Price Discovery in High-Frequency Equity Markets: Evidence from Retail and Institutional Trades

Jan Harren
University of Münster, Germany

1. Story

Summary

Using high-frequency trades and quotes (TAQ) data, I quantify the information content of retail and institutional trades in equity markets. I find evidence of a heterogenous price impact among retailers and institutionals. Consistent with theory, I show that information frictions, illiquidity, and information drive differences in the price impact of retail and institutional investors. A size-neutral trading strategy on institutional investors' price impact yields sizeable returns, beats the market, and is not explained by established risk factors. Furthermore, I find that episodes of coordinated trading by Robinhood investors reduce the price impact of institutional investors. This is consistent with indirect liquidity provision from retailers to institutionals via wholesalers due to internalization of retail trades.

Motivation

- January 2021: GME stock showed that **retailers** can act as an "angry mob" and move markets in unfavorable directions for **institutional** traders
- Different market participants
- Retailers: Information via r/wallstreetbets → e.g. GME
- Institutionals: Information via huge research departments
- However: Wholesalers' internalization of retail trades prevents direct interactions between both market participants
- Wholesalers' **provide liquidity** from one group to the other, especially when stock liquidity is limited (i.e. when price impact of trades is high)
- High price impact proxies these indirect trading costs (Barardehi et al., 2021)

This paper

- *How does retail activity affect the price impact/trading costs of institutionals?*

Questions

- Is there a heterogenous price impact of retailers/institutionals?
- What is the **economic intuition** behind the price impact?
- For which participants is the price impact (trading costs) **priced** in the cross-section?
- Do retailers **provide liquidity** to institutionals and reduce their trading costs?

What I do

- I measure permanent price impact of retailers' and institutionals' trades (Hasbrouck, 1991; Ranaldo and Somogyi, 2021)
- I show that there is a risk premium for institutionals price impact, which vanishes with increased retail activity in 2018
- I show that retailers reduce the trading costs of institutions by indirect liquidity provision via wholesalers'

What I find

- Retailer and institutional trades have a permanent **price impact** on quote changes, where institutionals impact exceeds that of retailers
- Retailers and instis have the largest price impact on **illiquid stocks** with large **informational frictions** (stocks with higher limits to arbitrage and level of riskiness) → Confirms price impact as measure of trading costs
- A size-neutral **trading strategy** on institutionals price impact reveals significant returns and SR after transaction costs until 2018
- Diff-in-diff matching: High directional trading on Robinhood by retailers reduces the price impact (and hence trading costs) of institutions → **liquidity provision** by retailers to institutions via internalization of wholesaler trades

2. Data and Methodology

VAR approach

Methodology → Aggregate trades on 5 min intervals

- z_t^j = Signed net volume = $\text{buy}_t^j - \text{sell}_t^j$ with $j \in C = \{IN, RE\}$ → Boehmer et al. (2021)
- $T_t^j = 1$ if $z_t^j > 0$ and $T_t^j = -1$ if $z_t^j < 0$ and $T_t^j = 0$ otherwise
- Price impact should capture illiquidity premia/informed trading
- Control for log size of net order flow that is orthogonal to T_t^j

We follow Hasbrouck (1991) and Ranaldo & Somogyi (2021) and estimate for each stock k

$$r_t = \sum_{i=1}^{10} \rho_i r_{t-i} + \sum_{j \in C} \left(\sum_{i=0}^{10} \beta_i^j T_{t-i}^j + \sum_{i=0}^{10} \phi_i^j \tilde{S}_{t-i}^j \right) + \zeta_{1,l} D_{l,t} + \epsilon_{r,t} \tag{1}$$

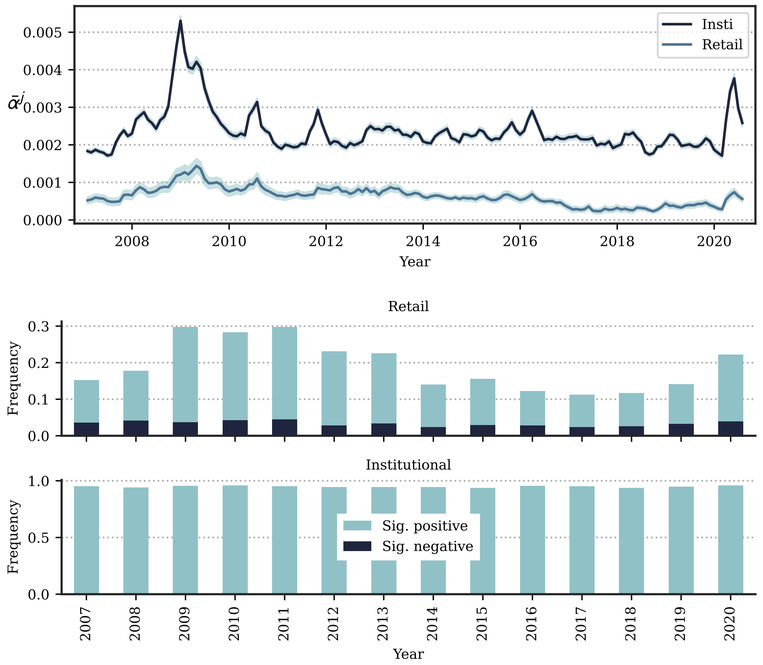
$$T_t = \sum_{i=1}^{10} \gamma_i r_{t-i} + \sum_{j \in C} \left(\sum_{i=1}^{10} \delta_i^j T_{t-i}^j + \sum_{i=1}^{10} \omega_i^j \tilde{S}_{t-i}^j \right) + \zeta_{2,l} D_{l,t} + \epsilon_{T,t}, \tag{2}$$

and calculate the permanent price impact (expected asymmetric/private information) as

$$\alpha_{10}^{j,k} = \sum_{t=0}^{10} \beta_t^{j,k}, \tag{3}$$

where $\epsilon_{r,t}$ and $\epsilon_{T,t}$ are unexpected public and private information respectively.

Permanent price impact of instis and retailers



Data

TAQ

- Trades and Quotes (TAQ) data → all transactions for stocks listed on US-exchanges

CRSP

- Monthly stock files from 2006 to 2020
- Common shares with share code 10 and 11 and standard exchanges (NYSE, AMEX, NASDAQ)
- Filters: $P_t < 5$ and exclude smallest quintile of the cross-sectional distribution each month

Robinhood → Robintrack website

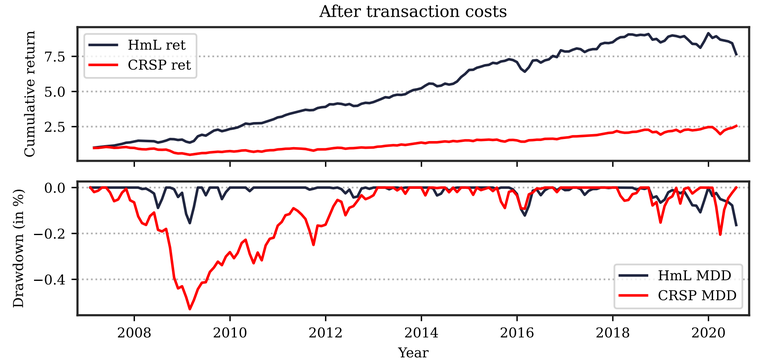
- Website scraped hourly user holdings for all equities on Robinhood
- API was active from May 5, 2018, to August 13, 2020

3. Empirical Results

Inf. frictions/Trading costs explain price impact

	α_{10}^{IN}					α_{10}^{RE}				
(in bps)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
α_{10}^{IN}						1.20***	1.12***	1.28***	0.96***	0.76***
α_{10}^{RE}	1.58***	1.65***	1.30***	1.38***	0.72***					
Age	−1.02					−1.57***				
Illiq	7.79					43.09***				
IVOL	−0.04					0.14				
Size	−22.59***					−0.70***				
Analyst Cov.	−2.04***					−0.05				
Insti Own.	−2.58***					−0.57***				
Inf. Friction		10.80***		9.75***			1.17***		0.96***	
$\ln(VPIN_t)$			4.80***	3.24***				1.16***	0.98***	
ES_t					10.22***					1.39
QS_t					−1.23**					1.24*
Entity effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Obs.	121015	121015	126427	121002	126426	121015	121015	126427	121002	126426
within R^2 [%]	18.8	20.8	7.4	22.7	20.6	-1.6	3.5	2.5	4.0	3.9

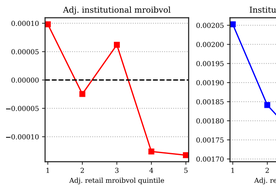
Profitable size-neutral trading strategy



- Significant value-weighted returns with Sharpe ratio of 1.66 (1.30) before (after) transaction costs compared to market with 0.56 (0.52)
- With rising retail activity: No superior returns for trading on institutional price impact

Retailers provide liquidity to institutionals

$$\Delta\alpha_{i,t}^{IN,10} = \hat{\beta} \cdot \mathbb{1}\left(\Delta w_{i,t}^{ARH} - w_{i,t-1}^{ARH} > 30\right)_{i,t} + \hat{e}_{i,t}. \tag{4}$$



	Δw_t							
	α_{10}^{IN}							
in (%)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}_{\Delta 30}$	−0.02	−0.01	−0.02	−0.01	−0.01	−0.02	−0.02	−0.02
	[−4.56]	[−2.08]	[−4.45]	[−2.50]	[−2.35]	[−4.08]	[−2.96]	[−3.58]
Obs.	17731	17731	17730	17730	17730	17731	17731	17730
R^2 [%]	0.5	0.1	0.2	0.1	0.1	0.2	0.2	0.2
ln(Size)		X		X	X		X	
Ret.			X	X	X			X
Std.					X	X	X	X
Entity effects	Y	Y	Y	Y	Y	Y	Y	Y
Time effects	Y	Y	Y	Y	Y	Y	Y	Y

- Barardehi et al. (2021): $|Mroib|$ → good proxy for wholesalers' liquidity provision intensity
- $Mroib$ and institutional ANcerno trade imbalances are negatively correlated → same here
- Higher retail activity reduces the price impact of institutionals → **Retailers provide liquidity to institutions** when liquidity is scarce

- Contact of corresponding author: jan.harren@wiwi.uni-muenster.de