Informed Trading under the Microscope: Evidence from 30 Years of Daily Hedge Fund Trades

JinGi Ha[1] Jianfeng Hu[2] Yuehua Tang[3]

[1]Soongsil University
[2]Singapore Management University, Heritage Capital Management
[3]University of Florida

June, 2026

U Maryland / SMU / UBS Quant Investment Forum: AI and Finance

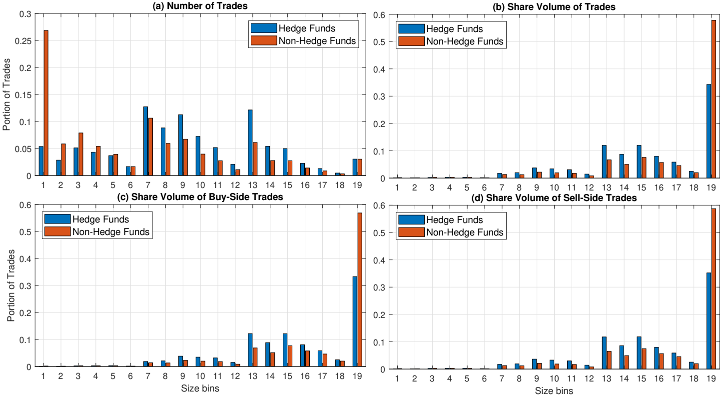
Summary and Contributions

- We develop a method that allows investigation of hedge fund behavior at daily or higher frequency using public data
- Huang, Tan, Wermers (2020) find institutional trading follows news arrival. We find HF trades *before* news predict event returns.
- We pin down a source of hedge fund information edge – the use of alternative data, a new effect of the data related to Zhu (2019), Katona, et al. (2024), Dessaint, et al. (2024), and Bonelli and Foucault (2024).
- We provide high-frequency evidence of how hedge fund trades contribute to price discovery and improve price efficiency (Boehmer and Kelly (2009), Akbas, et al. (2015), and Cao, et al. (2018)).

Use Trade Size to Estimate Hedge Fund Trades

- Trade size has been used to distinguish institutional and retail trades
- Campbell, Ramadorai, and Schwartz (CRS, 2009) use the whole spectrum of trade size bins instead of a single cutoff:
  1. Estimate relations between quarterly institutional ownership changes in 13F to quarterly order flow across nineteen trade size bins in TAQ
  2. Extrapolate the quarterly relation using daily TAQ order flow
- If HF and NHF exhibit different trading styles concerning order sizes, we can distinguish HF and NHF trades using a similar method to CRS.

HF and NHF Execution Methods (Ancerno Sample)



- NHF place more small orders concerning slicing and dicing
- HF execute more trades and volume in middle sized bins

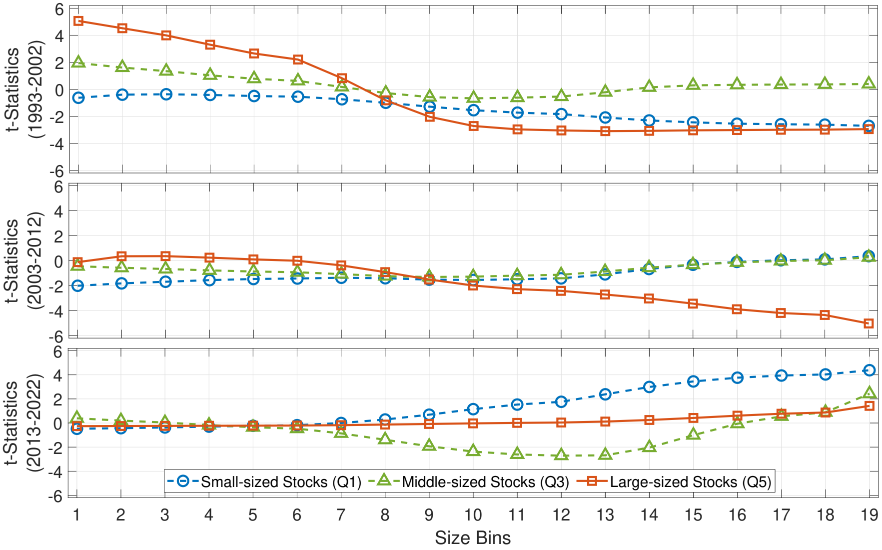
Estimation of Relation in Quarterly Data

We first estimate the quarterly relation between changes in hedge fund ownership (13F) and order imbalances across trade size bins (TAQ):

$$\Delta Y_{i,q} = \alpha_q + \rho \Delta Y_{i,q-1} + \varphi Y_{i,q-1} + \beta_U U_{i,q} + \sum_{Z=1}^{19} \beta(Z, Y_{i,q-1}) F_{Z,i,q} + \epsilon_{i,q}$$

- $\Delta Y_{i,q}$: Change HF/NHF ownership for stock i in quarter q.
- $F_{Z,i,q}$: Order imbalance in trade size bin Z.
- $\beta(Z, Y_{i,q-1})$: Coefficient for each trade size bin Z, modeled using a yield curve function.
- $U_{i,q}$: Unclassified trades scaled by shares outstanding.

Difference of Trade-Size Sensitivities $\left(\beta^{HF}(Z,v) - \beta^{NHF}(Z,v)\right)$



Summary Statistics of Fitted Daily HF and NHF

Period	Full	Early	Middle	Late
Mean HF	0.009	0.005	0.012	0.010
S.D. HF	0.025	0.015	0.025	0.036
Mean NHF	0.024	0.023	0.032	0.014
S.D. NHF	0.072	0.063	0.079	0.074
Correlation (HF, NHF)	0.671	0.547	0.668	0.798

- HF is smaller than NHF.
- HF and NHF become more similar over time.

What Affects Institutional Trading?

	HF_m	NHF_m
$MktBeta_m$	0.020***	0.048***
	(7.81)	(9.18)
$MktCap_m$	0.006***	0.094***
	(2.71)	(8.01)
MB_m	-0.067***	-0.007
	(-2.60)	(-0.09)
REV_m	-0.062***	-0.063***
	(-7.03)	(-3.30)
MOM_m	-0.036***	-0.051***
	(-6.39)	(-4.74)
$SPRD_m$	-3.128***	-6.265***
	(-5.11)	(-4.61)
$TURN_m$	0.079***	0.250***
	(22.11)	(11.43)
$IdioRisk_m$	0.180***	-0.080
	(5.14)	(-1.29)
MISP excl. MOM_m	0.016**	-0.017
	(2.29)	(-0.73)

- At the monthly frequency, HF and NHF respond to firm characteristics in similar ways.

Trading on Corporate Events and News

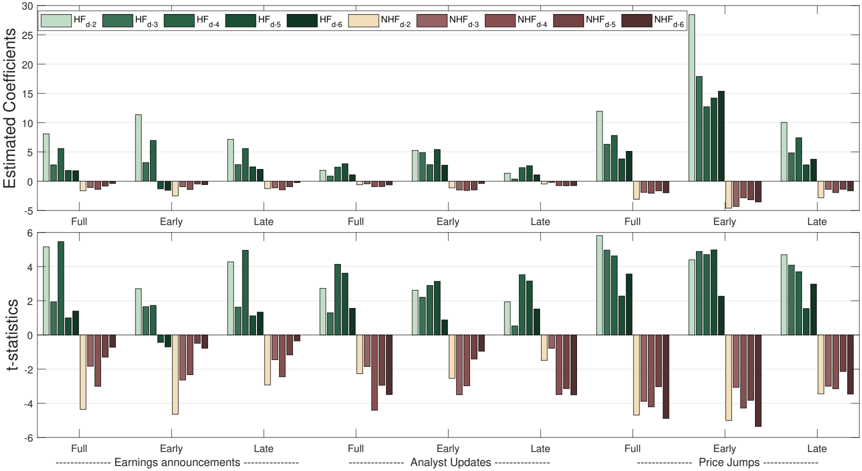
- How do hedge funds react to corporate events at a higher frequency?

$$\mathrm{CAR}_{i,t-1,t+1} = \alpha_i + \alpha_y + \sum_{k=2}^{6} \beta_k^{HF} \mathrm{HF}_{i,t-k} + \sum_{k=2}^{6} \beta_k^{NHF} \mathrm{NHF}_{i,t-k} + \mathrm{Control}_{i,t} + \epsilon_{i,t},$$

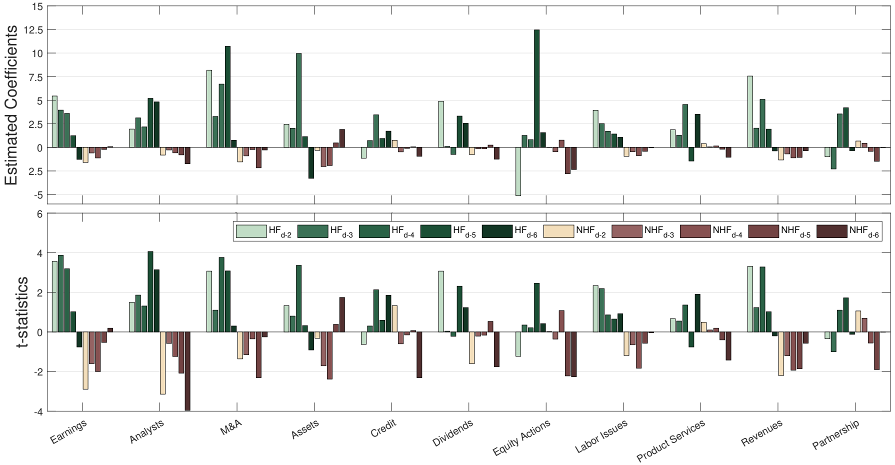
- Events include earnings announcements, analyst rating upgrade, and price jumps. We also use corporate news covered by RavenPack including both fundamental and non-fundamental news.

- In all the events, HF positively predicts the subsequent abnormal return and NHF does not.

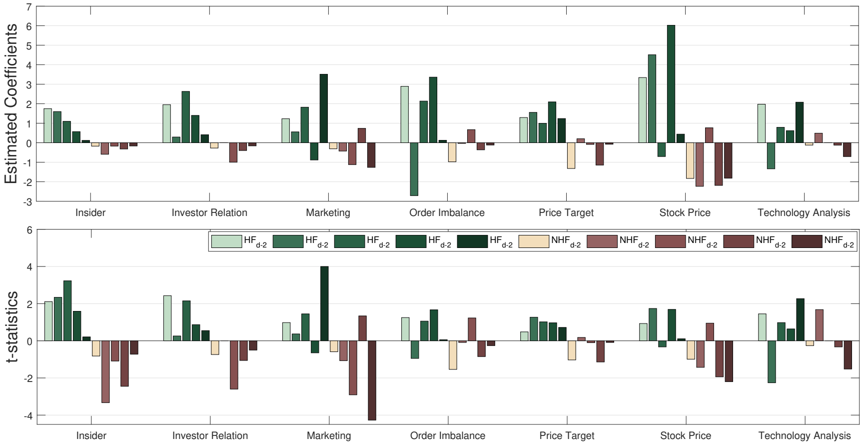
Corporate Event Results



RavenPack Fundamental News Results

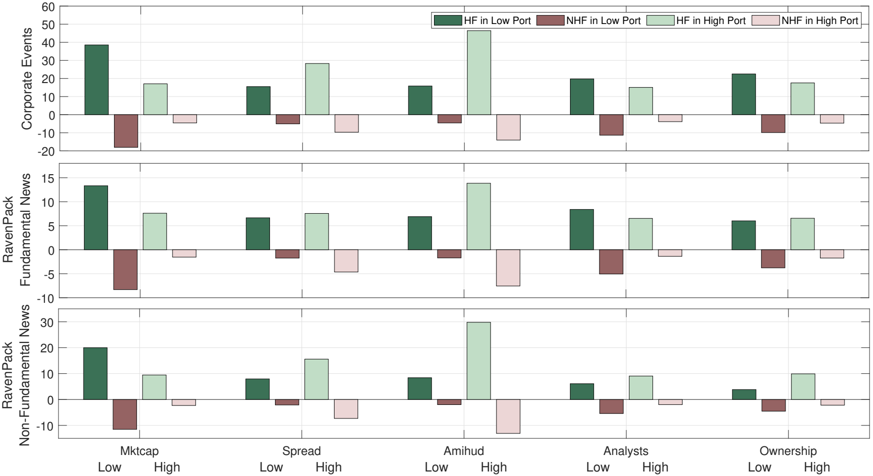


RavenPack Non-Fundamental News Results



Cross-sectional Variation in Predicting Event Returns

- The estimated effects are stronger for firms with high information asymmetry.



The Information Edge: Alternative Data

- RS Metrics provides state-level parking lot data through satellite images.
- Sample includes 48 covered stocks matched with three non-covered stocks from the same industry, 3 years before and after initiation of coverage.
- Difference-in-differences model with firm and year fixed effects:

$$\text{Return}_{i,d} = \alpha_i + \alpha_y + \beta_1 \text{TREAT} \times \text{POST} \times \text{HF}_{d-1} + \ldots + \epsilon_{i,d}$$

  - TREAT is a dummy for stocks covered by RS Metrics
  - POST indicates the post-data period
  - Controls for firm characteristics, including market capitalization and other trading variables.

DiD Test Results

	(1)	(2)
TREAT × POST × HF_{d-1}	0.319***	0.305***
	(2.96)	(2.67)
TREAT × POST × NHF_{d-1}	-0.532	-0.571
	(-1.45)	(-1.49)
TREAT × POST × TOF_{d-1}	-0.151	-0.189
	(-0.69)	(-0.85)
TREAT × HF_{d-1}	-0.027	-0.070
	(-0.05)	(-0.14)
POST × HF_{d-1}	-0.540	-0.567
	(-0.71)	(-0.73)
HF_{d-1}	0.459	0.515
	(0.69)	(0.77)
Control	No	Yes

	HF_w	NHF_w
$Traffic_{w-1} - AvgTraffic_{q-1}$	7.002***	6.483
	[3.05]	[0.86]
HF_{w-1}	0.261	0.202
	[0.92]	[0.36]
NHF_{w-1}	0.024	0.241*
	[0.73]	[1.91]
TOI_{w-1}	0.014	0.046
	[0.89]	[1.37]
$Amihud_{w-1}$	-2.128	-17.443
	[-0.19]	[-0.79]
$SPRd_{w-1}$	1.601	-6.445
	[0.05]	[-0.08]
NAT_{q-1}	-0.217	-0.258
	[-0.86]	[-0.46]
$ARETQ_{w-1}$	0.000	0.002**
	[0.84]	[2.10]

- HF increases purchase after satellite images show abnormal increases in traffic outside stores.

Out-of-Sample Return Predictability

- We reestimate the models every year using ten years of data and calculate fitted HF and NHF for the next year.

	2003-2022	2003-2012	2013-2022
OOS HF_{d-1}	0.798***	0.839***	0.757***
	(9.47)	(5.96)	(8.38)
OOS HF_{d-2}	0.118	-0.044	0.280***
	(1.52)	(-0.33)	(3.75)
OOS HF_{d-3}	0.178**	0.275**	0.080
	(2.29)	(2.04)	(1.06)
OOS HF_{d-4}	0.151*	0.176	0.126*
	(1.93)	(1.28)	(1.70)
OOS HF_{d-5}	0.062	0.020	0.105
	(0.88)	(0.16)	(1.51)
Adjusted R^2	0.036	0.028	0.043

HF and Market Efficiency: Earnings Announcements

(1) y = $CAR_{d+1,d+61}$		(2) y = Jump Ratio Rank		(3) y = ΔVR	
HF_d	-7.980**	$\lvert HF\rvert_{d-21,d-1}$	-1.042***	$\lvert HF\rvert_d$	-2.626**
	(-2.16)		(-10.16)		(-2.41)
NHF_d	3.183**	$\lvert NHF\rvert_{d-21,d-1}$	-0.654***	$\lvert NHF\rvert_d$	-0.565
	(2.03)		(-13.36)		(-1.17)
TOF_d	-0.490	$\lvert TOF\rvert_{d-21,d-1}$	1.305***	$\lvert TOF\rvert_d$	-0.422**
	(-0.74)		(34.31)		(-2.52)

- HF increases market efficiency by reducing the magnitude of the market response to information shocks at earnings announcements.

Fama and MacBeth Regressions of Variance Ratio

	1993-2022	1993-2002	2003-2012	2013-2022
$\lvert\mathrm{HF}\rvert_{d-1}$	-3.301***	-5.773***	-1.539***	-2.598***
	(-5.71)	(-3.60)	(-3.26)	(-6.08)
$\lvert\mathrm{NHF}\rvert_{d-1}$	-0.108	0.658**	-0.300	-0.680***
	(-0.75)	(1.98)	(-1.55)	(-3.50)
$\lvert\mathrm{TOF}\rvert_{d-1}$	-0.568***	0.038	-0.258***	-1.482***
	(-8.46)	(0.26)	(-2.96)	(-20.14)
VR_{d-3}	-2.549***	-1.984***	-2.359***	-3.301***
	(-42.55)	(-16.02)	(-28.58)	(-43.05)
$\lvert\mathrm{NAT}\rvert_{q-1}$	0.880***	1.904***	0.100	0.638***
	(3.80)	(2.81)	(0.63)	(4.47)

- HF significantly reduces the variance ratio, making the stock price more efficient.

Conclusion

- Hedge fund order flows have predictive power, especially when utilizing alternative data sources like satellite imagery, contributing to more accurate stock price adjustments.
- Hedge funds play a significant role in price discovery and improving market efficiency, particularly around major corporate events such as earnings announcements and news arrivals.
- The method also allows future research to further investigate into the role of hedge fund trading across different market conditions.