográfica

Neste capítulo, serão apresentados os principais assuntos relativo ao impacto de mercado. Além disso, será introduzido o modelo de Cont *et al.* [12] que relaciona a variação de preço com os eventos do livro de ordens. Uma das questões fundamentais do problema de execução é o impacto que a própria execução causa nos preços. Por causa disso, também será apresentado o problema de execução ótima do trabalho de Almgren-Chriss [3].

1.1 Modelos de impacto de mercado

O impacto de mercado é o efeito que o participante realiza quando ele compra ou vende um ativo. Normalmente, esta influência resulta em um efeito adverso criando custos adicionais. Em uma operação de compra movendo o preço contra o comprador, aumentando o preço na execução, e em uma operação venda baixando o preço na execução. O impacto de uma ordem está muito relacionado com a liquidez. Especialmente para grandes investidores, o impacto é uma fator essencial a ser considerada antes de qualquer decisão de movimentação financeira entre mercados. Se a quantidade financeira a ser é movimentada é grande (relativa ao giro deste ativo em questão), então o impacto de mercado pode ser alguns pontos percentuais e necessita de ser avaliado além de outros custos de transação.

O impacto de mercado pode surgir porque o preço necessita mover para atender outros investidores para comprar ou vender o ativo (contrapartes), mas também porque investidores profissionais podem se posicionar para lucrar do conhecimento que um grande investidor que está ativo em uma direção. A instituição financeira que está buscando gerenciar o impacto de mercado necessita limitar o ritmo da sua atividade (por exemplo mantendo a atividade abaixo de um certo nível do volume diário) para evitar distorções no preço. Para medir o impacto de mercado existem várias medidas estatísticas, a medição do impacto de mercado é, no entanto, uma tarefa muito difícil e erros de estimação podem afetar bastante estratégias de execução ótima.

Esta seção será dividida em três subseções. Primeiramente, será apresentado o modelo de impacto mais simples o λ de Kyle. Na segunda subseção, serão mostradas algumas considerações que foram feitas no trabalho do Bouchaud [9] em relação a dependência do volume e o comportamento temporal do impacto. Na ultima subseção, é introduzido o modelo apresentado no trabalho [12] que utiliza uma variável dada pelo fluxo de liquido das ordens do primeiro nível de preço do livro de ofertas para realizar a estimação do impacto de preço.

1.1.1 O impacto linear e permanente: o modelo de Kyle

A hipótese mais simples é que o impacto é linear com volume negociado e permanente no tempo. Esta suposição pode ser parcialmente justificada dentro do modelo de Kyle [23], no qual um agente informado e um agente não-informado (ruído) submetem ordens que são executadas pelo formador de mercado (MM-*market maker*) em cada intervalo de tempo Δt . Neste modelo, a regra de ajuste de preço Δp do MM é linear com o total de volume com sinal negociado ϵv , isto é ¹

$$\Delta p = \lambda \epsilon v, \quad (1.1.1)$$

onde λ é a medida de impacto, e é inversamente proporcional a liquidez do mercado. Além disso, este ajuste de preço é *permanente*, isto é, a mudança de preço entre o tempo $t = 0$ e $t = T = N\Delta t$ é:

$$p_T = p_0 + \sum_{n=0}^{N-1} \Delta p_n = p_0 + \lambda \sum_{n=0}^{N-1} \epsilon_n v_n. \quad (1.1.2)$$

Esta fórmula assume que o impacto $\lambda \epsilon_n v_n$ dos negócios dos n -ésimos intervalos de tempo persiste até o tempo T . [20] e [15] mostraram que os ajustes de preço linear do modelo de Kyle é a única especificação que não permite manipulações de preço. Da equação acima, é claro que os sinais dos negócios devem ser serialmente descorrelacionados se o preço segue um passeio aleatório não previsível. Dentro da configuração do modelo de Kyle, o cronograma de execução de um agente informado é tal que ϵ_n são descorrelacionados [23]. No entanto, dados reais de mercado mostram correlação no sinal do volume negociado em longos intervalos de tempo [9].

1.1.2 O impacto não é tão simples

Esta seção irá apresentar um breve resumo de alguns fatos empíricos que surgiram nos últimos 20 anos, em relação a dependência do volume e o comportamento temporal do impacto apresentadas no trabalho [9].

Medidas de impacto de preço

Uma medida interessante do impacto de preço é a correlação entre a mudança de preço entre 0 e T e o sinal do negócio no tempo 0, normalmente definida por:

$$R(T) = E[(p_T - p_0) \cdot \epsilon_0] - E[(p_T - p_0)]E[\epsilon_0].$$

Também é assumido que o *drift* pode ser negligenciado então somente o primeiro termo é diferente de zero. Este é o caso para T escolhido suficientemente pequeno e/ou não existe nenhuma assimetria de compra/venda ($E[\epsilon_0] = 0$).

Dentro do modelo de Kyle que ϵ é descorrelacionado. $R(T)$ é facilmente calculado e é dado por:

$$R(T) = \lambda E[v].$$

Neste modelo, a função de impacto $R(T)$ é então independente de valores anteriores. Pode-se também definir a função de impacto dependente do volume condicionando ao volume de chegada, isto é:

$$R(T, v) = E[(p_T - p_0) \cdot \epsilon_0 | v_0 = v],$$

¹ $\epsilon = +1$ se o volume de compra v_b é maior do que o volume de venda v_s no intervalo de tempo Δt , e $\epsilon = -1$ é o caso oposto, e $v = |v_b - v_s|$

que, no modelo de Kyle, é igual a λv , para todo T .

Outra medida de impacto comumente usada é a correlação $\rho(T)$ entre a mudança de preço de 0 a T e o volume total com sinal no mesmo intervalo, ou mais precisamente:

$$\rho_T = \frac{E[(p_T - p_0) \cdot \sum_{n=0}^{N-1} \epsilon_n v_n]}{\sqrt{E[(p_T - p_0)]^2 E[(\sum_{n=0}^{N-1} \epsilon_n v_n)^2]}} \quad (1.1.3)$$

No modelo de Kyle esta correlação é trivialmente igual a 1, mas esta correlação pode diminuir se a liquidez flutua (λ se torna dependente do tempo), ou se outras fontes de volatilidades são levadas em conta. Por exemplo, os MM podem revisar as suas cotações na ausência de qualquer negócio, por causa de alguma notícia que está próxima de ser anunciada. Neste modelo pode ser adicionado um termo extra na equação 1.1.2 acima:

$$p_T = p_0 + \sum_{n=0}^{N-1} \Delta p_n = p_0 + \lambda \sum_{n=0}^{N-1} \epsilon_n v_n + \sum_{n=0}^{N-1} \eta_n,$$

onde η_n são incrementos descorrelacionados, representando variações de preço não diretamente relacionadas a execução.

Normalmente, é útil generalizar a definição acima de $\rho(T)$ substituindo os volume v_n por v_n^ψ , onde ψ é um certo expoente.

Dependência concava do volume

O modelo de Kyle assume uma dependência linear do impacto com o volume negociado. No entanto, pode-se medir a função de impacto *instantânea* dependente do volume $R(T = \Delta t, v)$ para diferentes escalas de tempo Δt , variando do tempo médio de uma transação até várias horas ou dias. Também, pode-se assumir que a dependência do impacto com volume é *sublinear* e bem descrita pela lei de potência:

$$R(T = \Delta t, v) \propto v^{\psi(\Delta t)}; \psi(\Delta t) \leq 1.$$

O expoente ψ aumenta com o intervalo de tempo Δt , tomando valores pequenos de $\psi \simeq 0,1-0,3$ para um único negócio, e aumentando em direção de $\psi = 1$ quando Δt corresponde a milhares de negócios, com alguma concavidade para grandes volumes [19], [29]. O coeficiente de correlação $\rho(T = \Delta t)$ similarmente aumenta com Δt , e alcança valores maiores que 0,5 para retornos diários de ações, futuros e moedas [14].

Note que valores pequenos de ψ somente para um negócio significa que o impacto é apenas fracamente dependente do volume, em linha com vários estudos que mostram forte correlação das mudanças de preço com número de negócios ao invés do volume negociado [21]. O valor pequeno de ψ normalmente é interpretado em termos de particularidades da execução: grandes ordens a mercado são submetidas quando há grandes volume dando liquidez na melhor cotação, uma condição que mitiga o impacto de grandes ordens.

O impacto não pode ser permanente

Como observado acima, o impacto permanente leva a mudanças de preço imprevisíveis somente se o sinal do volume é descorrelacionado. No entanto, dados reais mostram que para uma grande variedade de mercados a autocorrelação do sinal ϵ_n decai extremamente devagar com o tempo [9]. Por causa disso, o fluxo de ordens é extremamente persistente e previsível. Isto vem do fato

que mesmo mercados altamente líquidos somente oferecem pequenas quantidades de volume para execução imediata. O fato da liquidez instantânea ser tão pequena tem consequências imediatas: as execuções devem ser fragmentadas, e necessitam de várias horas ou dias para serem completadas. Isto claramente cria memória longa no sinal do fluxo de ordens e mostra que informação privada somente pode ser lentamente incorporada no mercado. Esta observação, no entanto, é incompatível com modelo de impacto permanente da equação 1.1.2, que levaria *trends*, isto é, mudanças de preço altamente correlacionadas.

O impacto não linear e transiente

Em ordem de reconciliar o persistente fluxo de ordem com as quase imprevisíveis mudanças de preço, pode-se postular um modelo de impacto não linear e transiente que generaliza a equação 1.1.2 acima:

$$p_T = p_{-\infty} + \lambda \sum_{n=-\infty}^{N-1} G(N-n)\epsilon_n v_n^\psi$$

onde $G(l)$ descreve o comportamento temporal do impacto. Pode-se mostrar que é possível escolher certo formato de decaimento para função $G(l)$ tal que compense exatamente a autocorrelação do fluxo de ordem e que garanta que as mudanças de preço são um ruído branco.

Assumindo por simplicidade que os volume são iguais: $v_n \equiv v, \forall n$. Pode-se mostrar que se $C(l) = E[\epsilon_n \cdot \epsilon_{n+l}]$ decai para l grande como $l^{-\gamma}$ com $\gamma < 1$ (tipicamente $\gamma \simeq 0,5$ para ações), então $G(l)$ deveria decair para zero como $l^{-\beta}$ com $\beta = (1-\gamma)/2$ [10]. A componente de impacto permanente $G(l \rightarrow \infty) > 0$ somente é compatível com a natureza randômica do preço se $C(l)$ decai suficientemente rápido ($\gamma > 1$). Dentro deste modelo de impacto transiente com volume fixo, a relação entre a função de impacto no preço $R(T)$, $G(l)$, e $C(l)$ ficam [10]

$$R(T = l\Delta t) = \lambda v^\psi [G(l) + \sum_{0 < j < l} G(l-j)C(j) + \sum_{j > 0} [G_0(l+j) - G_0(j)]C(j)].$$

Em outras palavras, o impacto $G(l)$ de um único negócio isolado é diferente do impacto medido diretamente $R(T)$, que pega contribuições do fato que os negócios tendem a se repetir na mesma direção.

Note que mesmo que o impacto $G(l)$ individual de um negócio decaia com o tempo, pode-se mostrar que o impacto total $R(T)$ e a correlação $\rho(T)$ definida em 1.1.3 tendem para um limite diferente de zero quando T é grande sempre que a relação $\beta = (1-\gamma)/2$ mantém-se.

Finalmente, como mencionado acima o modelo de impacto linear, correspondente a $\psi = 1$, é a única opção consistente para que não haja estratégias de manipulação de preço caso impacto seja permanente ($\beta = 0$). No entanto o impacto é transiente, outros valores de $\psi \leq 1$ são aceitáveis. Gatheral recentemente mostrou que se $\beta + \psi \geq 1$, a manipulação de preço não é possível [16].

1.1.3 Fluxo líquido de ordens e o impacto de preço

O trabalho [12] conduziu uma investigação empírica do impacto instantâneo dos eventos do livro de ordens: ordens a mercado, ordens limitadas e cancelamentos no preço das ações. Foi mostrado que o efeito instantâneo no preço destes eventos pode ser modelado através de uma única variável, o *order flow imbalance* (OFI). Esta variável representa o fluxo líquido das ordens no primeiro nível de preço do livro de ordens, isto é: aumentando toda vez que a quantidade

da compra (*bid*) aumenta, a quantidade da venda (*ask*) diminui ou os preços da compra/venda (*bid/ask*) aumentam; e decrementando toda vez que a quantidade da compra diminui, a quantidade da venda aumenta ou preços de compra/venda diminuem.

Curiosamente, esta variável trata ordem a mercado de venda e o cancelamento de uma ordem de compra de mesma quantidade equivalentemente, já que elas possuem o mesmo efeito na quantidade da melhor cotação de compra. Esta variável agregada explicou variações no preço em intervalos curtos de tempo de forma linear, para uma grande amostra de ações americanas, com uma média de R^2 de 65%. Em contraste, o fluxo de ordens mais profundo do livro de ordens não contribui substancialmente para as mudanças de preço. O modelo baseado no OFI relaciona preço, negócios, ordens limitadas e cancelamentos de uma forma simples: ele é linear, requer a estimação de um único *coeficiente de impacto de preço* e é robusto entre ativos e intervalos de tempo.

O coeficiente de impacto de preço possui substancial variabilidade intra diária coincidindo com os padrões intra diários de *spread*, profundidade do mercado e volatilidade de preço [1],[6], [24], [25]. Há uma ligação forte entre volatilidade de preço e o desvio padrão do OFI que sugere que este modelo é uma melhor estimação do λ de Kyle do que estimativas tradicionais baseadas no volume negociado.

Nesta seção será apresentado este modelo mais detalhadamente. Para isto, vai ser introduzido um livro de ordens estilizado, e o método utilizado para a estimação.

Livro de ofertas estilizado

Para entender melhor o impacto de mercado, será apresentado um modelo teórico de um livro de ofertas que será chamado de livro de ofertas estilizado, para ilustrar a relação entre o fluxo de ordens e o impacto no preço.

Considere um livro de ofertas que a quantidade de ações em cada nível de preço além do melhor *bid* e *ask* é igual a D . As chegadas de novas ordens e cancelamentos ocorrem somente no melhor *bid* e *ask*. Além disso, quando o *bid* (ou *ask*) atingem um quantidade D , a próxima chegada de ordem limitada acontece um *tick* acima (ou abaixo) da melhor cotação, criando um novo melhor *bid* (ou *ask*). Considere o intervalo de tempo $[t_{k-1}, t_k]$ e seja L_k^b, C_k^b respectivamente as quantidades totais de ordens de compra que chegam e são canceladas do melhor *bid* durante este intervalo de tempo. Também, definimos M_k^b como a quantidade total das ordens a mercado de compra (agressões) que chegam contra o melhor *ask* durante o intervalo $[t_{k-1}, t_k]$, e P_k^b como o melhor preço do *bid* no tempo t_k . As quantidades L_k^s, C_k^s, M_k^s para as ordens de venda são definidas analogamente e P_k^s como preço do *ask* no tempo t_k .

Neste modelo de livro de ofertas simplificado existe uma relação linear entre o fluxo de ordens $L_k^{b,s}, C_k^{b,s}, M_k^{b,s}$ e as mudanças de preço $\Delta P_k^{b,s} = (P_k^{b,s} - P_{k-1}^{b,s})$ (também ilustradas nas figuras 1.1 - 1.2):

$$\Delta P_k^b = \delta \left[\frac{L_k^b - C_k^b - M_k^s}{D} \right] \quad (1.1.4)$$

$$\Delta P_k^s = -\delta \left[\frac{L_k^s - C_k^s - M_k^b}{D} \right], \quad (1.1.5)$$

Onde δ é o tamanho de um *tick* (mínimo incremento de preço). Fazendo esta mesma análise para as variações de preço do *mid* normalizada pelo *tick* $P_k = \frac{P_k^b + P_k^s}{2\delta}$, obtemos as seguintes relações:

$$\Delta P_k = \frac{OFI_k}{2D} + \epsilon_k, \quad (1.1.6)$$

$$OFI_k = L_k^b - C_k^b - M_k^s - L_k^s + C_k^s + M_k^b \quad (1.1.7)$$

onde OFI_k é o *order flow imbalance* (ou fluxo de ordens líquido) e ϵ_k é o erro de truncamento. Também, podemos reescrever 1.1.6 como:

$$\Delta P_k = \frac{TI_k}{2D} + \eta_k, \quad (1.1.8)$$

$$TI_k = M_k^b - M_k^s \quad (1.1.9)$$

onde TI_k é o desbalanceamento dos negócios e $\eta_k = \frac{L_k^b - C_k^b - L_k^s + C_k^s}{2D} + \epsilon_k$. Quando a atividade das ordens limitadas aumentam, isto é, os valores absolutos dos termos $|L_k^{b,s}|, |C_k^{b,s}|$ são muito maiores do que $|M_k^{b,s}|$, a correlação das mudanças de preço com TI_k são muito mais fracas do que com OFI_k , porque as ordens limitadas e cancelamentos manifestam-se como ruído em 1.1.8.

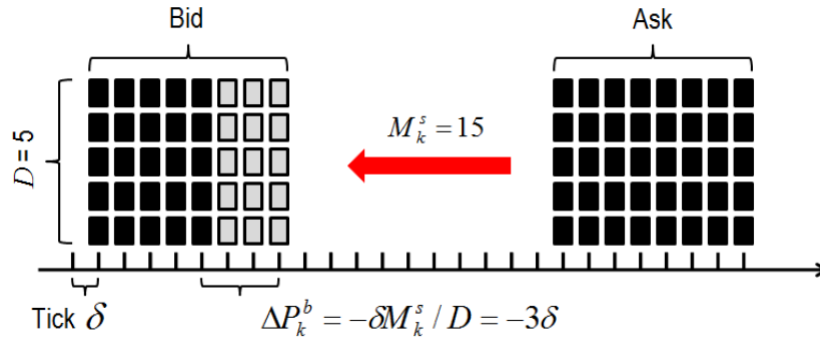


Figura 1.1: Ordem a mercado de venda remove quantidade $M_k^s = 15$ do *bid*

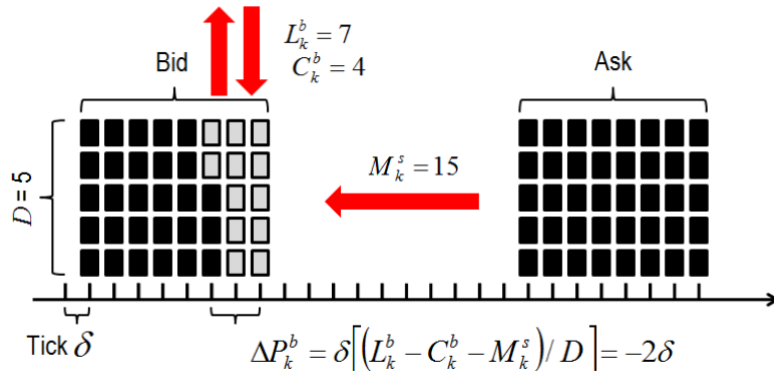
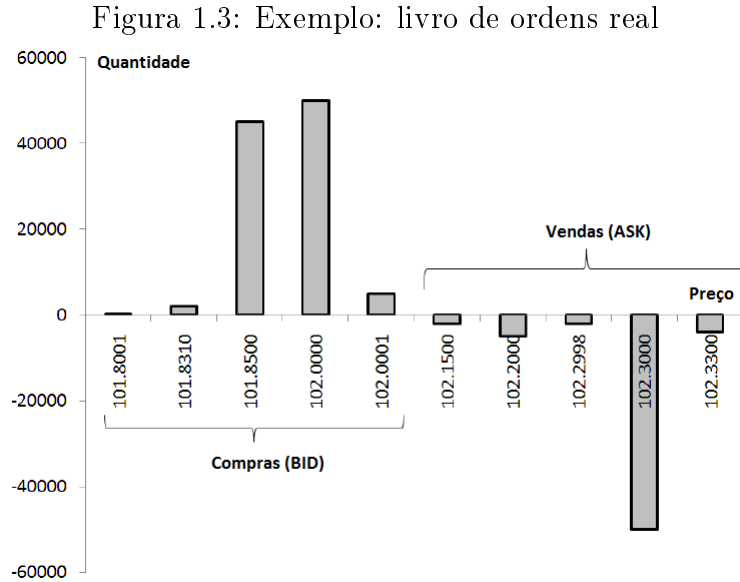


Figura 1.2: Ordem a mercado de venda e cancelamento da compra removem quantidade $M_k^s + C_k^b = 19$ do *bid*, enquanto ordem de compra limitada adiciona $L_k^b = 7$ no *bid*

Especificação do modelo

No mercado financeiro, o livro de ofertas reais possui uma dinâmica muito mais complexa do que a apresentada no livro de ofertas estilizado, podendo ocorrer chegadas e cancelamentos de ordens em todos níveis de preço e a distribuição de profundidade através dos níveis não é constante como a assumida no livro simplificado, a figura 1.1.3 é um exemplo de um livro de ordens real.



Motivado pelo modelo teórico do livro de ofertas estilizado, é assumida uma relação entre as mudanças de preço e o OFI , que se mantém localmente para intervalos de tempo curtos $[t_{k-1,i}, t_{k,i}] \subset [T_{i-1}, T_i]$, onde $[T_{i-1}, T_i]$ são intervalos mais longos [12].

$$\Delta P_{k,i} = \beta_i OFI_{k,i} + \epsilon_{k,i} \quad (1.1.10)$$

Neste modelo, β_i é o coeficiente de impacto no preço para o i -ésimo intervalo de tempo e $\epsilon_{k,i}$ é um termo de ruído que resume a influência de outros fatores. É permitido que β_i e a distribuição de $\epsilon_{k,i}$ se alterem com o índice i , por causa dos efeitos da sazonalidade intradiária. A discussão da seção anterior do livro simplificado permite interpretar $1/2\beta_i$ como a profundidade implícita do livro de ofertas, ou liquidez instantânea do mercado. O livro de ofertas simplificado sugere que o coeficiente de impacto no preço β_i possui uma relação inversa com a profundidade do mercado D_i , e é considerado o seguinte modelo:

$$\beta_i = \frac{c}{D_i^\lambda} + \nu_i, \quad (1.1.11)$$

onde c, λ são constantes e ν_i é o termo de ruído. Para o livro de ofertas estilizado os parâmetros correspondentes são $c = 1/2, \lambda = 1$. Também pode-se considerar a relação entre as mudanças de preço e os negócios:

$$\Delta P_{k,i} = \beta_i^T TI_{k,i} + \eta_{k,i}, \quad (1.1.12)$$

mas espera-se com muito mais ruído do que 1.1.10.

As equações (1.1.10) e (1.1.11) podem ser consideradas como o modelo de impacto no preço dos eventos do livro de oferta que ocorrem no intervalo de tempo $[t_{k-1}, t_k]$. Uma ordem submetida ou cancelada no tempo $\tau \in [t_{k-1}, t_k]$ contribui com uma quantidade e_τ para a

oferta/demanda. Para qualquer intervalo de tempo, estas contribuições são provavelmente desbalanceadas, levando para um fluxo líquido de ordens OFI_k , que afeta a oferta/demanda e implica em um ajuste de preço correspondente.

1.2 Estratégia de execução ótima

O problema de execução ótima foi apresentado por Bertsimas e Lo em [8], em um modelo que o preço do ativo segue um passeio aleatório discreto no tempo e a execução causa um impacto permanente no preço que é linear com a quantidade negociada. Neste modelo a estratégia que minimiza o custo de execução esperado tem velocidade constante. Uma contribuição posterior fundamental, foi feita por Almgren-Chriss em [3] que adicionou o impacto de preço temporário no qual o ativo recupera o preço no próximo período de tempo e daí obteve a melhor estratégia determinística a partir do critério de média-variância, com isso ao levar a variância em questão, um agente averso ao risco executa mais rapidamente do que uma estratégia de velocidade constante. Almgren and Lorenz [4] investigaram estratégias que reagem a mudanças de preço. Schied et al. [31] estudou o problema de execução ótima para um agente com utilidade esperada exponencial em um modelo contínuo no tempo com ruído Lévy e encontrou um modelo contínuo no tempo análogo a estratégia ótima de Almgren-Chriss. A resiliência do impacto de preço com diferentes decaimentos foi estudada em [18], [2] e [27]. Os trabalhos [7] e [26] modelaram o impacto de preço estocástico com um processo de Poisson.

Esta seção irá apresentar o modelo proposto por Almgren-Chriss em [3], será formalizada a definição desta estratégia de execução para o problema de venda (liquidação) de um único ativo. A definição e os resultados são análogas para o problema de compra (aquisição).

1.2.1 Definição da estratégia

Seja um bloco de ordem de quantidade X de um certo ativo que desejamos liquidar completamente antes do tempo T . O tempo T é dividido em N intervalos de tamanho $\tau = T/N$, e $t_k = k\tau$ são tempos discretos de $k = 0, \dots, N$. A trajetória de negociação é definida pela a lista x_0, \dots, x_N onde x_k é a quantidade remanescente no tempo t_k . A quantidade inicial é $x_0 = X$, e a liquidação até o tempo T requer $x_N = 0$.

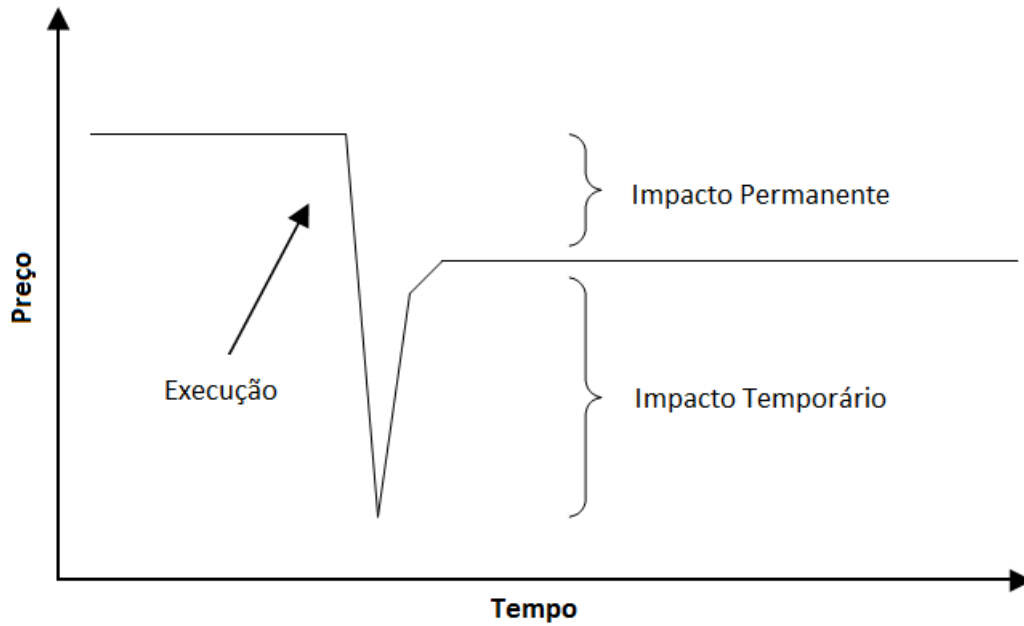
Também, pode ser definida a lista dos negócios n_1, \dots, n_N , onde $n_k = x_{k-1} - x_k$ é a quantidade negociada entre os tempos t_{k-1} e t_k . Claramente, x_k e n_k são relacionadas por:

$$x_k = X - \sum_{j=1}^k n_j = \sum_{j=k+1}^N n_j \quad (1.2.1)$$

Por simplicidade, são tomados todos os intervalos de igual tamanho τ , mas esta restrição não é essencial. Embora isto não seja discutido, em todos resultados obtidos é fácil passar o tempo contínuo fazendo limite de $N \rightarrow \infty, \tau \rightarrow 0$.

Então, é definida a estratégia de execução como a regra de determinar n_k baseado na informação disponível no tempo t_{k-1} . São definidos dois tipos de estratégias: dinâmica e estática. Estratégias estáticas são determinadas no início da execução, ou seja, a regra para determinar n_k depende somente da informação disponível no tempo t_0 . As Estratégias dinâmicas dependem de toda informação incluindo o tempo t_{k-1} .

Figura 1.4: Impacto permanente vs temporário



Dinâmica de preço

Suponha que o preço inicial do ativo seja S_0 , então o valor inicial da posição é $X S_0$. O preço do ativo evolui de acordo com dois fatores exógenos: volatilidade e o *drift*, e um fator endógeno: o impacto de preço. A volatilidade e o *drift* são resultados das forças do mercado que ocorrem aleatoriamente e independentemente da nossa execução.

O impacto de mercado é um mecanismo natural da formação de preço, a medida que os participantes do mercado detectam um grande volume sendo vendido (comprado) eles naturalmente ajustam as ofertas de compra (venda) para baixo (cima). São apresentados dois tipos de impacto de mercado. O impacto de mercado temporário referente ao desbalanceamento temporário da oferta e demanda, ou seja, a nossa execução causa movimentos de preço fora do equilíbrio. O impacto de preço permanente que são mudanças causadas no equilíbrio do preço pela nossa execução, pelo menos até o final da liquidação. A figura 1.4 ilustra a relação destes dois tipos de impacto de mercado.

É assumido, que o preço do ativo evolui de acordo com um passeio aleatório aritmético discreto

$$S_k = S_{k-1} + \sigma \tau^{1/2} \xi_k - \tau g\left(\frac{n_k}{\tau}\right), \quad (1.2.2)$$

para $k = 1, \dots, N$. Aqui σ representa a volatilidade do ativo, o ξ_k são choques independentes da variável aleatória normal com média zero e variância unitária, e o impacto permanente $g(v)$ é função da taxa média de execução $v = n_k/\tau$ durante o intervalo de t_{k-1} a t_k . A equação 1.2.2 não possui o termo de *drift*, é assumido que não possuímos informações da direção dos futuros movimentos de preço.

Impacto de mercado temporário

A intuição por trás do impacto de mercado temporário é que um agente ao planejar vender uma certa quantidade n_k entre os tempos t_{k-1} e t_k , ele pode trabalhar a ordem em vários pedaços pequenos para buscar o ponto ótimo de liquidação. Se a quantidade total n_k é suficientemente

grande, o preço de execução pode diminuir no intervalo de t_{k-1} e t_k , em parte devido a exaustão da liquidez da oferta em cada nível de preço. É assumido que o efeito é curto, e em particular, a liquidez retorna em cada período de tempo e um novo equilíbrio de preço é estabelecido.

Este efeito é modelado introduzindo uma função de impacto de preço temporário $h(v)$, a queda temporária no preço médio causada pela execução da taxa média hv durante um intervalo de tempo. Com isso, o preço real executado em k é

$$\tilde{S}_k = S_{k-1} - h\left(\frac{n_k}{\tau}\right), \quad (1.2.3)$$

mas o efeito de $h(v)$ não aparece no próximo preço de mercado S_k .

As funções $g(v)$ em 1.2.2 e $h(v)$ em 1.2.3 podem ser escolhidas por qualquer modelo de preferencia de microestrutura de mercado, sujeitas apenas algumas condições naturais de convexidade.

Custos das trajetórias de execução

Agora será discutido os lucros/prejuízos resultantes da execução ao longo de certa trajetória. É definido o resultado da trajetória como a receita de toda execução após completado todos os negócios. Isto é a soma do produto das quantidades n_k vendidas em cada intervalo de tempo vezes o preço efetivo recebido \tilde{S}_k pela venda, calculado pela seguinte equação:

$$\sum_{k=1}^N n_k \tilde{S}_k = X S_0 + \sum_{k=1}^N (\sigma \tau^{1/2} \xi_k - \tau g(\frac{n_k}{\tau})) x_k - \sum_{k=1}^N n_k h(\frac{n_k}{\tau}). \quad (1.2.4)$$

O primeiro termo do lado direito de 1.2.4 é o valor inicial de mercado da posição, e cada termo adicional representa o ganho ou perda devido um fator específico de mercado.

O primeiro termo do tipo $\sum \sigma \tau^{1/2} \xi_k x_k$, representa o efeito total da volatilidade. O termo $-\sum \tau x_k g(n_k/\tau)$ é o impacto de preço permanente e representa uma perda no valor da posição, causado pelo impacto de preço permanente da queda de preço associada a execução da posição até aquele instante. E o termo $\sum n_k h(n_k/\tau)$ é o impacto temporário, que representa uma queda no preço de venda, somente na quantidade executada no k -ésimo intervalo.

O custo total da execução é a diferença $X S_0 - \sum n_k \tilde{S}_k$ entre o valor inicial e o resultado da execução. Isto é a medida padrão de custo de transação usado para avaliar a performance, e é o que essencialmente Perold [28] chamou de *implementation shortfall*. Neste modelo, *implementation shortfall* é uma variável aleatória, então pode-se escrever $E(x)$ como *expected shortfall* e $V(x)$ como *variance shortfall*. Dada a dinâmica de preço, estas variáveis são calculados por:

$$E(x) = \sum_{k=1}^N \tau x_k g\left(\frac{n_k}{\tau}\right) + \sum_{k=1}^N n_k h\left(\frac{n_k}{\tau}\right) \quad (1.2.5)$$

$$V(x) = \sigma^2 \sum_{k=1}^N \tau x_k^2 \quad (1.2.6)$$

A distribuição do *shortfall* é exatamente Gaussiana se ξ_k é Gaussiano, e para qualquer caso de N grande é muito próximo de uma Gaussiana.

O problema de execução ótima, é encontrar as trajetórias de execução que minimizam $E(x) + \lambda V(x)$ para um certo valor de λ , ou seja, para um certo nível de risco. Foi mostrado que cada valor de λ corresponde a uma única trajetória x tal que $E(x) + \lambda V(x)$ é mínimo [3].

Funções de impacto linear

O problema de calcular as trajetórias ótimas se torna muito mais simples se utilizarmos funções lineares para o impacto permanente e temporário, embora para a formalização do problema não foi necessária esta restrição.

Para o impacto permanente linear, utilizamos $g(v)$ como:

$$g(v) = \gamma v, \quad (1.2.7)$$

no qual a constante γ tem unidade $(\$/quantidade)/quantidade$. Desta forma, cada n unidades que são vendidas o preço é diminuído por γn , independente do tempo que gastamos para executar n unidades, então a equação 1.2.2 fica:

$$S_k = S_{k-1} + \sigma\tau^{1/2}\xi_k - \gamma(X - x_k), \quad (1.2.8)$$

A soma das partes, do termo impacto permanente em 1.2.4 fica

$$\begin{aligned} \sum_{k=1}^N \tau x_k g\left(\frac{n_k}{\tau}\right) &= \gamma \sum_{k=1}^N x_k n_k = \gamma \sum_{k=1}^N x_k (x_{k-1} - x_k) = \\ &= \frac{1}{2}\gamma \sum_{k=1}^N (x_{k-1}^2 - x_k^2 - (x_k - x_{k-1})^2) = \frac{1}{2}\gamma X^2 - \frac{1}{2}\gamma \sum_{k=1}^N n_k^2. \end{aligned} \quad (1.2.9)$$

$$(1.2.10)$$

Similarmente, para o impacto temporário temos:

$$h\left(\frac{n_k}{\tau}\right) = \epsilon \operatorname{sgn}(n_k) + \frac{\eta}{\tau} n_k, \quad (1.2.11)$$

onde sgn é uma função sinal.

A unidade de ϵ é $\$/quantidade$, e a de η é $(\$/quantidade)/(quantidade/tempo)$. A estimativa plausível para ϵ é um custo fixo, tal que seja metade do *spread* mais taxas. É muito mais difícil estimar η já que depende de aspectos internos e temporários da microestrutura do mercado. Neste termo que é esperado efeitos não lineares, e a aproximação de 1.2.11 pode ser duvidosa.

O modelo linear em 1.2.11 é normalmente chamado de custo quadrático porque o custo total envolvido na execução de n unidades em uma única unidade de tempo é

$$nh\left(\frac{n}{\tau}\right) = \epsilon|n| + \frac{\eta}{\tau}n^2 \quad (1.2.12)$$

Com os dois custos de impacto lineares em 1.2.7, 1.2.11, o valor esperado dos custos de impacto em 1.2.5 é

$$E(x) = \frac{1}{2}\gamma X^2 + \epsilon \sum_{k=1}^N |n_k| + \frac{\tilde{\eta}}{\tau} \sum_{k=1}^N n_k^2 \quad (1.2.13)$$

no qual

$$\tilde{\eta} = \eta - \frac{1}{2}\gamma\tau \quad (1.2.14)$$

Claramente, E é uma função estritamente convexa deste que $\tilde{\eta} > 0$.

1.2.2 Fronteira eficiente para execução ótima

Nesta seção, será calculado as trajetórias de execução ótima, e mostrada a relação precisa entre aversão ao risco e a definição de otimalidade. Em particular, é mostrado que para cada nível de aversão ao risco existe uma única trajetória ótima de execução.

Definição da fronteira

O agente racional sempre vai buscar minimizar a perda esperada dado um certo nível de variância da perda. Por causa disso, uma execução é eficiente ou ótima se não existe outra estratégia que tenha menor variância para o mesmo nível de perda esperada.

Então pode-se construir estratégias eficientes resolvendo o problema de otimização restrita

$$\min_{x:V(x)\leq V_*} E(x). \quad (1.2.15)$$

Isto é, para um nível máximo de variância $V_* \geq 0$, buscamos um estratégia que tenha o mínimo valor esperado de custos de transação. Como $V(x)$ é convexa, o conjunto $V(x) \leq V_*$ é convexo, e desde que $E(x)$ seja estritamente convexa, existe um único minimizador $x_*(V_*)$.

Independentemente da preferência de risco e retorno, cada outra solução de x que tenha $V(x) \leq V_*$ possui custos maiores do que $x_*(V_*)$ para o mesmo nível ou menor variância, que nunca pode ser eficiente. Por causa disso, a família de todas as possíveis estratégias eficientes são parametrizadas pela única variável V_* , representando todos possíveis níveis de variância e custos de transação. Esta família é chamada de fronteira eficiente das estratégias de execuções ótimas.

Para resolver o problema de otimização restrita em 1.2.15 é introduzido o multiplicador de Lagrange λ , resolvendo o problema irrestrito

$$\min_x (E(x) + \lambda V(x)). \quad (1.2.16)$$

Se $\lambda > 0$, então $E + \lambda V$ é estritamente convexa, e 1.2.16 possui uma única solução $x^*(\lambda)$. Ao variar λ o $x^*(\lambda)$ varia traçando uma fronteira eficiente.

O parâmetro λ possui um interpretação financeira direta. Pode-se observar de 1.2.16 que λ é uma medida de aversão ao risco, medindo o quanto penalizamos a variância relativo ao custo esperado.

Dados os valores dos parâmetros, o problema 1.2.16 pode ser resolvido por diversas técnicas numéricas dependendo da forma das funções $g(v)$ e $h(v)$. No caso especial de funções lineares, a solução pode ser escrita explicitamente dando uma ideia de estratégias de execução.

A figura 1.5, ilustra a fronteira eficiente obtida no trabalho [3] para funções de impacto de preço lineares. A região sombreada são um conjunto de variâncias e esperanças alcançáveis por algum tipo de estratégia. A curva sólida é a fronteira eficiente, a curva pontilhada são estratégias que possuem maior variância para o mesmo nível de custo esperado. O ponto B é a estratégia que minimiza o custo esperado sem considerar o variância (neutro ao risco). Os pontos A, B, C são ilustrados da na figura 1.6.

A trajetória A possui $\lambda > 0$, que seria escolhida por um agente averso ao risco que prefere vender rapidamente para reduzir a exposição ao risco de volatilidade, mesmo acarretando custos de execução maior.

A trajetória B possui $\lambda = 0$. Esta é a chamada estratégia simples, já que representa a estratégia ótima para minimizar o custo esperado sem considerar a variância. Para um ativo com arrasto zero e custos de transação linear, corresponde a redução linear da posição ao longo do tempo conhecido como TWAP (*time-weighted average price*).

A trajetória C possui $\lambda < 0$, que seria escolhida por um agente que gosta de risco. Ela atrasa a execução gerando custos esperados maiores devido a rápida execução no final, e maior variância durante período que mantém a posição.

Figura 1.5: Fronteira eficiente

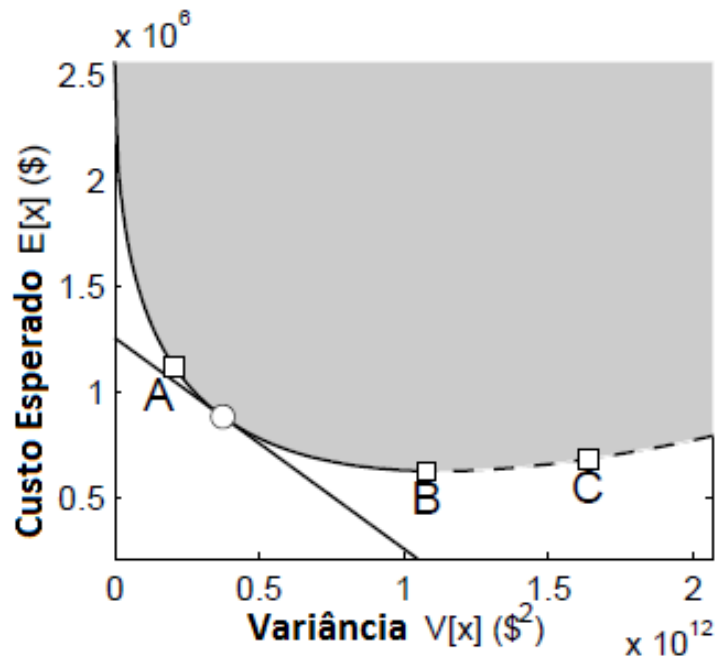
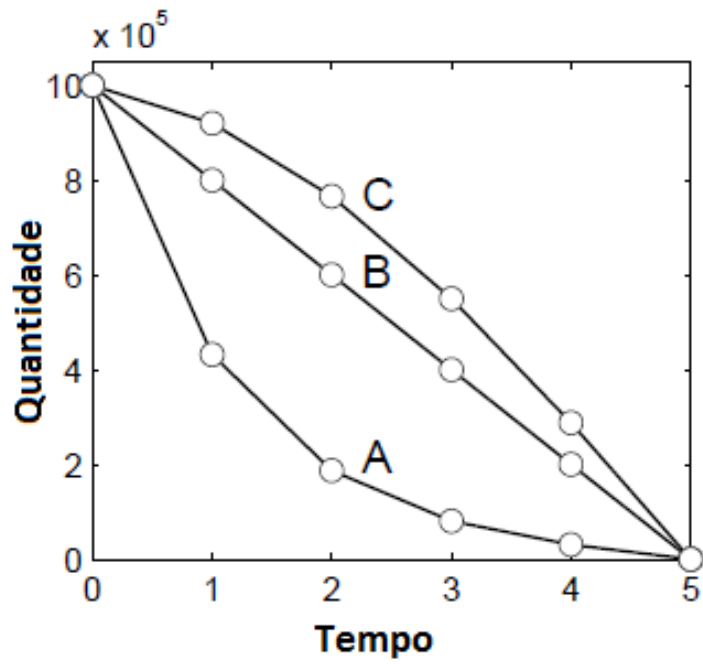


Figura 1.6: Trajetórias



Capítulo 2

Estimação mercado brasileiro

2.1 Dados

Neste trabalho foram utilizados dois tipos de dados em alta frequência: cotações consolidadas das melhores ofertas de compra/venda e os negócios intra-diários. As figuras 2.1 e 2.2 ilustram o formato dos dois tipos de dados: cotações e negócios.

Como mostrado na figura 2.1 qualquer variação de quantidade ou preço do primeiro nível do livro de ofertas resulta em uma nova amostra. Foi observado que aproximadamente há 10 vezes mais dados de cotações do que de negócios, corroborando a ideia deste trabalho de utilizar o fluxo líquido de ordens ao invés de apenas os negócios para calcular o coeficiente de impacto no preço por possuir muito mais informação da microestrutura do mercado. Os dados de negócio também possuem a informação de qual lado do livro de ofertas foi agredido sendo 1 a oferta de compra foi agredida, 2 a oferta de venda foi agredida e 3 para diretas.

Os dados utilizados são dos principais mercados brasileiros: futuro do Índice Bovespa (1^o vencimento), futuro de Dolar (1^o vencimento), futuro de Juros (DI1F17 contrato futuro mais negociado neste período) e seis ações do Ibovespa divididas em três categorias de liquidez: muito volume (PETR4, VALE5), volume intermediário (CYRE3, CSAN3) e pouco volume (GOLL4, ELET6). A base de dados é de 6 de setembro de 2013 até 21 de março de 2014, aproximadamente seis meses de dados (133 dias). Esta base foi obtida do fornecedor Bloomberg. A tabela 2.1 possui algumas estatísticas descritivas destes ativos que foram calculadas neste período. Esta tabela mostra na terceira coluna o número de total de atualizações (eventos) da primeira posição do livro de ordens e na quarta coluna o número total de negócios que ocorreram neste período na ordem de milhões (MM). A quinta coluna mostra o *spread* médio em número de ticks destes ativos. A sexta coluna possui o volume médio em bilhões (Bi) de reais negociados neste período.

2.2 Definição das variáveis

Cada observação de cotação consiste de P^b e P^s sendo os melhores preços de compra/venda e q^b e q^s as quantidades do melhor nível de compra/venda respectivamente. Enumerando por n e calculando as diferenças entre as observações consecutivas $(P_n^b, q_n^b, P_n^s, q_n^s)$ como segue:

$$e_n = q_n^b 1_{\{P_n^b \geq P_{n-1}^b\}} - q_{n-1}^b 1_{\{P_n^b \leq P_{n-1}^b\}} - q_n^s 1_{\{P_n^s \leq P_{n-1}^s\}} + q_{n-1}^s 1_{\{P_n^s \geq P_{n-1}^s\}} \quad (2.2.1)$$

A variável e_n é a contribuição dos eventos do livro de ofertas para oferta/demanda. Quando uma ordem passiva de compra chega, q^b aumenta mas P^b continua o mesmo, levando $e_n =$

Figura 2.1: Exemplo Cotações

tempo	bid qtd	bid preço	ask preço	ask qtd
2013-11-01 10:10:00	5100	20,09	20,10	12900
2013-11-01 10:10:01	4000	20,09	20,10	12900
2013-11-01 10:10:01	4000	20,09	20,10	13000
2013-11-01 10:10:01	3000	20,09	20,10	13000
2013-11-01 10:10:01	2800	20,09	20,10	13000
2013-11-01 10:10:02	1800	20,08	20,10	13000
2013-11-01 10:10:02	4600	20,08	20,10	13000
2013-11-01 10:10:02	4600	20,08	20,10	13100
2013-11-01 10:10:02	200	20,09	20,10	13100
2013-11-01 10:10:02	4600	20,08	20,10	13100
2013-11-01 10:10:02	1800	20,08	20,10	13100
2013-11-01 10:10:02	2800	20,09	20,10	13100
2013-11-01 10:10:02	1800	20,08	20,10	13100

Figura 2.2: Exemplo Negócios

tempo	preço	quantidade	lado
2013-11-01 10:10:00	20,09	1100	1
2013-11-01 10:10:01	20,09	1100	1
2013-11-01 10:10:01	20,09	1000	1
2013-11-01 10:10:01	20,09	200	1
2013-11-01 10:10:02	20,09	200	3
2013-11-01 10:10:02	20,09	1500	2
2013-11-01 10:10:03	20,09	1400	2
2013-11-01 10:10:08	20,09	1500	2
2013-11-01 10:10:08	20,09	1500	2
2013-11-01 10:10:21	20,1	9100	2
2013-11-01 10:10:21	20,1	800	2
2013-11-01 10:10:21	20,1	100	2
2013-11-01 10:10:21	20,1	500	2

Tabela 2.1: Estatísticas Descritivas - Dados 06/09/2013 - 21/03/2014 (133 dias)

Mercado	Nome	Nº cotações (MM)	Nº negócios (MM)	Méd. spread (ticks)	Méd. vol. (Bi)
Juros	DI1F17	3,07	0,63	1,08	16,8421
Dólar	1º futuro	10,35	2,62	1,19	35,8921
Índice	1º futuro	10,76	1,26	3,39	3,3770
Ações	PETR4	12,15	4,28	1,22	0,4754
Ações	VALE5	10,94	3,47	1,41	0,4963
Ações	CYRE3	3,02	1,16	1,46	0,0489
Ações	CSAN3	2,55	0,62	3,61	0,0484
Ações	GOLL4	1,55	0,56	1,81	0,0154
Ações	ELET6	2,02	0,71	1,50	0,0177

$q_n^b - q_{n-1}^b$ que é o tamanho desta nova ordem. Se q^b diminui, temos que $e_n = q_n^b - q_{n-1}^b$, representando o tamanho de uma ordem a mercado de venda ou cancelamento de um ordem de compra. Se P^b muda, então $e_n = q_n^b$ ou $e_n = -q_{n-1}^b$, representando respectivamente o tamanho da nova ordem que melhorou o preço ou o tamanho da fila anterior que foi removida. Esta mesma análise é simétrica para o lado de venda do livro de ofertas.

Neste trabalho, são usados dois intervalos de tempo T_0, \dots, T_I e $t_{0,0}, \dots, t_{I,K}$ com tamanhos de $T_i - T_{i-1}$ e $t_{k,i} - t_{k-1,i} = \Delta t$. Com cada intervalo de tempo longo $[T_{i-1}, T_i]$ nós calculamos n mudanças de preço e de *order flow imbalances* indexados por k :

$$\Delta P_{k,i} = \frac{P_{N(t_{k,i})}^b + P_{N(t_{k,i})}^s}{2\delta} - \frac{P_{N(t_{k-1,i})}^b + P_{N(t_{k-1,i})}^s}{2\delta} \quad (2.2.2)$$

$$OFI_{k,i} = \sum_{n=N(t_{k-1,i})+1}^{N(t_{k,i})} e_n, \quad (2.2.3)$$

sendo $N(t_{k-1,i})+1$ e $N(t_{k,i})$ são os índices do primeiro e do último evento do livro de ofertas no intervalo $[t_{k-1,i}, t_{k,i}]$. O δ é o tamanho de um tick (mínimo incremento de preço).

Também, definimos o *trade imbalance* durante o intervalo $[t_{k-1,i}, t_{k,i}]$ como a diferença entre o volume dos compradores e vendedores durante este intervalo de tempo, e também definimos volume negociado em cada intervalo:

$$TI_{k,i} = \sum_{n=N(t_{k-1,i})+1}^{N(t_{k,i})} b_n - s_n \quad (2.2.4)$$

$$VOL_{k,i} = \sum_{n=N(t_{k-1,i})+1}^{N(t_{k,i})} b_n + s_n \quad (2.2.5)$$

sendo b_n, s_n as quantidades iniciadas pelos compradores e vendedores que ocorreram na n -ésima cotação.

2.3 Resultados empíricos

Esta seção apresenta os principais resultados obtidos neste trabalho, são mostrados os resultados das estimações do modelo nos principais mercados brasileiros, assim como tabelas e gráficos ilustrando a aderência do modelo.

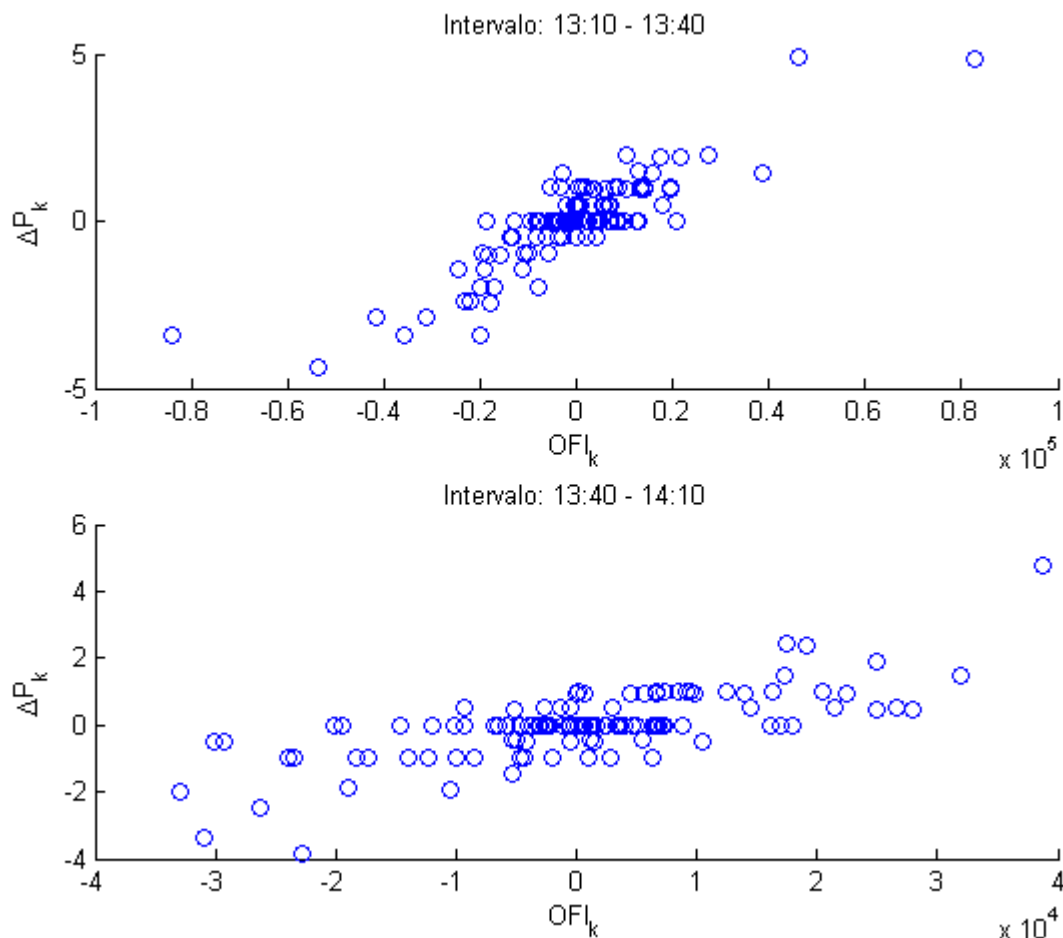
Estimação pela variável OFI

O modelo foi aplicado nos principais ativos negociados eletronicamente no Brasil, e foi estimado utilizando a seguinte regressão linear simples:

$$\Delta P_{k,i} = \hat{\alpha}_i + \hat{\beta}_i OFI_{k,i} + \hat{\epsilon}_{k,i} \quad (2.3.1)$$

em amostras de 30 minutos indexadas por i , cada intervalo de 30 minutos possui 120 dados de $\Delta P_{k,i}$ e $OFI_{k,i}$ calculados a cada 15 segundos utilizando as formulas apresentadas na seção 2.2. A figura 2.3 mostra os gráficos de dispersão do $\Delta P_{k,i}$ contra $OFI_{k,i}$ para duas amostras de 30 minutos de um dia sendo uma no intervalo de 13:10 a 13:40 e outra no período de 13:40 a 14:10, também pode-se notar a diferença de inclinação dos dois gráficos ilustrando que o coeficiente de impacto no preço (inclinação da reta) é diferente nestes dois intervalos.

Figura 2.3: Gráfico de dispersão $OFI_k \times \Delta P_k$ da VALE5 no dia 17/01/2014



Para realizar as estimações, foi necessário criar um programa para poder amostrar a base de dados nestas duas frequências, pois os dados de cotações e negócios estão na sequência de tempo das atualizações do mercado (descontínuos) e não igualmente espaçados no tempo (amostrados). Após a amostragem, foi calculado $\Delta P_{k,i}$ e $OFI_{k,i}$ para cada intervalo curto de 15 segundos, e em cada intervalo longo de 30 minutos foram realizadas as regressões utilizando 120 subamostras de 15 segundos do intervalo de 30 minutos. Estas estimações foram estendidas para todos os dias e todos os ativos. A tabela seguinte, mostra a média destes resultados para todos intervalos de 30 minutos das estimações realizadas no período de 6 de setembro de 2013 até 21 de março de 2014 (133 dias).

Ativo	R^2	$\hat{\alpha}_i$	$\hat{\beta}_i$	$\hat{D} = 1/2\hat{\beta}_i$	$t(\hat{\alpha}_i = 0)$	$t(\hat{\beta}_i = 0)$
IBOVESPA ¹	72%	0,0010	0,02940	17	0,01	19,68
PETR4	53%	-0,0054	0,00002	24408	-0,11	12,12
VALE5	56%	0,0027	0,00005	9828	0,04	13,04
CYRE3	52%	0,0026	0,00017	2931	0,09	12,08
CSAN3	48%	0,0251	0,00066	751	0,26	10,90
GOLL4	43%	-0,0011	0,00018	2727	0,02	10,10
ELET6	40%	0,0044	0,00014	3503	0,14	09,29
DOLAR ¹	55%	-0,0008	0,00240	209	-0,03	12,74
DI1F17	33%	0,0006	0,00038	1300	0,02	07,76

Como pode ser visto na tabela anterior, os resultados obtidos mostraram-se bastante aderentes com uma média de $R^2 = 50\%$, o t -statistic do parâmetro β (coeficiente de impacto no preço) bastante significativo e do parâmetro α praticamente insignificante, o que já era esperado pelo modelo do livro de ofertas estilizado mostrado na seção 1.1.3. Também, pode-se observar que o ativo que obteve o menor R^2 foi o juros futuro com um valor de $R^2 = 33\%$. Apesar deste mercado ser bastante líquido, acredita-se que a variável OFI_k não explicou bem as variações no preço ΔP_k , pois este ativo possui um tick (mínimo incremento de preço) muito caro, então há muita atividade de ordens alterando a variável OFI_k , mas sem uma alteração no preço.

Também, vale ressaltar que esta tabela pode ser utilizada como uma referência de liquidez para execução de ordens nestes mercados. Por exemplo, para o índice Bovespa futuro foi encontrado uma média de profundidade do livro de $D = 17$, ou seja, para um agente negociar 170 lotes deste contrato ele vai impactar 5 ticks ($\Delta P = OFI/2D$) no preço deste ativo. Então o parâmetro β , coeficiente de impacto no preço, pode ser visto como o impacto temporário utilizado no modelo de execução ótima de Almgren-Chris [3].

Estimação pela variável TI

A estimação do impacto de preço também foi realizada utilizando a variável *trade imbalance* (TI). Para isto foram calculadas as variáveis $\Delta P_k, TI_k$ para cada intervalo de 1 minuto tendo em vista que os negócios acontecem em uma frequência menor do que as cotações, então utilizar a amostragem de 15 segundos como utilizada no modelo do OFI poderia resultar em muitas amostras sem volume negociado. Para realizar a estimação foi utilizada a seguinte regressão e dentro de cada dia a regressão foi calculada em intervalos independentes de 100 minutos. A média dos resultados obtidos segue na tabela abaixo.

$$\Delta P_{k,i} = \hat{\alpha}_i^T + \hat{\beta}_i^T TI_{k,i} + \hat{\epsilon}_{k,i}, \quad (2.3.2)$$

¹Contrato Futuro

Ativo	R^2	$\hat{\alpha}_i^T$	$\hat{\beta}_i^T$	\hat{D}_i^T	$t(\hat{\alpha}_i^T = 0)$	$t(\hat{\beta}_i^T = 0)$
IBOVESPA ¹	14%	0,13	0,021500	46	0,27	3,51
PETR4	10%	0,04	0,000012	83218	0,28	2,87
VALE5	7,7%	0,03	0,000022	44510	0,13	2,36
CYRE3	6,8%	0,00	0,000071	13997	-0,02	1,74
CSAN3	7,2%	0,03	0,000276	3622	0,08	1,63
GOLL4	13%	0,03	0,000125	7991	0,23	3,12
ELET6	8,2%	0,02	0,000090	11057	0,26	2,14
DOLAR ¹	7,6%	0,10	0,000770	1298	0,45	2,99
JUROS ¹	7,7%	0,00	0,000182	5469	-0,15	2,98

Na tabela anterior, é possível observar que o modelo utilizando a variável TI mostrou-se bem menos aderente que o modelo do OFI com uma média de $R^2 = 9,1\%$ contra a média de $R^2 = 50\%$ do modelo anterior. Tal fato ilustra que o modelo utilizando os eventos do livro de ordens possui muito mais informação da microestrutura sendo um melhor estimador para o impacto de preço.

Estimação diária modelo de Kyle

O modelo de Kyle foi utilizado para estimação do impacto de mercado, para isto foram calculadas as variáveis ΔP_k , TI_k como definidas na seção 1.1.3 para dados diários, isto é, para cada dia foi calculada a variação de preço que ocorreu neste dia (ΔP_k) e a diferença das quantidades das ordens a mercado (TI_k) deste dia, com isto foi estimada a seguinte regressão para todos os ativos. Os resultados obtidos encontram-se na tabela seguinte.

$$\Delta P_k = \hat{\alpha}^\lambda + \hat{\beta}^\lambda TI_k + \hat{\epsilon}_k, \quad (2.3.3)$$

Ativo	R^2	$\hat{\alpha}^\lambda$	$\hat{\beta}^\lambda$	\hat{D}^λ	$t(\hat{\alpha}^\lambda = 0)$	$t(\hat{\beta}^\lambda = 0)$
IBOVESPA ¹	13%	13,51	0,012200	41	2,41	4,41
PETR4	17%	2,08	0,000004	124180	0,77	5,14
VALE5	25%	3,25	0,000011	43458	0,89	6,54
CYRE3	8,1%	0,27	0,000009	52827	0,12	3,40
CSAN3	7,0%	8,55	0,000054	9264	2,83	3,14
GOLL4	22%	5,72	0,000042	11781	2,23	6,07
ELET6	11%	1,74	0,000019	25491	0,97	4,17
DOLAR ¹	15%	18,06	0,000836	597	3,48	4,73
JUROS ¹	10%	1,19	0,000111	4481	1,28	3,78

Pode-se observar que este modelo não é tão aderente como o modelo do OFI apresentado anteriormente, mas também não se pode desprezar a sua significância estatística. Tendo em vista que a estimação foi feita utilizando dados diários, então este modelo pode ser utilizado para calcular o impacto de preço permanente do modelo de execução de Almgren-Chriss [3].

Capítulo 3

Análise temporal

Ao realizar as estimações foi verificada uma variação no coeficiente de impacto no preço β durante os intervalos do dia. Sendo que certos intervalos possui maior impacto do que outros, devido alguma particularidade do mercado como horário de almoço, abertura e fechamento . A tabela 3.1, tem o intuito de ilustrar estas variações intra-diárias. Ela mostra a média dos parâmetros durante os intervalos do dia para o índice Bovespa futuro, no apêndice C estão as tabelas equivalentes para os outros ativos da base de dados.

Tabela 3.1: Média dos parâmetros do modelo OFI: índice Bovespa futuro

Intervalo	$\hat{\beta}$	\hat{D}	R^2
9:30	0.036335	13.8	53%
10:00	0.036911	13.5	73%
10:30	0.033139	15.1	75%
11:00	0.031018	16.1	78%
11:30	0.029987	16.7	77%
12:00	0.030728	16.3	78%
12:30	0.029230	17.1	75%
13:00	0.030336	16.5	75%
13:30	0.032271	15.5	79%
14:00	0.030723	16.3	77%
14:30	0.029900	16.7	78%
15:00	0.028123	17.8	74%
15:30	0.026643	18.8	72%
16:00	0.025418	19.7	70%
16:30	0.024978	20.0	69%
17:00	0.025391	19.7	69%
17:30	0.023848	21.0	65%
18:00	0.024841	20.1	61%
Média Total	0.029434	17.3	72%

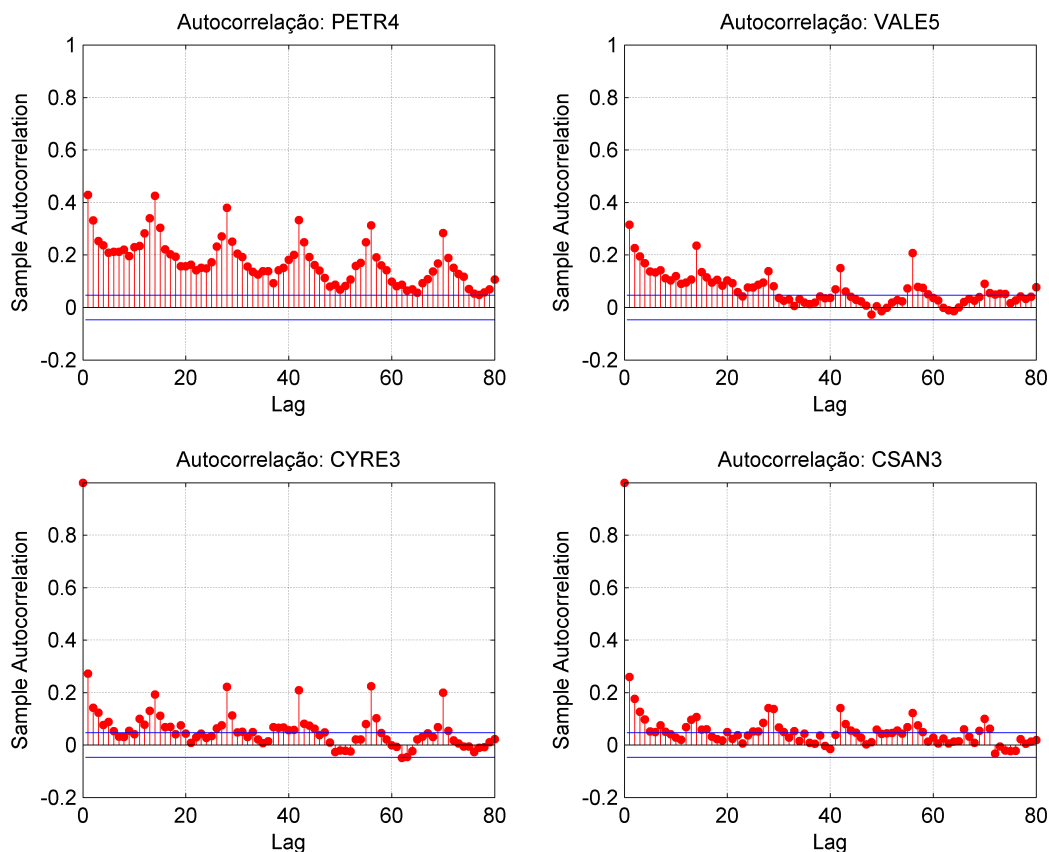
Na tabela 3.1 pode-se verificar que na abertura do mercado de índice futuro entre as 9:00 e 9:30 horas da manhã, o parâmetro $\hat{\beta}$ possui uma média em torno de 0,036 que é bem maior do que a média do intervalo das 17:00 as 17:30 dada por 0,024, sendo que este último intervalo engloba o ajuste do índice futuro e o fechamento do mercado de ações, que naturalmente é um período de maior liquidez por possuir uma maior quantidade de agentes negociando, e a abertura do índice futuro não possui tanta liquidez devido o contrato futuro abrir antes que o

mercado de ações. Com isso pode-se observar uma grande variação no coeficiente de impacto no preço durante os intervalos do dia, o que nos sugeriu fazer uma análise mais detalhada desta variável.

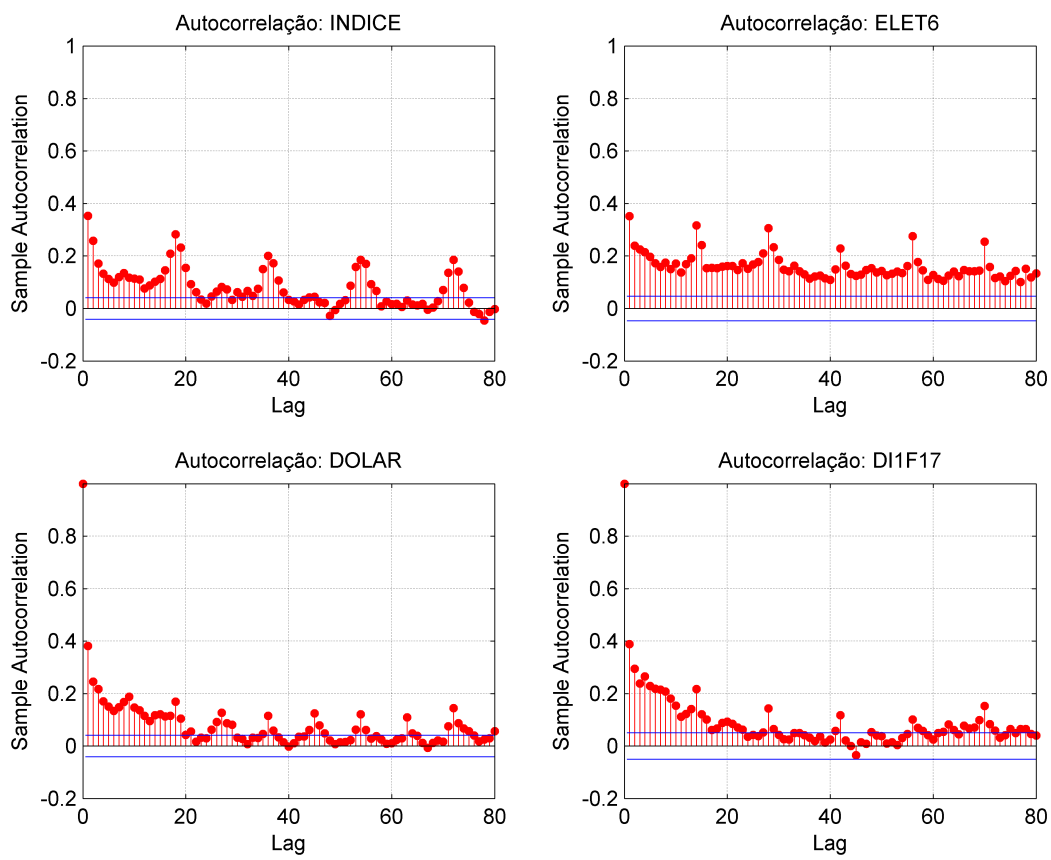
3.1 Autocorrelação no impacto de preço

Esta seção irá apresentar uma análise mais detalhada que foi realizada na série do parâmetro β (coeficiente de impacto no preço). Analisando as estimações do parâmetro β , foi possível verificar uma estrutura de autocorrelação na série destes dados. Tal fato estimulou buscar um modelo para incorporar esta ideia na modelagem do impacto no preço. As figuras 3.1 e 3.2 mostram as funções de autocorrelação dos ativos estudados. Como pode ser observado nas figuras, as funções de autocorrelação possuem uma forte autocorrelação com uma certa sazonalidade o que já era esperado coincidindo com padrões intra diários de volume, spread e profundidade do mercado [1],[6], [24] e [25]. Baseado nesta estrutura de autocorrelação foi calibrado um modelo SARIMA (sazonal auto regressivo integrado média móvel) na série do parâmetro β do ativo PETR4 com intuito de ilustrar como modelar a estrutura temporal deste parâmetro.

Figura 3.1: Funções autocorrelação do β : PETR4, VALE5, CYRE3, CSAN3

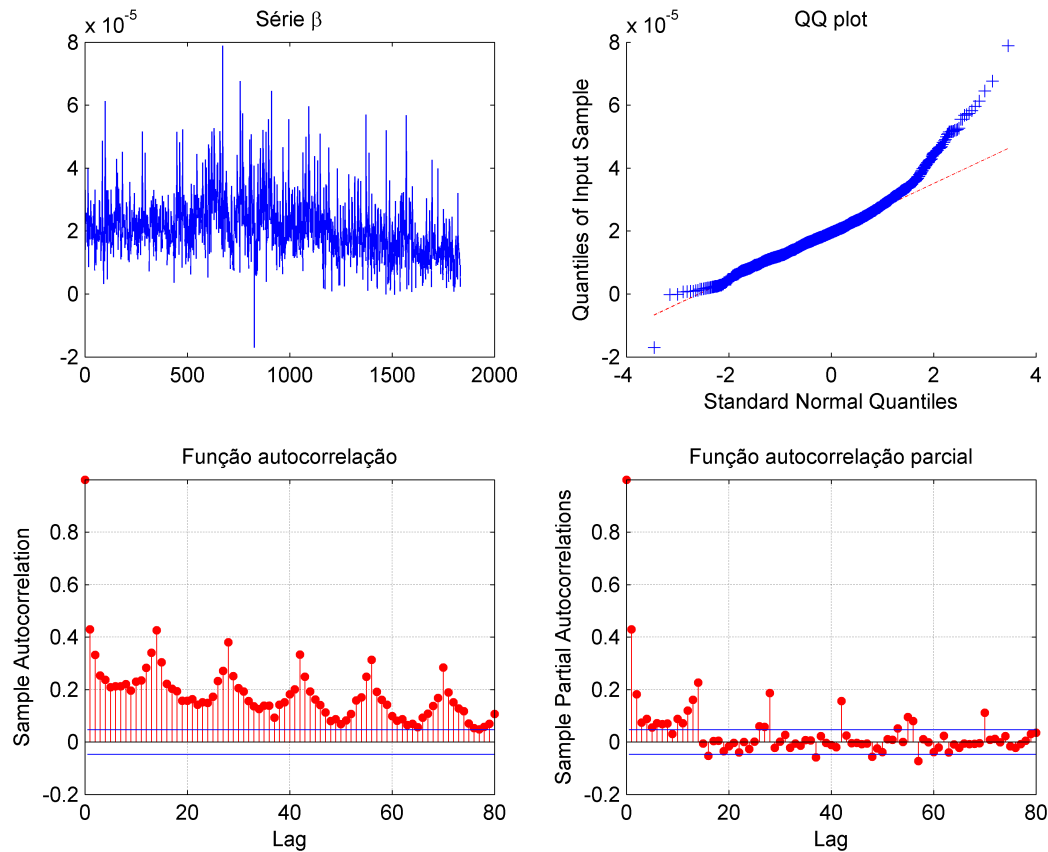


Primeiramente para decidir o modelo utilizado na calibragem foi feita uma análise visual na série deste parâmetro. Como pode ser observado na figura 3.3, de 14 em 14 lags há uma forte autocorrelação sugerindo um modelo sazonal de lag 14, que é condizente com o mercado, pois como cada parâmetro β foi estimado utilizando 30 minutos de dados a lag 14 significa 7 horas

Figura 3.2: Funções autocorrelação do β : INDICE, ELET6, DOLAR, DI1F17

de diferença que é exatamente 1 dia do mercado de ações que abre as 10:00 horas e as 16:55 começa o leilão de fechamento. Além disso, também verificou-se que um modelo ARMA(1,1) é aderente para capturar a estrutura de autocorrelação de lags próximas.

Figura 3.3: Análise série β : PETR4



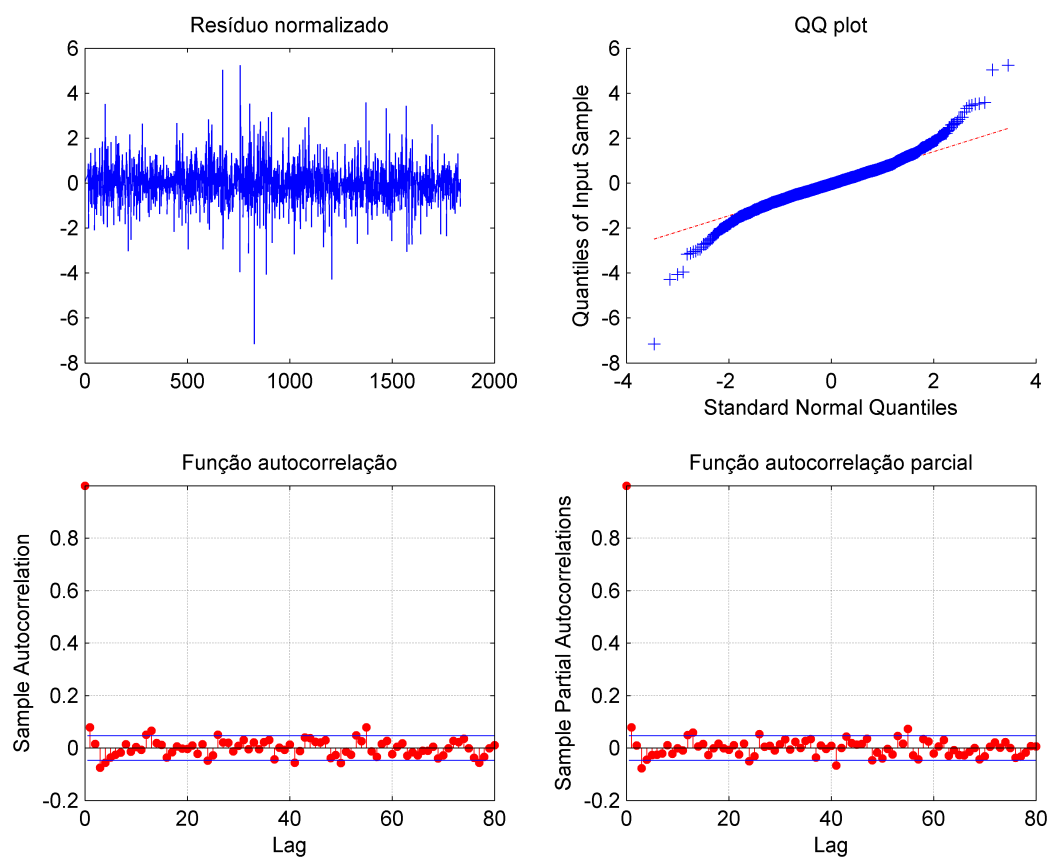
Os resultados obtidos da calibragem estão na tabela 3.2, como pode-se observar todos os parâmetros possuem um t -statistic bastante alta, mostrando a significância estatística dos parâmetros do modelo utilizado.

Tabela 3.2: ARIMA(1,0,1) sazonalmente integrado com sazonalidade MA(14): PETR4

Parâmetro	Valor	Erro Padrão	T-statistic
Constante	0	fixo	fixo
AR(1)	0,94	0,015	61,84
MA(1)	-0,79	0,025	-31,16
SMA(14)	-0,93	0,008	-112,26
Variância	6,4e-11	1,7e-08	0,003

Após calibrar o modelo SARIMA, também foi analisada a série dos resíduos. A figura 3.4 mostra os gráficos do resíduo normalizado, QQ plot, funções autocorrelação e autocorrelação parcial do resíduo. Como pode-se observar a calibragem do modelo SARIMA retirou a estrutura de autocorrelação dos dados mostrando que o modelo escolhido é aderente.

Figura 3.4: Análise do resíduo após a calibragem: PETR4



3.2 Aplicações

Os resultados obtidos neste trabalho sugerem que assumir o impacto de preço constante durante o dia pode afetar negativamente a performance de estratégias de execução ótima. Uma possível alteração neste algoritmo de execução ótima, é utilizar-se de um estratégia dinâmica que a função de impacto temporário varia durante o dia, com isso executando mais aonde impacto esperado é menor e executando menos onde o impacto esperado é maior. Com isso a distribuição de quantidade a ser executada por janela vai variar dependendo do impacto esperado para próxima janela. O seguinte modelo apresentado em [11] pode ser usado para incorporar esta dinâmica temporal do impacto temporária em uma estratégias de execução ótima.

Considerando o problema de liquidação da posição de um ativo de $X \in R_+$ quantidades até um dado tempo $T \in R_+$. O intervalo $[0, T]$ é dividido em N subintervalos de tamanho Δt e tem que decidir em cada tempo $t_{n-1} = (n-1)\Delta t$ qual a quantidade y_n para vender no intervalo $(t_{n-1}, t_n]$. É assumido que as ordens y_n são executadas por um programa tático que submete uma combinação de ordens limitadas e ordens a mercado para obter o preço ótimo, e isto resulta em um preço de execução:

$$\tilde{S}_n = S_n - \eta_n y_n, \quad (3.2.1)$$

onde S_n segue a dinâmica

$$S_n = S_{n-1} + \sigma_n \xi_n - c y_n. \quad (3.2.2)$$

O S_n pode ser visto como uma combinação convexa dos preços de compra e venda: $S_n = \lambda S_n^b + (1-\lambda)S_n^a$. Para $\lambda = 1/2$, S_n é o preço do meio. Mas dependendo do ativo negociado ele pode ser próximo da compra ou venda. (ξ_n) é uma sequência de variáveis aleatórias normais independentes com média zero e variância Δt . (σ_n) é a volatilidade estocástica. $c \in R_+$ é uma constante descrevendo o impacto permanente e (η_n) é o processo estocástico da liquidez modelando o impacto de preço temporário.

Podemos supor que S_n, σ_n, η_n são observáveis e a filtração que eles geram denotada por \mathcal{F}_n . As variáveis S_n e \tilde{S}_n podem ser observados diretamente. Se $y_n > 0$, η_n pode ser deduzido de 3.2.1. Caso contrario, ele tem que ser inferido do *spread*. σ_n tem que ser estimado da 3.2.2 ou por um procedimento estatístico separado usando observações recentes. O procedimento de vender as quantidades de um ativo é dado por $\sum_{n=1}^N y_n \tilde{S}_n$. A estratégia de liquidação é descrita em termos das quantidades remanescentes $x_n = X - \sum_{i=1}^n y_i$ e é chamada de admissível se as seguintes condições são aceitas:

$$x_0 = X, \quad x_{n-1} \geq x_n, \quad x_N = 0, \quad (x_n) \text{ é predizível com respeito a } (\mathcal{F}_n). \quad (3.2.3)$$

Isto significa que y_n deve ser não-negativo e \mathcal{F}_{n-1} mensurável tal que $\sum_{n=1}^N y_n = X$. Uma estratégia admissível é completamente especificada por x_n para $1 \leq n \leq N-1$. \mathcal{A} é conjunto das estratégia admissíveis.

O custo de execução a é diferença entre o valor inicial e o financeiro executado: $C(x) := X S_0 - \sum_{n=1}^N y_n \tilde{S}_n$, já que $x_N = 0$, ele pode ser escrito por:

$$C(x) = \frac{cX^2}{2} + \sum_{n=1}^N (x_{n-1} - x_n)^2 \left(\eta_n - \frac{c}{2} \right) - x_n \sigma_n \xi_n. \quad (3.2.4)$$

Para encontrar a estratégia de execução ótima basta minimizar o valor esperado de $C(x)$. Nestes dois casos pode-se eliminar a constante $cX^2/2$ de $C(x)$ e trabalhar com a formula quadrática:

$$Q(x) = \sum_{n=1}^N (x_{n-1} - x_n)^2 \left(\eta_n - \frac{c}{2} \right) - x_n \sigma_n \xi_n \quad (3.2.5)$$

sem alterar a solução ótima.

Se o objetivo é minimizar o valor esperado de $Q(x)$, a dinâmica de (S_n, σ_n, η_n) pode ser escolhida de forma mais geral do que o critério não linear. Assumindo que ξ_n é independente de $(\mathcal{F}_{n-1}, \sigma_n, \eta_n)$, (σ_n, η_n) e toma valores no subconjunto finito $V \subseteq \mathbb{R}_+^2$, e $\tilde{\eta}_n := \eta_n - c/2 > 0$. Supondo que condicionado a \mathcal{F}_{n-1} , a distribuição (σ_n, η_n) apenas depende de $(\sigma_{n-1}, \eta_{n-1}, \varphi_{n-1})$, onde φ_{n-1} é definida como $\varphi_{n-1} := \varphi(\sigma_{n-1}, \xi_{n-1})$ para um inteiro positivo k e uma função mensurável $\varphi : \mathbb{R} \rightarrow \{1, \dots, k\}$ (por exemplo: $\varphi(s) = 1$ para $s < 0$ e $\varphi(s) = 2$ para $s \geq 0$). Isto segue que $(\sigma_n, \eta_n, \varphi_n)$ é uma cadeia de Markov com espaço finito $V^k := V \times \{1, \dots, k\}$ e probabilidades de transição dependentes do tempo

$$p_{n-1}^{vw} := P[(\sigma_n, \eta_n, \varphi_n) = w | (\sigma_{n-1}, \eta_{n-1}, \varphi_{n-1}) = v], \quad v, w \in V^k \quad (3.2.6)$$

Isto cobre o modelo que (σ_n, η_n) pode ser dependente das inovações de preço passadas. Também, pode-se assumir que (σ_n, η_n) é uma cadeia de Markov com espaço de estado V que é independente das inovações (ξ_n) . As suas probabilidades de transição são dadas por

$$p_{n-1}^{vw} := P[(\sigma_n, \eta_n) = w | (\sigma_{n-1}, \eta_{n-1}) = v], \quad v, w \in V \quad (3.2.7)$$

Caso neutro ao risco

No caso neutro ao risco o objetivo é encontrar a estratégia admissível x que minimiza o valor esperado de $Q(x)$. Nesta seção $(\sigma_n, \eta_n, \varphi_n)$ é assumido ser uma cadeia de Markov com espaço finito $V^k := V \times \{1, \dots, k\} \subseteq \mathbb{R}_+^2 \times \mathbb{N}$ e probabilidades de transição

$$p_{n-1}^{vw} := P[(\sigma_n, \eta_n, \varphi_n) = w | (\sigma_{n-1}, \eta_{n-1}, \varphi_{n-1}) = v], \quad v, w \in V^k. \quad (3.2.8)$$

Devido a forma especial de $Q(x)$, é suficiente observar para uma estratégia ótima o processo $(\sigma_n, \eta_n, \varphi_n)$, $n \geq 0$. Por causa disso, o objetivo é minimizar $E_0^v[Q(x)]$, que para todos os estados possíveis v da cadeia de Markov $(\sigma_n, \eta_n, \varphi_n)$, E_n^v denota a esperança condicional $E[\cdot | (\sigma_n, \eta_n, \varphi_n) = v]$. O teorema seguinte dá o valor ótimo e a estratégia ótima para este problema.

Teorema 1.

$$\min_{x \in \mathcal{A}} E_0^v[Q(x)] = X^2 a_0^v, \quad (3.2.9)$$

a única solução ótima da estratégia é dada por

$$x_n^* |_{x_{n-1}^*, (\sigma_{n-1}, \eta_{n-1}, \varphi_{n-1}) = v} = x_{n-1}^* \frac{E_{n-1}^v[\tilde{\eta}_n]}{E_{n-1}^v[\tilde{\eta}_n] + \sum_{w \in V^k} p_{n-1}^{vw} a_n^{vw}}, \quad n = 1, \dots, N-1 \quad (3.2.10)$$

onde o coeficiente a_n^v satisfaz a recursão de frente para trás:

$$a_{N-1}^v = E_{N-1}^v[\tilde{\eta}_N], \quad a_{N-1}^v = \frac{E_{n-1}^v[\tilde{\eta}_n] \sum_{w \in V^k} p_{n-1}^{vw} a_n^{vw}}{E_{n-1}^v[\tilde{\eta}_n] + \sum_{w \in V^k} p_{n-1}^{vw} a_n^{vw}}, \quad n \leq N-1. \quad (3.2.11)$$

O sistema de recursão 3.2.10-3.2.11 define a estratégia de execução ótima em função da liquidez disponível no mercado. Ele diminui a posição em cada passo por uma fração que depende da liquidez esperada no próximo período e a soma ponderada dos termos da liquidez futura.

Para o caso particular da liquidez constante $\eta_n \equiv \eta$, o teorema 1 resulta

$$a_n^v = \frac{\eta - c/2}{N - n}, \quad x_n^* = x_{n-1}^* \frac{N - n}{N - n + 1}, \quad (3.2.12)$$

o que leva a uma estratégia determinística de velocidade constante

$$x_n^* = X \frac{N - n}{N} \tag{3.2.13}$$

e que é consistente com os resultados de Berstimas e Lo em [8].

Capítulo 4

Conclusão

Neste trabalho, foi realizado um estudo sobre o impacto de preço e utilizou-se modelos que relacionam a variação do preço com a variação da demanda e da oferta no livro das ordens para fazer estimações sobre o impacto de preço nos principais ativos do mercado brasileiro. As estimações se mostraram bastante aderentes no mercado brasileiro com uma média de $R^2 = 50\%$. Além disso, foi feita uma análise econométrica da série dos parâmetros de impacto de preço estimados, foi observado uma forte autocorrelação nestes dados indicando sazonalidade. Por causa disso, foi calibrado um modelo econométrico SARIMA que mostrou-se bastante aderente e que pode ser utilizado para realizar melhores estimações do impacto de preço nos próximos intervalos. Os resultados obtidos neste trabalho sugerem que assumir o impacto de preço constante durante o dia pode degradar a performance de estratégias de execução. Por isso, foi apresentada uma estratégia de execução ótima que leve em consideração a estrutura temporal do impacto de preço (liquidez estocástica).

Apêndice A

Provas

Prova do Teorema 1

Para provar o teorema é utilizada indução reversa. Já que $E_0^v[Q(x)] = E_0^v[R_0(x)]$ para

$$R_n(x) = \sum_{i=n+1}^N (x_{i-1} - x_i)^2 \tilde{\eta}_i,$$

e denotamos $\mathcal{A}_n(z)$ o conjunto (\mathcal{F}_n) -predizível das estratégias $(x_i)_{i=n}^N$ satisfazendo $x_n = z$, $x_{i-1} \geq x_i$, $x_N = 0$, e definimos:

$$J_n^v(z) := \min_{x \in \mathcal{A}_n(z)} E_n^v[R_n(x)].$$

Então

$$J_{N-1}^v(x_{N-1}) = x_{N-1}^2 E_{N-1}^v[\tilde{\eta}_N] = x_{N-1}^2 a_{N-1}^v.$$

e por indução,

$$\begin{aligned} J_{n-1}^v(x_{n-1}) &= \min_{x \in \mathcal{A}_{n-1}(x_{n-1})} E_{n-1}^v[R_{n-1}(x)] \\ &= \min_{x \in \mathcal{A}_{n-1}(x_{n-1})} (x_{n-1} - x_n)^2 E_{n-1}^v[\tilde{\eta}_n] + E_{n-1}^v[R_n(x_n)] \\ &= \min_{x_n} (x_{n-1} - x_n)^2 E_{n-1}^v[\tilde{\eta}_n] + x_n^2 \sum_{w \in V^k} p_{n-1}^{vw} a_n^w, n \leq N-1. \end{aligned}$$

Isto segue que a única estratégia ótima é dada por

$$x_n^* = x_{n-1}^* \frac{E_{n-1}^v[\tilde{\eta}_n]}{E_{n-1}^v[\tilde{\eta}_n] + \sum_{w \in V^k} p_{n-1}^{vw} a_n^w},$$

e $J_{n-1}^v(x_{n-1}^*)$ que fica

$$(x_{n-1}^*)^2 \frac{E_{n-1}^v[\tilde{\eta}_n] \sum_{w \in V^k} p_{n-1}^{vw} a_n^w}{E_{n-1}^v[\tilde{\eta}_n] + \sum_{w \in V^k} p_{n-1}^{vw} a_n^w} = (x_{n-1}^*)^2 a_{n-1}^v.$$

Em particular, $J_0^v(X) = X^2 a_0^v$.

Apêndice B

Códigos

Código para a estimação do coeficiente de impacto de preço utilizando a variável *order flow imbalance* (OFI) dado pela equação 1.1.10:

```
%----- PARAMETROS -----
timeOpenMkt = [09,01,00];
timeCloseMkt = [17,59,00];
janelaLonga = 30*60; %segundos
janelaCurta = 15; %segundos
tol = 1e-4;

%----- CRIAR O EN -----
nbEvents = size(Z,1);
en = zeros(nbEvents,1);
qb = Z(:,2); %QTY TOP_BID
pb = Z(:,3); %PX TOP_BID
pa = Z(:,4); %PX TOP_ASK
qa = Z(:,5); %QTY TOP_ASK
en(2:end) = qb(2:end).*(MaiorOuIgual(pb(2:end),pb(1:end-1),tol)) ...
    -qb(1:end-1).*(MenorOuIgual(pb(2:end),pb(1:end-1),tol)) ...
    -qa(2:end).*(MenorOuIgual(pa(2:end),pa(1:end-1),tol)) ...
    +qa(1:end-1).*(MaiorOuIgual(pa(2:end),pa(1:end-1),tol)));

%----- OFIk, DPk, Tk -----
nbDias = size(IF_Z,1); %IF_Z: matriz com indice de inicio e fim de cada dia
begD = datenum(2013, 01, 01, timeOpenMkt(1), timeOpenMkt(2), timeOpenMkt(3));
endD = datenum(2013, 01, 01, timeCloseMkt(1), timeCloseMkt(2), timeCloseMkt(3));
nk = length(begD+janelaCurta/3600/24:janelaCurta/3600/24:endD+janelaCurta/3600/24);
%mapK = repmat(2, nk*nbDias);
OFI = zeros(nk*nbDias,1);
DP = zeros(nk*nbDias,1);
Tk = zeros(nk*nbDias,1);
Bk = zeros(nk*nbDias,1);
Mk = zeros(nk*nbDias,1);

%----- DIVIDIR O DIA EM JANELAS CURTAS -----
for d = 1:nbDias
    iniZ = IF_Z(d,1);
    fimZ = IF_Z(d,2);

    %DIVIDIR O DIA EM JANELAS CURTAS
    dt = datevec(Z(iniZ,1));
    year = dt(1); month = dt(2); day = dt(3);
```

```

begD = datenum(year, month, day, timeOpenMkt(1), timeOpenMkt(2), timeOpenMkt(3));
endD = datenum(year, month, day, timeCloseMkt(1), timeCloseMkt(2), timeCloseMkt(3));
vecK = begD+janelaCurta/3600/24:endD+janelaCurta/3600/24;

Tk((d-1)*nk + 1: d*nk) = vecK;

lastK = iniZ;
for k = 1:nk
    if k == 1
        indK = (lastK-1) + find(Z(lastK:fimZ,1)<vecK(k));
    else
        indK = (lastK-1) + find(and(Z(lastK:fimZ,1)>=vecK(k-1),Z(lastK:fimZ,1)<vecK(k)));
    end
    if isempty(indK)
        continue;
    end
    %mapK((d-1)*nk + k,1) = indK(1);
    %mapK((d-1)*nk + k,2) = indK(end);
    OFI((d-1)*nk + k) = sum(en(indK));
    pbk0 = Z(lastK,3);
    pak0 = Z(lastK,4);
    pbk1 = Z(indK(end), 3);
    pak1 = Z(indK(end), 4);
    DP((d-1)*nk + k) = ((pbk1+pak1)-(pbk0+pak0))/(2.0*ticker);
    if OFI((d-1)*nk + k) ~= 0
        Bk((d-1)*nk + k) = DP((d-1)*nk + k)/OFI((d-1)*nk + k);
    end
    lastK = indK(end)+1;
end
end

%PARAMETROS - JANELA LONGA -----
begD = datenum(2013, 01, 01, timeOpenMkt(1), timeOpenMkt(2), timeOpenMkt(3));
endD = datenum(2013, 01, 01, timeCloseMkt(1), timeCloseMkt(2), timeCloseMkt(3));
ni = length(begD+janelaLonga/3600/24:janelaLonga/3600/24:endD+janelaLonga/3600/24);

R2 = zeros(nbDias, ni);
alphas = zeros(nbDias, ni);
betas = zeros(nbDias, ni);
t_alphas = zeros(nbDias, ni);
t_betas = zeros(nbDias, ni);

lastI = Tk(1);
%-----DIVIDIR O DIA EM JANELAS LONGAS-----
for d = 1:nbDias
    iniZ = IF_Z(d,1);
    fimZ = IF_Z(d,2);

    %DIVIDIR O DIA EM JANELAS LONGAS
    dt = datevec(Z(iniZ,1));
    year = dt(1); month = dt(2); day = dt(3);
    begD = datenum(year, month, day, timeOpenMkt(1), timeOpenMkt(2), timeOpenMkt(3));
    endD = datenum(year, month, day, timeCloseMkt(1), timeCloseMkt(2), timeCloseMkt(3));
    vecI = begD+janelaLonga/3600/24:janelaLonga/3600/24:endD+janelaLonga/3600/24;

    for i = 1:ni
        indI = find(and(Tk >= lastI,Tk < vecI(i)));
        if isempty(indI)

```

```

        continue;
    end
    if i == 1
        indD(1) = indI(1);
    end
    lastI = Tk(min(indI(end)+1, length(Tk)));

    mdl = fitlm(OFI(indI), DP(indI));
    alphas(d, i) = mdl.Coefficients.Estimate(1);
    betas(d, i) = mdl.Coefficients.Estimate(2);
    t_alphas(d, i) = mdl.Coefficients.tStat(1);
    t_betas(d, i) = mdl.Coefficients.tStat(2);
    R2(d,i) = mdl.Rsquared.Ordinary;
end
end

%----- OUTPUT -----
[lins,~] = find(isnan(R2));
linCut = unique(lins);
R2(linCut,:) = [];
alphas(linCut,:) = [];
betas(linCut,:) = [];
t_alphas(linCut,:) = [];
t_betas(linCut,:) = [];
[~,cols] = find(isnan(R2));
colCut = unique(cols);
R2(:,colCut) = [];
alphas(:,colCut) = [];
betas(:,colCut) = [];
t_alphas(:,colCut) = [];
t_betas(:,colCut) = [];

meanR2 = mean(mean(R2))
meanAlphas = mean(mean(alphas))
meanBetas = mean(mean(betas))
meanD = 1/(2*meanBetas)
meanTAlphas = mean(mean(t_alphas))
meanTBetas = mean(mean(t_betas))

meanR2_i = (mean(R2))';
meanBetas_i = (mean(betas))';
meanD_i = 1./(2*mean(betas)');
table = [meanR2_i meanBetas_i meanD_i];
vecBetas = reshape(betas', size(betas,1)*size(betas,2),1);

boxplot(betas);
autocorr(vecBetas,80);

%ativo = 'DI1F17';
%save(strcat(ativo,'_mean_intraday.mat'), 'table');
%save(strcat(ativo,'_betas.mat'), 'vecBetas');

Código para a estimação do impacto de mercado pelo modelo de Kyle para dados diários
dado pela equação 1.1.1:

%-----PARAMETROS-----
timeOpenMkt = [10,01,00];

```

```

timeCloseMkt = [16,54,00];

DP_day = [];
DV_day = [];
nbDias = size(IF_M,1);
for d = 1:nbDias
    iniM = IF_M(d,1);
    fimM = IF_M(d,2);

    dt = datevec(M(iniM,1));
    year = dt(1); month = dt(2); day = dt(3);
    timeBeg = datenum(year, month, day, timeOpenMkt(1), timeOpenMkt(2), timeOpenMkt(3));
    timeEnd = datenum(year, month, day, timeCloseMkt(1), timeCloseMkt(2), timeCloseMkt(3));

    begD = (iniM-1) + find(M(iniM:fimM,1)>timeBeg,1,'first');
    endD = (iniM-1) + find(M(iniM:fimM,1)>timeEnd,1,'first');

    indSell = (begD-1) + find(M(begD:endD, 4) == 1);
    indBuy = (begD-1) + find(M(begD:endD, 4) == 2);

    if isempty(indBuy) || isempty(indSell)
        continue;
    end

    DP_day = [DP_day; (M(endD,2)-M(begD,2))/ticker];
    DV_day = [DV_day; (sum(M(indBuy,3))-sum(M(indSell,3)))];
end

scatter(DV_day,DP_day);
mdl = fitlm(DV_day,DP_day);

R2 = mdl.Rsquared.Ordinary
alpha = mdl.Coefficients.Estimate(1)
beta = mdl.Coefficients.Estimate(2)
D = 1/(2*beta)
t_alpha = mdl.Coefficients.tStat(1)
t_beta = mdl.Coefficients.tStat(2)

```

Código para amostragem dos negócios

```

%-----INPUTS-----
freq = 1*60; %segundos
timeOpen = '10:00:00';
timeClose = '16:54:00';

nbDias = size(IF_M,1);
B = zeros(nbDias*(10*(3600/freq)),9); % BARRA PARA 10 HORAS DE MERCADO
IF_B = zeros(nbDias, 2);
for d = 1:nbDias
    iniM = IF_M(d, 1);
    fimM = IF_M(d, 2);
    tBeg = M(iniM,1);
    tEnd = M(fimM,1);

    tBeg = floor(tBeg) + datenum(timeOpen, 'HH:MM:SS')-floor(datenum(timeOpen, 'HH:MM:SS'))
    tEnd = floor(tEnd) + datenum(timeClose, 'HH:MM:SS')-floor(datenum(timeClose, 'HH:MM:SS'))

    j = 1;

```



```

dt = [];
while true
    newT = tBeg + j*(freq/(24*3600));
    if newT > tEnd
        dt = [dt; tEnd];
        break;
    end
    dt = [dt; newT];
    j = j + 1;
end

nbBars = length(dt);
if d == 1
    IF_B(d,1) = 1;
    IF_B(d,2) = nbBars;
else
    IF_B(d,1) = IF_B(d-1,2)+1;
    IF_B(d,2) = IF_B(d,1) + nbBars-1;
end
iniB = IF_B(d,1);
fimB = IF_B(d,2);
B(iniB:fimB,1) = dt;

for i = iniB:fimB
    if i == iniB
        ind = (iniM-1) + find(M(iniM:fimM,1) < B(i,1));
    else
        ind = (iniM-1) + find(and(M(iniM:fimM,1) >= B(i-1,1), M(iniM:fimM,1) < B(i,1)));
    end

    if isempty(ind) ~= 1
        open = M(ind(1),2);
        [high, t1] = max(M(ind,2));
        [low, t2] = min(M(ind,2));
        close = M(ind(end),2);
        volume = sum(M(ind,3));
        vwap = (M(ind,2)'*M(ind,3))/volume;

        indBuy = (ind(1)-1) + find(M(ind,4)==2);
        indSell = (ind(1)-1) + find(M(ind,4)==1);
        volumeBuy = sum(M(indBuy,3));
        volumeSell = sum(M(indSell,3));

        B(i, 2) = open;
        B(i, 3) = high;
        B(i, 4) = low;
        B(i, 5) = close;
        B(i, 6) = vwap;
        B(i, 7) = volume;
        B(i, 8) = volumeBuy;
        B(i, 9) = volumeSell;
    else
        if i > iniB
            B(i,2:6) = B(i-1,5);
        end
        B(i, 7) = 0;
        B(i, 8) = 0;
        B(i, 9) = 0;
    end
end

```

```

        end
    end
end

B((IF_B(end,2)+1):end, :) = [];
clear ans begDate endDate close d dt endD fimB fimD fimM flag freq high i ind iniB iniD in

```

Código para estimação do coeficiente de impacto de preço utilizando a variável *trade imbalance* (TI) da equação 1.1.10:

```

%-----PARAMETROS-----
janela = 100; %numero de amostras para estimar

DP = (B(:,5)-B(:,2))/ticker;
DV = (B(:,8)-B(:,9));

n = length(DP);
R2 = [];
alphas = [];
betas = [];
t_alphas = [];
t_betas = [];

nbDias = size(IF_B,1);

for d = 1:nbDias
    ind = IF_B(d,1):IF_B(d,2);
    n = floor(length(ind)/janela);

    for i = 1:n
        if i < n
            idx = ind(1) + (i-1)*janela:ind(1) + i*janela;
        else
            idx = ind(1) + (i-1)*janela:ind(end);
        end

        mdl = fitlm(DV(idx),DP(idx));
        R2 = [R2; mdl.Rsquared.Ordinary];
        alphas = [alphas; mdl.Coefficients.Estimate(1)];
        betas = [betas; mdl.Coefficients.Estimate(2)];
        t_alphas = [t_alphas; mdl.Coefficients.tStat(1)];
        t_betas = [t_betas; mdl.Coefficients.tStat(2)];
    end
end

linCut = isnan(R2);
alphas(linCut) = [];
betas(linCut) = [];
R2(linCut) = [];
t_alphas(linCut) = [];
t_betas(linCut) = [];

meanR2 = mean(R2)
meanAlpha = mean(alphas)
meanBeta = mean(betas)
meanD = 1/meanBeta
meanTAlpha = mean(t_alphas)
meanTBeta = mean(t_betas)

```

Código utilizado para a fazer a calibragem econométrica (SARIMA) na série do parâmetro β

```

y = vecBetas;

figure(1);
subplot(2,2,1)
plot(y)
title('Serie \beta')
subplot(2,2,2)
qqplot(y);
title('QQ plot')
subplot(2,2,3)
autocorr(y,80)
subplot(2,2,4)
parcorr(y,80)

SLags = 14;
model = arima('SMALags', SLags, 'ARLags', 1, 'MALags', 1,...
    'Seasonality',SLags, 'Constant', 0);

[fit, CV, logL] = estimate(model,y);
[res, V] = infer(fit,y);
stdr = res./sqrt(V);

figure(2);
subplot(2,2,1)
plot(stdr)
title('Residuo normalizado')
subplot(2,2,2)
qqplot(stdr);
title('QQ plot')
subplot(2,2,3)
autocorr(stdr,80)
subplot(2,2,4)
parcorr(stdr,80)

```

Funções auxiliares para maior ou igual, e menor ou igual para mitigar erros de arredondamento decimal

```

function r = MaiorOuIgual(v1,v2,tol)
    r = or(abs(v1-v2) < tol, v1 - v2 > tol);
end

function r = MenorOuIgual(v1,v2,tol)
    r = or(abs(v1-v2) < tol, v2-v1 > tol);
end

```


Apêndice C

Tabelas

Tabela C.1: Média dos parâmetros do modelo OFI: PETR4

Intervalo	β	\hat{D}	R^2
10:30	0.000031	16219	45%
11:00	0.000025	20253	59%
11:30	0.000024	21166	57%
12:00	0.000022	22366	56%
12:30	0.000022	22707	59%
13:00	0.000020	25354	59%
13:30	0.000019	25895	57%
14:00	0.000019	26206	53%
14:30	0.000019	26205	52%
15:00	0.000019	26952	52%
15:30	0.000018	27899	52%
16:00	0.000018	28307	50%
16:30	0.000016	31314	49%
17:00	0.000016	31292	45%
Média Total	0.000020	25153	53%

Tabela C.2: Média dos parâmetros do modelo OFI: VALE5

Intervalo	$\hat{\beta}$	\hat{D}	R^2
10:30	0.000070	7102	43%
11:00	0.000056	8950	61%
11:30	0.000054	9230	61%
12:00	0.000054	9272	60%
12:30	0.000054	9247	60%
13:00	0.000052	9600	63%
13:30	0.000049	10141	59%
14:00	0.000050	9954	59%
14:30	0.000049	10225	56%
15:00	0.000048	10450	56%
15:30	0.000046	10788	55%
16:00	0.000044	11358	54%
16:30	0.000043	11502	54%
17:00	0.000042	12031	50%
Média Total	0.000051	9989	57%

Tabela C.3: Média dos parâmetros do modelo OFI: CYRE3

Intervalo	$\hat{\beta}$	\hat{D}	R^2
10:30	0.000284	1762	47%
11:00	0.000212	2358	62%
11:30	0.000184	2710	57%
12:00	0.000181	2769	57%
12:30	0.000178	2810	57%
13:00	0.000171	2921	54%
13:30	0.000167	2988	53%
14:00	0.000155	3225	52%
14:30	0.000145	3437	47%
15:00	0.000161	3114	52%
15:30	0.000148	3375	50%
16:00	0.000144	3467	48%
16:30	0.000133	3756	49%
17:00	0.000124	4042	46%
Média Total	0.000171	3052	52%

Tabela C.4: Média dos parâmetros do modelo OFI: CSAN3

Intervalo	$\hat{\beta}$	\hat{D}	R^2
10:30	0.001032	484	39%
11:00	0.000809	618	53%
11:30	0.000728	686	54%
12:00	0.000703	712	52%
12:30	0.000660	757	52%
13:00	0.000639	782	49%
13:30	0.000642	779	50%
14:00	0.000666	751	47%
14:30	0.000618	809	46%
15:00	0.000588	851	47%
15:30	0.000598	836	49%
16:00	0.000565	885	46%
16:30	0.000542	922	46%
17:00	0.000529	945	43%
Média Total	0.000666	773	48%

Tabela C.5: Média dos parâmetros do modelo OFI: GOLL4

Intervalo	$\hat{\beta}$	\hat{D}	R^2
10:30	0.000285	1755	48%
11:00	0.000230	2175	53%
11:30	0.000199	2513	50%
12:00	0.000186	2684	49%
12:30	0.000202	2473	51%
13:00	0.000186	2683	49%
13:30	0.000161	3104	43%
14:00	0.000152	3292	40%
14:30	0.000165	3038	40%
15:00	0.000146	3423	39%
15:30	0.000200	2500	39%
16:00	0.000166	3012	39%
16:30	0.000154	3251	37%
17:00	0.000134	3722	32%
Média Total	0.000183	2830	43%

Tabela C.6: Média dos parâmetros do modelo OFI: ELET6

Intervalo	$\hat{\beta}$	\hat{D}	R^2
10:30	0.000232	2157	42%
11:00	0.000171	2918	49%
11:30	0.000145	3450	42%
12:00	0.000142	3523	42%
12:30	0.000147	3409	41%
13:00	0.000138	3613	40%
13:30	0.000143	3486	39%
14:00	0.000135	3710	38%
14:30	0.000137	3652	39%
15:00	0.000137	3638	41%
15:30	0.000129	3861	41%
16:00	0.000119	4214	37%
16:30	0.000113	4422	38%
17:00	0.000109	4582	36%
Média Total	0.000143	3617	40%

Tabela C.7: Média dos parâmetros do modelo OFI: dolar futuro (1º vencimento)

Intervalo	$\hat{\beta}$	\hat{D}	R^2
9:30	0.003366	148.6	59%
10:00	0.002756	181.5	62%
10:30	0.002080	240.4	56%
11:00	0.002448	204.2	60%
11:30	0.002101	238.0	57%
12:00	0.002267	220.5	60%
12:30	0.001986	251.7	55%
13:00	0.002255	221.7	59%
13:30	0.002362	211.7	56%
14:00	0.002939	170.1	61%
14:30	0.002560	195.3	55%
15:00	0.002358	212.0	56%
15:30	0.002198	227.5	54%
16:00	0.002093	238.9	52%
16:30	0.002251	222.1	52%
17:00	0.002422	206.4	50%
17:30	0.002429	205.8	46%
18:00	0.002244	222.8	40%
Média Total	0.002395	212.2	55%

Tabela C.8: Média dos parâmetros do modelo OFI: o juros futuro (DI1F17)

Intervalo	$\hat{\beta}$	\hat{D}	R^2
9:30	0.000525	952	36%
10:00	0.000412	1215	35%
10:30	0.000386	1295	34%
11:00	0.000335	1492	33%
11:30	0.000355	1410	35%
12:00	0.000370	1352	34%
12:30	0.000363	1378	33%
13:00	0.000396	1264	32%
13:30	0.000459	1089	35%
14:00	0.000468	1069	35%
14:30	0.000421	1188	33%
15:00	0.000365	1370	31%
15:30	0.000323	1546	29%
16:00	0.000226	2211	22%
Média Total	0.000386	1345	33%

Referências Bibliográficas

- [1] Hee-Joon Ahn, Kee-Hong Bae, and Kalok Chan, *Limit orders, depth, and volatility: evidence from the stock exchange of hong kong*, The Journal of Finance **56** (2001), no. 2, 767–788.
- [2] Aurélien Alfonsi and Alexander Schied, *Capacitary measures for completely monotone kernels via singular control*, SIAM Journal on Control and Optimization **51** (2013), no. 2, 1758–1780.
- [3] Robert Almgren and Neil Chriss, *Optimal execution of portfolio transactions*, Journal of Risk **3** (2001), 5–40.
- [4] Robert Almgren and Julian Lorenz, *Adaptive arrival price*, Trading **2007** (2007), no. 1, 59–66.
- [5] Robert Almgren, Chee Thum, Emmanuel Hauptmann, and Hong Li, *Direct estimation of equity market impact*, Risk **57** (2005).
- [6] Torben G Andersen and Tim Bollerslev, *Deutsche mark–dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies*, the Journal of Finance **53** (1998), no. 1, 219–265.
- [7] Erhan Bayraktar and Michael Ludkovski, *Optimal trade execution in illiquid markets*, Mathematical Finance **21** (2011), no. 4, 681–701.
- [8] Dimitris Bertsimas and Andrew W Lo, *Optimal control of execution costs*, Journal of Financial Markets **1** (1998), no. 1, 1–50.
- [9] Jean-Philippe Bouchaud, J Doyne Farmer, and Fabrizio Lillo, *How markets slowly digest changes in supply and demand*, Handbook of financial markets: dynamics and evolution **1** (2009), 57.
- [10] Jean-Philippe Bouchaud, Yuval Gefen, Marc Potters, and Matthieu Wyart, *Fluctuations and response in financial markets: the subtle nature of ‘random’ price changes*, Quantitative Finance **4** (2004), no. 2, 176–190.
- [11] Patrick Cheridito and Tardu Sepin, *Optimal trade execution under stochastic volatility and liquidity*, Applied Mathematical Finance (2014), no. ahead-of-print, 1–21.
- [12] Rama Cont, Arseniy Kukanov, and Sasha Stoikov, *The price impact of order book events*, arXiv preprint arXiv:1011.6402 (2010).
- [13] Jonathan Donier, *Market impact with autocorrelated order flow under perfect competition*, Available at SSRN 2191660 (2012).

- [14] Martin DD Evans and Richard K Lyons, *Order flow and exchange rate dynamics*, Tech. report, National bureau of economic research, 1999.
- [15] J Doyne Farmer, *Market force, ecology and evolution*, *Industrial and Corporate Change* **11** (2002), no. 5, 895–953.
- [16] Jim Gatheral, *No-dynamic-arbitrage and market impact*, *Quantitative Finance* **10** (2010), no. 7, 749–759.
- [17] Jim Gatheral and Alexander Schied, *Dynamical models of market impact and algorithms for order execution*, Available at SSRN 2034178 (2013).
- [18] Jim Gatheral, Alexander Schied, and Alla Slynko, *Transient linear price impact and fredholm integral equations*, *Mathematical Finance* **22** (2012), no. 3, 445–474.
- [19] Joel Hasbrouck and Duane J Seppi, *Common factors in prices, order flows, and liquidity*, *Journal of financial Economics* **59** (2001), no. 3, 383–411.
- [20] Gur Huberman and Werner Stanzl, *Price manipulation and quasi-arbitrage*, *Econometrica* **72** (2004), no. 4, 1247–1275.
- [21] Charles M Jones, Gautam Kaul, and Marc L Lipson, *Transactions, volume, and volatility*, *Review of Financial Studies* **7** (1994), no. 4, 631–651.
- [22] Albert Pete Kyle and Anna Obizhaeva, *Market microstructure invariants: Theory and implications of calibration*, Available at SSRN 1978932 (2011).
- [23] Albert S Kyle, *Continuous auctions and insider trading*, *Econometrica: Journal of the Econometric Society* (1985), 1315–1335.
- [24] Charles MC Lee, Belinda Mucklow, and Mark J Ready, *Spreads, depths, and the impact of earnings information: An intraday analysis*, *Review of Financial Studies* **6** (1993), no. 2, 345–374.
- [25] Thomas H McNish and Robert A Wood, *An analysis of intraday patterns in bid/ask spreads for nyse stocks*, *the Journal of Finance* **47** (1992), no. 2, 753–764.
- [26] Somayeh Moazeni, Thomas F Coleman, and Yuying Li, *Optimal execution under jump models for uncertain price impact*, *Journal of Computational Finance* **16** (2013), no. 4, 1–44.
- [27] Anna A Obizhaeva and Jiang Wang, *Optimal trading strategy and supply/demand dynamics*, *Journal of Financial Markets* (2012).
- [28] Andre F Perold, *The implementation shortfall: Paper versus reality*, *Streetwise: the best of the Journal of portfolio management* (1998), 106.
- [29] Vasiliki Plerou, Parameswaran Gopikrishnan, Xavier Gabaix, and H Eugene Stanley, *Quantifying stock-price response to demand fluctuations*, *Physical Review E* **66** (2002), no. 2, 027104.
- [30] Marc Potters and Jean-Philippe Bouchaud, *More statistical properties of order books and price impact*, *Physica A: Statistical Mechanics and its Applications* **324** (2003), no. 1, 133–140.

- [31] Alexander Schied, Torsten Schöneborn, and Michael Tehranchi, *Optimal basket liquidation for cara investors is deterministic*, Applied Mathematical Finance **17** (2010), no. 6, 471–489.