"Information Acquisition and Usage of Retail Investors: Evidence from Web Views and Watchlists"
by Yuecheng Jia, Shu Yan, Hongyu Zhang

Discussion by Ahmed Guecioueur (INSEAD)

FMA Annual Meeting

23 October 2021

Summary

Goal

- Studying information acquisition by **retail** investors:

  - Which stocks to acquire information on?

  - How is the information used?

- Chinese setting
  - **Hithink Royal Flush Information Network** is the main data provider for the paper
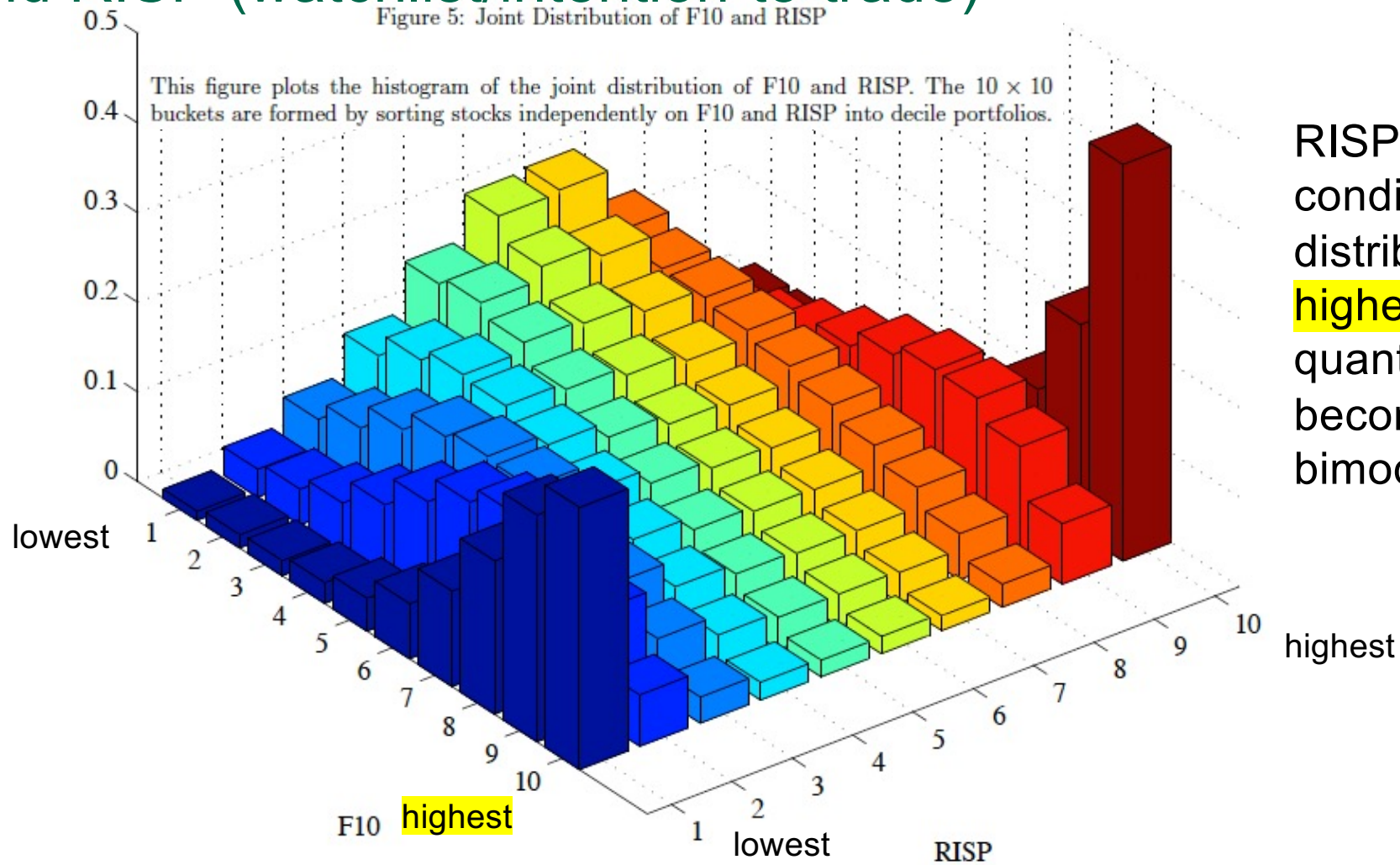
Key selling point of the paper (in my opinion)

- Stock-level proxies for both stages of a retail investor's decision:

    - Information acquisition (**F10**): visits to company hub using the Hithink software

    - Intention-to-purchase (**RISP**): additions to user watchlists

- In the paper:

    - Determinants of investor decision rules for each stage

    - Joint distribution of F10 + RISP

Joint distribution of F10 (info. acquisition) and RISP (watchlist/intention-to-trade)
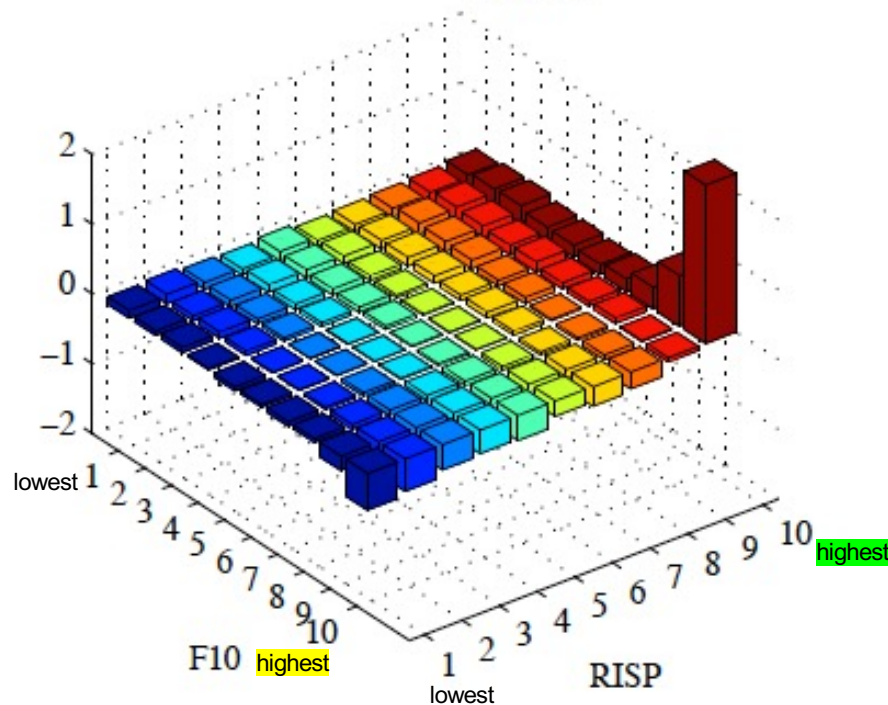

Figure 5: Joint Distribution of F10 and RISP

This figure plots the histogram of the joint distribution of F10 and RISP. The 10 × 10 buckets are formed by sorting stocks independently on F10 and RISP into decile portfolios.
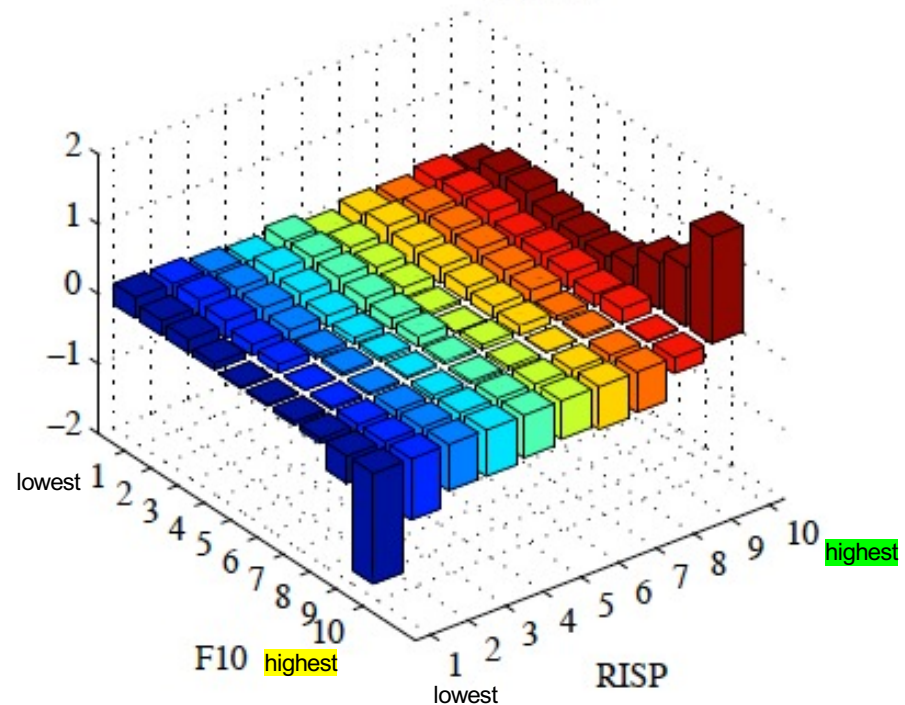
RISP conditional distribution at highest F10 quantiles becomes bimodal

EW *ex post* next-day and next-week returns

Panel A: $R_{t+1,t+2}$ — Panel B: $R_{t+1,t+5}$

- **High F10** stocks mostly earn negative returns
- Except for **High F10** + **High RISP** stocks
- All **High F10** turn negative (i.e. return reversal) at longer horizons

Retail investor decision rules *ex ante* for stocks with high…

F10 (information acquisition)

- Larger market cap
- Higher turnover
- Lower B/M
- i.e. attention-grabbing / "glittering"

RISP (intention-to-buy)

- Low idiosyncratic volatility
- Less complicated
- i.e. easier to value
- Lower turnover

High F10+RISP stocks (with +ve short-term returns) have high *future* news sentiment ⇒ skilled retail invs.

Table 8: Future News Sentiment and F10–RISP Interactions

This table reports estimation results of Fama–MacBeth regressions of future news sentiment on F10, RISP, and their interactions. The dependent variable is future news sentiment. The independent variables are defined as following. We sort stocks on F10 and RISP independently. Each stock has two rankings: r_{F10} for F10 and r_{RISP} for RISP. We normalize the rankings so that they are between 0 and 1. RD is defined as difference of the two rankings: $RD = r_{RISP} - r_{F10}$. Consensus is the dummy variable which equals to 1 if a stock is in the top F10 decile and top RISP decile, and zero otherwise. Deviation is the dummy variable which equals to 1 if a stock is in the top F10 decile and the bottom RISP decile, and zero otherwise. For

Panel A: Content Sentiment

	(1)	(2)	(3)	(4)	(5)	(6)
F10	-0.296***	-0.265***	-0.295***	-0.269***	-0.270***	-0.245***
	[-13.53]	[-10.19]	[-11.85]	[-9.25]	[-12.60]	[-8.68]
RISP	0.468***	0.722***	0.457***	0.738***	0.408***	0.632***
	[6.24]	[7.23]	[5.25]	[6.63]	[5.13]	[5.11]
Consensus		0.112***		0.108***		0.099**
		[2.76]		[2.89]		[2.51]
Deviation			0.314	-0.023		0.111
			[0.34]	[-0.25]		[0.12]
RD					0.001***	0.001***
					[2.72]	[2.67]
Adj. R^2	0.057	0.059	0.045	0.061	0.061	0.067

Panel B: Overall Sentiment

	(1)	(2)	(3)	(4)	(5)	(6)
F10	-0.256***	-0.316***	-0.232***	-0.296***	-0.235***	-0.183***
	[-11.57]	[-11.52]	[-9.21]	[-9.73]	[-10.46]	[-6.17]
RISP	0.319***	0.629***	0.385***	0.708***	0.349***	0.577***
	[4.92]	[6.00]	[5.14]	[6.15]	[4.12]	[4.18]
Consensus		0.111***		0.117***		0.128***
		[3.70]		[3.95]		[2.91]
Deviation			0.433*	0.427		0.091
			[1.69]	[1.24]		[0.38]
RD					0.001***	0.001**
					[2.88]	[2.19]
Adj. R^2	0.062	0.069	0.066	0.069	0.063	0.070

Panel C: Title Sentiment

	(1)	(2)	(3)	(4)	(5)	(6)
F10	-0.218***	-0.256***	-0.196***	-0.236***	-0.178***	-0.193**
	[-6.62]	[-6.20]	[-5.30]	[-5.20]	[-5.04]	[-2.43]
RISP	0.414***	0.652***	0.549***	0.702***	0.602***	0.479**
	[3.97]	[3.91]	[3.92]	[3.54]	[3.79]	[2.20]
Consensus		0.724		0.740		0.491
		[1.51]		[1.54]		[1.00]
Deviation			0.265	0.309		0.284
			[1.32]	[1.52]		[1.57]
RD					0.002	0.002
					[0.12]	[0.19]
Adj. R^2	0.058	0.062	0.062	0.066	0.055	0.061

Panel D: Reference Sentiment

	(1)	(2)	(3)	(4)	(5)	(6)
F10	-0.208***	-0.195***	-0.212***	-0.186***	-0.215***	-0.184***
	[-10.94]	[-6.49]	[-7.58]	[-5.57]	[-8.81]	[-5.62]
RISP	0.343***	0.588***	0.295***	0.559***	0.348***	0.565***
	[3.88]	[5.10]	[3.02]	[4.53]	[3.73]	[4.05]
Consensus		0.102**		0.945***		0.978**
		[2.36]		[2.42]		[2.36]
Deviation			0.271	-0.236		-0.223
			[0.27]	[-0.23]		[-0.21]
RD					0.122**	0.072*
					[2.45]	[1.78]
Adj. R^2	0.054	0.056	0.057	0.059	0.058	0.063

Comments

Platform & population

- Are the statistics provided by Hithink for desktop users only, or do they include mobile users?
  - Paper says the software is available on multiple platforms
  - Information display effects on different platforms/formats? See Hong, Lu & Pan's (2019) work on FinTech platform effects

- Any general demographics for the Hithink users and how they compare to the overall retail trading population in China?

- Could add English captions to Figures 1-2

News: past, present, future

- What is the **horizon** over which future news is calculated?
  - Story is that investors can correctly identify stocks with +ve **next-day** and **week-ahead** returns, potentially private info.

- Is there **persistence** in news volume and sentiment? Is "Consensus" (i.e. High F10 + High RISP) associated with contemporaneous or past positive-sentiment news?
  - **Alternative story**: investors learn (F10) about stocks with lots of past/contemporaneous news, and trade (RISP) in the same direction as the sentiment of this news
  - Can you rule that out? Maybe add as control

RISP as a measure of trading decisions

- *"A high value of RISP for a stock indicates a high level of intention to **purchase** the stock."*

  - Makes sense that it **includes** intentions-to-purchase

  - But maybe also captures **other retail investor decisions**: for example, investors learning how to time their stock sales may add **recently-sold** stocks to their watchlists to continue to follow them

  - Can you rule this out by checking that a stock's RISP does not increase when there is a relatively high amount of retail selling? (BSI should not capture this)

Wish list for more Hithink data

- Is it possible to get **investor-level** (or IP address-level) F10 and RISP measures?

  - Would be very interesting to describe what & how many stocks individual investors intend to trade

  - i.e. how sparse and persistent are their **choice sets**?

  - Related to Ralph Koijen's RFS editorial about understanding the extensive and intensive margin of **demand**

  - Can also help understand investors' **preferred habitats**.

Wrapping up

Conclusion

- Unique stock-level measures of the 2 stages to retail investors' trading decisions

- Empirical evidence on the associated decision rules

- Retail investors can be skilled at making use of their information

Thank you!

Bibliography

- Brunnermeier, M., Farhi, E., Koijen, R. S., Krishnamurthy, A., Ludvigson, S. C., Lustig, H., Nagel, S. & Piazzesi, M. (2021). Perspectives on the Future of Asset Pricing. *Review of Financial Studies*.

- Hong, C. Y., Lu, X., & Pan, J. (2019). FinTech Platforms and Mutual Fund Distribution (No. w26576). National Bureau of Economic Research.

- Jia, Y., Yan, S., & Zhang, H. (2021). Information Acquisition and Usage of Retail Investors: Evidence from Web Views and Watchlists.