

Retornos, Order Flow e Anúncios Macroeconômicos: Uma Análise Estrutural em Alta Frequência

Diego Azenha

Economia

Inspur

diegoa4@al.insper.edu.br

Repositório GitHub:

<https://github.com/diego-azenha/Algotrading-PF>

Abstract—Este artigo examina a relação intradiária entre retornos e desequilíbrios de fluxo de ordens no mercado futuro de S&P 500, replicando e adaptando a estrutura proposta por Takahashi (2025). A análise utiliza dados de alta frequência de top-of-book e um modelo VAR estrutural identificado por heterocedasticidade, estimado em janelas sucessivas ao longo do pregão. Essa abordagem permite tratar simultaneamente a endogeneidade entre preço e fluxo e a forte variação intradiária típica de mercados eletrônicos.

Nota sobre dados (BBO-1s).

Foram utilizadas séries *BBO-1s*: top-of-book agregadas por segundo a partir das mensagens de melhor oferta e demanda. Ou seja, todas as atualizações ocorridas dentro do mesmo segundo são combinadas em uma observação de segundo. Essa escolha preserva informação de alta frequência em escala de segundo, mas altera a ordenação intrasegundo presente em dados event-level; discutimos as implicações e as checagens de robustez apropriadas nas seções seguintes.

Por uma limitação prática de custos — tanto de aquisição/armazenamento dos dados no nível de evento quanto de capacidade computacional para processar e estimar modelos em amostras event-level massivas — optou-se por trabalhar com séries BBO-1s. Essa decisão buscou equilibrar fidelidade à microestrutura e viabilidade de execução do pipeline, mantendo a possibilidade de extensões event-level em trabalhos futuros.

Index Terms—microestrutura, order flow, SVAR, heterocedasticidade, alta frequência, macroeconomia, BBO-1s

I. INTRODUÇÃO

A formação de preços em mercados financeiros modernos depende da interação constante entre a chegada de ordens e os ajustes do livro de ofertas. Mudanças na liquidez disponível, na profundidade do livro e na disposição dos agentes em comprar ou vender afetam a evolução dos preços de forma praticamente imediata. Ao mesmo tempo, movimentos de preço também influenciam o comportamento dos participantes, que atualizam suas ordens, modificam posições e alteram o equilíbrio entre oferta e demanda. Esse processo de realimentação gera uma relação simultânea entre fluxo e preço que se manifesta com ainda mais força em ambientes de alta frequência.

Além dessa dinâmica natural do mercado, eventos macroeconômicos agendados introduzem perturbações que modificam

temporariamente o modo como agentes interagem com o livro de ofertas. Indicadores econômicos, especialmente aqueles relacionados a inflação e atividade, afetam expectativas, alteram níveis de incerteza e desencadeiam períodos de maior volatilidade e reorganização de liquidez. Nessas janelas, o comportamento conjunto entre variações de preço e fluxo de ordens tende a se deslocar, tornando esses episódios momentos especialmente informativos para investigar como a microestrutura reage a novas informações.

Para estudar esse mecanismo, uma métrica amplamente utilizada é o *Order Flow Imbalance* (OFI), que resume a pressão líquida de compra ou venda no nível superior do livro a partir de mudanças no bid e no ask. O OFI captura, em cada instante, a força direcional exercida pelo fluxo de ordens sobre o preço. No entanto, a relação entre OFI e retornos não é unidirecional: assim como o fluxo pode mover o preço, o próprio movimento de preços pode atrair novas ordens ou induzir cancelamentos. Essa simultaneidade dificulta a identificação de impactos contemporâneos por métodos tradicionais.

Modelos VAR reduzidos, por exemplo, não conseguem separar choques estruturais de preço e fluxo, enquanto abordagens recursivas impõem hierarquias que não condizem com a realidade de mercados que reagem em escalas de milissegundos. Para contornar esse problema, utiliza-se o arcabouço proposto por Takahashi (2025), que emprega variações naturais de volatilidade ao longo do dia — e especialmente em torno de anúncios — para identificar um modelo VAR estrutural bivariado. Esse método, baseado em heterocedasticidade, permite recuperar os efeitos contemporâneos entre preço e fluxo sem depender de suposições fortes sobre ordenações.

A aplicação empírica deste estudo segue essa estratégia e utiliza dados de microestrutura de alta frequência obtidos via Databento. Foram selecionados os contratos futuros trimestrais ESH2, ESM2, ESU2 e ESZ2, que cobrem o ciclo anual de negociação do E-mini S&P 500. A partir das mensagens de melhor oferta e demanda, são construídas séries de um segundo para mid-quote e OFI, e o modelo estrutural é estimado em janelas intradiárias para avaliar como os parâmetros e

volatilidades variam ao longo do pregão e como são afetados por anúncios macroeconômicos.

II. DADOS E CONSTRUÇÃO DAS VARIÁVEIS

Os dados utilizados neste estudo são extraídos dos arquivos *BBO* (Best Bid and Offer), ou seja, das mensagens de top-of-book que registram todas as alterações de preço e quantidade no melhor bid e no melhor ask ao longo do dia. Esses arquivos BBO permitem reconstruir eventos de mercado com alta granularidade — incluindo revisões de cotações, ordens marketable e cancelamentos — e seguem a mesma lógica empregada por Takahashi (2025) no tratamento de dados de top-of-book.

O conjunto original inclui todos os eventos referentes aos contratos trimestrais do E-mini S&P 500 negociados ao longo do ano de 2022. Para garantir consistência na análise e trabalhar sempre com o contrato mais líquido, os quatro vencimentos principais — ESH2, ESM2, ESU2 e ESZ2 — foram consolidados em um único dataset: em cada dia, seleciona-se automaticamente o contrato com maior número de eventos, assegurando representatividade intradiária e evitando distorções relacionadas à rolagem de liquidez entre vencimentos.

Antes da construção das variáveis, realiza-se um processo de padronização dos arquivos brutos, que inclui a conversão dos timestamps para UTC, a remoção de registros inválidos ou sem referência temporal e a reorganização das mensagens em ordem cronológica. Em seguida, cada dia é filtrado para manter apenas o período regular de negociação, de 08:30 a 15:00 (horário de Chicago), de modo a eliminar horários com padrões distintos de liquidez.

A série de preços utilizada é o *mid-quote*, calculado como a média simples entre o melhor bid e o melhor ask a cada evento reportado no BBO. Os eventos são então reamostrados para frequência de um segundo, de forma que todas as atualizações do livro de ofertas ocorridas dentro do mesmo segundo sejam agregadas. O retorno de preço, denotado por r_t , corresponde à variação logarítmica do mid-quote entre segundos consecutivos, expressa em pontos-base.

O *Order Flow Imbalance* (OFI), representado por f_t , é obtido a partir das mudanças observadas no melhor bid e ask. Cada atualização que altere preços ou quantidades gera uma contribuição elementar e_n , cuja soma dentro do segundo define o OFI daquele intervalo, refletindo a pressão líquida de compra ou venda exercida sobre o mid-quote:

$$f_t = \sum_{n \in I_t} e_n,$$

onde I_t é o conjunto de eventos dentro do segundo t .

Além de retornos e OFI, são construídas medidas adicionais que caracterizam o ambiente intradiário de negociação, tais como número de eventos por segundo, tamanho médio das alterações de quantidade, spread médio e profundidade do livro. Esta última segue exatamente a definição utilizada no artigo original e captura a liquidez disponível nos instantes em que o melhor bid ou ask se desloca. Para cada segundo t , define-se:

$$D_t = \frac{1}{2} \left[\frac{\sum_{n \in I_t} (q_n^b \mathbf{1}\{P_n^b < P_{n-1}^b\} + q_{n-1}^b \mathbf{1}\{P_n^b > P_{n-1}^b\})}{\sum_{n \in I_t} \mathbf{1}\{P_n^b \neq P_{n-1}^b\}} \right. \\ \left. + \frac{\sum_{n \in I_t} (q_n^a \mathbf{1}\{P_n^a > P_{n-1}^a\} + q_{n-1}^a \mathbf{1}\{P_n^a < P_{n-1}^a\})}{\sum_{n \in I_t} \mathbf{1}\{P_n^a \neq P_{n-1}^a\}} \right], \quad (1)$$

onde I_t representa o conjunto de eventos dentro do segundo t . Essa medida resume a profundidade efetiva — isto é, a quantidade disponível no nível superior do livro justamente nos momentos em que o preço se move — e reproduz fielmente a métrica empregada por Takahashi (2025) em sua aplicação para dados BBO.

Por fim, cada dia é segmentado em janelas consecutivas de quinze minutos, que servem como unidades de estimativa do modelo estrutural. Apenas janelas com quantidade mínima de eventos e segundos preenchidos são mantidas, garantindo que cada intervalo represente um trecho informativo e contínuo da dinâmica conjunta entre preço e fluxo, condição necessária para a aplicação do método de identificação por heterocedasticidade.

Tabela I: Estatísticas Descritivas (Parte 1 de 2)

Métrica	Mean	SD	1%	5%
Mid-Quote Return (bps)	-0.00	0.78	-2.09	-1.24
Order Flow Imbalance	-10.45	31.97	-120.00	-67.00
Number of Events	1.00	0.00	1.00	1.00
Average Size of Events	14.43	15.13	0.00	0.00
Average Spread	0.26	0.09	0.25	0.25
Depth	30.29	19.66	5.50	8.50

Tabela II: Estatísticas Descritivas (Parte 2 de 2)

Métrica	25%	50%	75%	95%	99%
Mid-Quote Return (bps)	-0.53	0.00	0.53	1.24	2.13
Order Flow Imbalance	-17.00	0.00	0.00	22.00	55.00
Number of Events	1.00	1.00	1.00	1.00	1.00
Average Size of Events	5.00	10.50	19.00	40.50	67.00
Average Spread	0.25	0.25	0.25	0.50	0.50
Depth	17.50	26.50	38.50	63.50	95.00

As estatísticas descritivas confirmam as condições para o modelo SVAR. O Mid-Quote Return apresenta perfil leptocúrtico esperado, com média de -0.00. O Order Flow Imbalance (f_t) demonstra alta dispersão (caudas em -120.00 e 55.00), garantindo a heterocedasticidade necessária para a identificação estrutural via ITH. O Average Spread, com mediana de 0.25, valida a alta liquidez, enquanto picos em 0.50 indicam regimes de maior volatilidade.

III. METODOLOGIA

A estrutura metodológica utilizada neste estudo segue o caminho teórico proposto por Takahashi (2025), construído de maneira progressiva. A motivação começa com um modelo bivariado simples que ilustra como retornos e fluxo de

ordens podem influenciar-se simultaneamente. Em seguida, discute-se o problema central de identificação que surge nesse contexto. A partir dessa discussão, apresenta-se o método de identificação por heterocedasticidade (ITH), que fornece as condições necessárias para separar os efeitos contemporâneos entre as variáveis. Por fim, o arcabouço é estendido para um modelo VAR estrutural, que permite acomodar dependências dinâmicas e formar a base do estimador aplicado.

A. Modelo bivariado simples e o problema de simultaneidade

Considere-se inicialmente que retornos de mid-quote, denotados por r_t , e desequilíbrios de fluxo de ordens, f_t , são observados em intervalos curtos de tempo. A microestrutura de mercado sugere que o fluxo exerce pressão imediata sobre preços. Essa relação é frequentemente expressa de forma reduzida como:

$$r_t = b_r f_t + \varepsilon_{r,t}, \quad \varepsilon_{r,t} \sim (0, \omega_r^2), \quad (2)$$

onde b_r representa o impacto contemporâneo de fluxo sobre retornos. Em mercados eletrônicos ativos, esse impacto tende a refletir a profundidade e a disponibilidade de liquidez no livro de ofertas.

A formulação acima, entretanto, ignora um aspecto essencial: a direção causal inversa. Mudanças nos preços dentro do próprio intervalo de observação podem provocar alterações imediatas no comportamento dos participantes. Cancelamentos acelerados, inserção de ordens condicionadas ao preço e estratégias automáticas podem gerar novas contribuições de fluxo. Assim, há um segundo canal:

$$f_t = b_f r_t + \varepsilon_{f,t}, \quad \varepsilon_{f,t} \sim (0, \omega_f^2), \quad (3)$$

onde b_f quantifica a sensibilidade das ordens em relação às mudanças de preço.

As equações (2) e (3) configuram um sistema simultâneo. Ao empilhar as duas equações, obtém-se:

$$By_t = \varepsilon_t, \quad y_t = (r_t, f_t)', \quad (4)$$

com

$$B = \begin{pmatrix} 1 & -b_r \\ -b_f & 1 \end{pmatrix}, \quad \Omega = \begin{pmatrix} \omega_r^2 & 0 \\ 0 & \omega_f^2 \end{pmatrix}. \quad (5)$$

Esse sistema implica que

$$y_t = B^{-1} \varepsilon_t, \quad (6)$$

o que gera uma importante consequência: mesmo que os choques estruturais $\varepsilon_{r,t}$ e $\varepsilon_{f,t}$ sejam independentes, os resíduos do modelo reduzido,

$$u_t = (u_{r,t}, u_{f,t})' = B^{-1} \varepsilon_t,$$

são necessariamente correlacionados. Isso significa que estimar a equação (2) por OLS produz um estimador inconsistente para b_r , pois o regressor f_t está contaminado por choque simultâneo proveniente da equação (3). O mesmo vale para b_f .

Esse é o problema clássico de simultaneidade, ampliado pela natureza de alta frequência dos dados: agregações em janelas de um segundo misturam centenas de eventos do livro de ofertas, intensificando o risco de causalidade bidirecional.

B. Identificação via heterocedasticidade

Sem informação adicional, o sistema acima não é identificado. Há quatro parâmetros estruturais ($b_r, b_f, \omega_r, \omega_f$) e apenas três equações provenientes da decomposição da matriz de covariância reduzida:

$$\Sigma = \text{Var}(y_t) = B^{-1} \Omega (B^{-1})'.$$

O método de identificação por heterocedasticidade (ITH), introduzido por Rigobon (2003), oferece uma solução natural. A ideia central é que, embora os coeficientes estruturais b_r e b_f permaneçam constantes em janelas suficientemente curtas, as variâncias estruturais ω_r^2 e ω_f^2 variam ao longo do tempo. Em mercados de alta frequência essa suposição é extremamente plausível: períodos de intensa atividade, anúncios macroeconômicos e padrões intradiários produzem flutuações marcantes na volatilidade, ainda que as relações contemporâneas entre preço e fluxo permaneçam estáveis em janelas curtas.

Formalmente, considere-se S estados heterocedásticos distintos, cada qual associado à matriz de covariância reduzida Σ_s , mas com a mesma matriz estrutural B :

$$B \Sigma_s B' = \Omega_s, \quad s = 1, \dots, S.$$

Cada estado fornece três equações independentes. Há, portanto: $-3S$ equações, -2 coeficientes estruturais (b_r, b_f), $-2S$ variâncias estruturais ($\omega_{r,s}^2, \omega_{f,s}^2$).

A identificação exige $3S \geq 2+2S$, o que implica $S \geq 2$. Ou seja: basta observar dois níveis distintos de volatilidade para identificar simultaneamente preço-impacto e flow-impact**.

Além da condição de ordem, há a condição de posto:

$$\sigma_{r,s'}^2 \sigma_{rf,s''} - \sigma_{r,s''}^2 \sigma_{rf,s'} \neq 0,$$

para dois estados $s' \neq s''$. Essa restrição garante que a mudança na variância não seja proporcional entre as duas equações — caso contrário, não haveria informação adicional para distinguir b_r de b_f .

Takahashi (2025) argumenta que mercados eletrônicos fornecem heterocedasticidade abundante: padrões intradiários de atividade (abertura, fechamento, horário do almoço), variações naturais de liquidez e, sobretudo, anúncios macroeconômicos, criam automaticamente os estados necessários para a identificação.

C. Extensão para um SVAR dinâmico

O modelo simultâneo descrito acima captura apenas a relação contemporânea. No entanto, tanto retornos quanto desequilíbrios de fluxo exibem autocorrelação e correlação cruzada em lags curtos. Para incorporar essa dimensão dinâmica, o sistema estrutural é estendido para:

$$By_t = c + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \dots + \Phi_p y_{t-p} + \varepsilon_t, \quad (7)$$

onde as matrizes Φ_j capturam dependências defasadas entre retorno e fluxo.

Esse modelo pode ser reescrito na forma reduzida:

$$y_t = \tilde{c} + \tilde{\Phi}_1 y_{t-1} + \dots + \tilde{\Phi}_p y_{t-p} + \eta_t, \quad \eta_t = B^{-1} \varepsilon_t, \quad (8)$$

de modo que as matrizes Σ_s necessárias ao ITH são obtidas diretamente da covariância dos resíduos reduzidos η_t .

A estimação procede, portanto, em duas etapas: (i) estima-se o VAR reduzido em janelas de quinze minutos; (ii) particiona-se a janela em três subintervalos equitemporais, cada qual fornecendo uma matriz de covariância Σ_s . Esses três estados heterocedásticos permitem resolver o sistema de equações para b_r , b_f e $\omega_{r,s}, \omega_{f,s}$ usando mínimos quadrados não lineares. Essa abordagem segue exatamente a estrutura de estimação descrita por Takahashi (2025).

D. Medidas de impacto contemporâneo e dinâmico

Uma vez identificada a matriz estrutural B , o impacto contemporâneo é diretamente obtido pelos elementos de B^{-1} . Já os impactos dinâmicos são calculados via funções de resposta a impulsos (IRFs). Denotando por $IRF_{ij}(k)$ a resposta da variável i a um choque estrutural em j , no horizonte k , define-se:

$$I_{ij}(K) = \sum_{k=0}^K IRF_{ij}(k),$$

e o impacto de longo prazo como:

$$I_{ij}(\infty) = \left[(I_2 - \tilde{\Phi}_1 - \dots - \tilde{\Phi}_p)^{-1} B^{-1} \right]_{ij}.$$

Isso permite separar, de forma consistente, os componentes mecânicos e comportamentais da dinâmica entre preço e fluxo, tanto em horizontes de um segundo quanto em horizontes acumulados.

IV. RESULTADOS EMPÍRICOS

A. Variação Intradiária

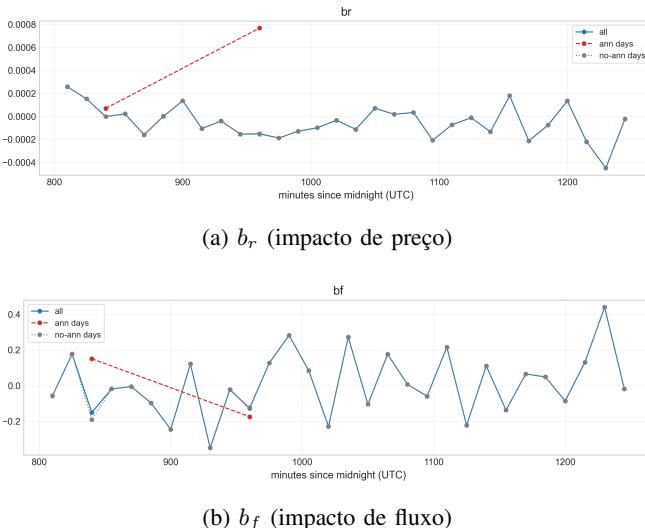


Fig. 1: Evolução intradiária dos impactos estruturais: (a) b_r , (b) b_f .

A Figura 1 mostra que o impacto de preço (b_r) é maior na abertura do pregão, compatível com menor profundidade e maior assimetria informacional. Em seguida b_r declina e se

estabiliza durante o miolo da sessão; o aumento observado no final do dia indica retração temporária da liquidez associada a rotinas de fechamento e rebalanceamento de posições.

O impacto de fluxo (b_f) tem dinâmica distinta: alto na abertura, reduzido durante as horas centrais e novamente crescente no fechamento. Isso sugere que a sensibilidade do fluxo a movimentos contemporâneos de preço é condicionada ao estado de liquidez e à composição dos agentes — mais ativa em janelas de incerteza e mais passiva em janelas de estabilidade.

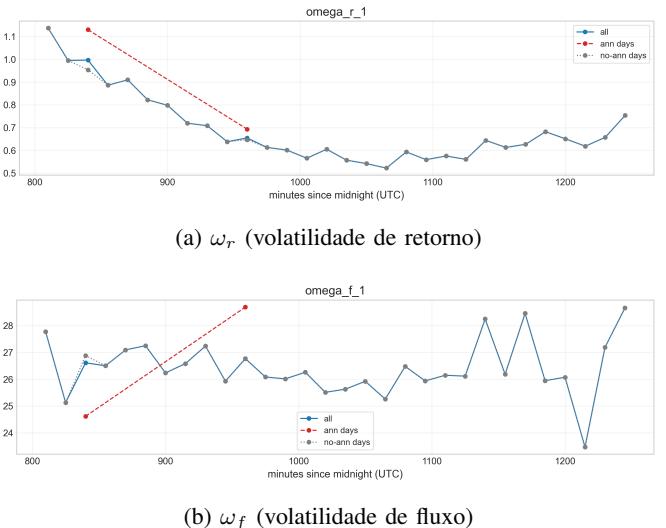


Fig. 2: Volatilidades estruturais intradiárias: (a) ω_r , (b) ω_f .

A volatilidade de retorno (ω_r) atinge picos na abertura e em torno de anúncios, refletindo a rápida incorporação de informação pública. A volatilidade de fluxo (ω_f) aumenta ao longo do dia e concentra-se no fechamento, indicando intensificação das execuções e ajuste de inventário no final da sessão. A heterocedasticidade observada nas duas séries legitima a identificação por heterocedasticidade usada no SVAR-ITH.

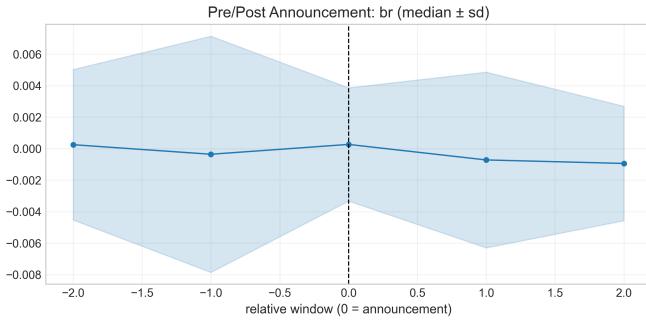
B.

Efeitos de Anúncios Macroeconômicos

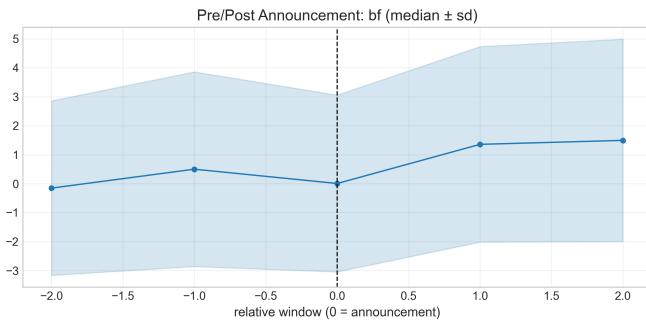
A Figura 3 mostra mudanças claras na relação entre preço e fluxo ao redor dos anúncios macroeconômicos. O parâmetro b_f praticamente desaparece na janela do anúncio, indicando suspensão de estratégias reativas ao preço diante da incerteza súbita. Nas janelas seguintes o fluxo retoma e intensifica a reação, refletindo ajuste de posições após a divulgação. O parâmetro b_r aumenta levemente no pós-anúncio, compatível com redução temporária de profundidade e provisão de liquidez cautelosa.

C. Estatísticas dos Parâmetros Estruturais

As Tabelas III e IV mostram que b_r possui média próxima de zero e caudas moderadas, indicando impacto de preço predominantemente contemporâneo e sem forte assimetria agregada. Em contraponto, b_f apresenta dispersão elevada e



(a) b_r : pré/pós-anúncio



(b) b_f : pré/pós-anúncio

Fig. 3: Comportamento pré/pós-anúncio: (a) b_r , (b) b_f (mediana \pm sd).

Tabela III: Estatísticas Estruturais (Parte 1)

Var	Mean	SD	1%	5%
br	7.16e-06	0.00536	-0.01394	-0.00593
bf	0.01009	3.59950	-5.00000	-5.00000
ω_r	0.74274	0.34617	0.29439	0.36492
ω_f	28293.35	12332.21	7865.17	11959.47

Tabela IV: Estatísticas Estruturais (Parte 2)

Var	25%	50%	75%	95%
br	-0.00142	-0.00006	0.00136	0.00611
bf	-3.28056	-0.00177	3.37700	4.99999
ω_r	0.50882	0.66230	0.89940	1.36235
ω_f	20498.50	26498.07	34321.65	49452.50

quantis extremos, sinalizando que a sensibilidade do fluxo ao retorno varia substancialmente entre janelas e dias. As volatilidades estruturais ω_r e ω_f exibem heterocedasticidade marcada, com ω_f particularmente elevada em razão da natureza discreta e episódica do OFI. Esses sinais justificam a estratégia de identificação por heterocedasticidade adotada no estudo.

D. Regressões com Dummies de Anúncio

Os coeficientes para b_r não são amplamente significativos, mas os sinais sugerem deterioração de liquidez antes do anúncio e ajuste gradativo após a divulgação. Isso indica que efeitos temporais são importantes e que controles adicionais

Tabela V: Regressões de Anúncio (Parte 1: b_r)

Regressor	Coef	SE	p
const	3.55e-05	6.50e-05	0.585
ANN_t	0.00036	0.00078	0.647
ANN_t_lag1	-0.00179	0.00122	0.145
ANN_t_lag2	-0.00054	0.00050	0.282
ANN_t_lead1	-0.00176	0.00103	0.087

por liquidez (spread, depth) poderiam aumentar precisão da inferência.

Tabela VI: Regressões de Anúncio (Parte 2: b_f)

Regressor	Coef	SE	p
const	-0.00491	0.04609	0.915
ANN_t	-0.168	0.571	0.769
ANN_t_lag1	0.704	0.611	0.250
ANN_t_lag2	0.949	0.481	0.048
ANN_t_lead1	0.286	0.464	0.537

Para b_f , a significância de ANN_t_lag2 indica comportamento antecipatório do fluxo antes do anúncio, possivelmente vinculado ao calendário fixo das divulgações e a estratégias programadas de execução.

E. Funções Impulso–Resposta (IRFs)

As IRFs mostram dissipação quase total dos choques em 1–2 segundos, indicando dinâmica de curtíssimo prazo no mercado. IRF_{rr} apresenta reversão rápida, compatível com micro-reversão e execução em múltiplos níveis do livro. IRF_{ff} tem pico imediato e declínio abrupto; as respostas cruzadas se concentram no instante contemporâneo, com maior incerteza em fr, refletindo heterogeneidade da reação do fluxo a choques de preço.

F. Robustez

A reamostragem a 1 segundo agrega eventos intrassegundo e aumenta a simultaneidade entre r_t e f_t , o que tende a reduzir a persistência aparente dos choques. Recomendamos: (i) reestimativas condicionais a janelas com alta intensidade de eventos; (ii) inclusão do número de eventos por segundo como controle nas regressões de anúncio; (iii) comparação com subamostras event-level quando disponíveis; (iv) verificação de sensibilidade a janelas de reamostragem menores.

V. CONCLUSÃO

Este estudo aplica o arcabouço de identificação estrutural via heterocedasticidade, originalmente explorado por Takahashi (2025), ao mercado futuro de S&P 500, com dados dos contratos ESH2, ESM2, ESU2 e ESZ2. A estimativa em janelas intradiárias de quinze minutos permite capturar de forma precisa a relação contemporânea entre retornos e desequilíbrios de fluxo (OFI), bem como documentar variações sistemáticas ao longo do pregão e em torno de anúncios macroeconômicos. Os resultados revelam padrões consistentes da microestrutura: maior impacto na abertura e no fechamento,

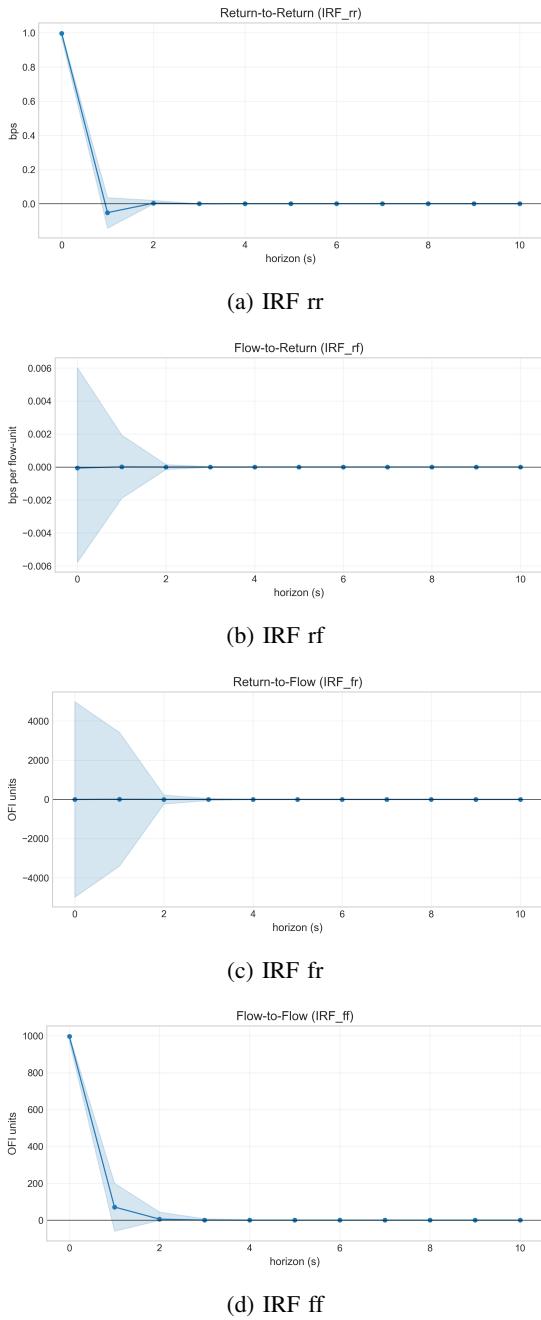


Fig. 4: IRFs estruturais medianas (faixas 5–95%). Painéis: (a) rr, (b) rf, (c) fr, (d) ff.

queda substancial no período de maior liquidez e forte assimetria entre os ajustes de preço e de fluxo na janela do anúncio.

Uma limitação importante do estudo decorre do uso de dados BBO agregados a cada segundo, em contraste com bases do tipo *Market-by-Price* (MBP) ou *Market-by-Level* (MBL), que preservam toda a profundidade visível do livro em nível de evento. Enquanto o BBO captura somente o melhor bid e ask, o MBP registra o volume disponível em cada nível de preço, permitindo mensurar profundidade verdadeira, elasticidade do livro e estrutura de liquidez além do primeiro nível.

Consequentemente, embora a reamostragem a 1 segundo seja suficiente para identificar relações contemporâneas e padrões intradiários amplos, ela comprime a sequência de eventos e reduz a capacidade de isolar mecanismos de propagação intrassegundo, especialmente durante períodos de alta intensidade de mensagens.

Apesar dessas limitações, o exercício confirma que a identificação via ITH permanece operacional mesmo com dados BBO agregados, produzindo estimativas plausíveis e coerentes com regularidades conhecidas da microestrutura. Ainda assim, os resultados devem ser interpretados com cautela: a combinação de agregação temporal a 1 segundo e ausência de profundidade para além do primeiro nível reduz a granularidade da análise e pode atenuar mecanismos que operam em horizontes intrassegundo. Extensões naturais incluem a replicação do arcabouço com dados MBP/MBL baseados em eventos, modelos multivariados que incorporem spreads, profundidade e intensidade, e a análise conjunta com ETFs, opções ou equity cash. Tais extensões poderiam oferecer uma caracterização mais completa da microestrutura, especialmente em condições de liquidez variável, mas representam desenvolvimentos além do alcance e da ambição empírica deste estudo.

REFERENCES

- [1] R. Rigobon, “Identification through heteroskedasticity,” 2003.
- [2] R. Cont et al., “The Price Impact of Order Book Events,” 2014.
- [3] T. Andersen et al., “Real-Time Price Discovery in Stock, Bond and FX Markets,” 2007.
- [4] H. Takahashi, “Return and Order Flow Imbalances: Intraday Dynamics and Macroeconomic News Effects,” 2025.