Resolving the Excessive Trading Puzzle:
An Integrated Approach Based on Surveys and Transactions

CEPR Fifth European Workshop on Household Finance

Hongqi Liu, Cameron Peng, Wei A. Xiong, and Wei Xiong

CUHK-SZ, LSE, Shenzhen Stock Exchange, and Princeton

The “Excessive Trading” Puzzle

- Behavioral finance has made significant advancement over the last few decades
  - by offering sharp insights on a wide range of anomalies in financial markets

- A byproduct:
  - multiple (perhaps too many) behavioral biases for each single anomaly

- As an example, consider the so-called “excessive trading” puzzle
  - retail investors appear to be trading too much (Odean 1999; Barber and Odean 2000)
    1. before fees: return lower than the market index
    2. transaction cost makes performance even worse
    3. those who trade the most often perform the worst
  - many behavioral explanations have been proposed
    - overconfidence, realization utility, gambling preference, sensation seeking, social interaction, and low financial literacy, …

The “Excessive Trading” Puzzle, cont’d

- The large set of behavioral explanations we face is not satisfying
  - unlikely that all the biases are equally important
  - possible that certain biases would be subsumed by others
  - a comparison seems necessary

- More generally, it is important to consolidate the multiple explanations for each anomaly
  - eventually develop a unified conceptual framework

Challenges of Consolidation

- Many existing explanations, by design, share similar predictions on the targeted anomaly
  - they may offer different predictions on more subtle dimensions, but the power is constrained by the availability of administrative data
  - it is even harder to run horse races among multiple explanations

- Recent literature has turned to survey-based approaches
  - elicit investors' own perspectives on the drivers of their trading and investment decisions
    - e.g., Greenwood and Shleifer (2014), Choi and Robertson (2019), and Chinco, Hartzmark and Sussman (2019)
  - advantages:
    - collect information for many mechanisms quickly
    - permit horse races among different explanations
  - concerns:
    - respondents may not truthfully report their answers
    - survey responses may not translate into real actions

This Paper

- Adopt a new approach to address the excessive trading puzzle
  - by combining information from two different sources: surveys and transactions
  - overcome the challenges faced by existing approaches

- A nation-wide survey among Chinese retail investors
  - more than 10,000 individuals randomized across provinces, brokerages, and branches
  - questions designed to measure an exhaustive list of trading motives
    - financial literacy, beliefs, preferences, information sources, other trading motives
  - a serious comparison among competing explanations for trading volume

- Merge survey responses with account-level transaction data at the Shenzhen Stock Exchange
  - survey responses are largely consistent with trading behavior
  - e.g., gambling preference → buy lottery-like stocks
  - justification to the use of surveys

- Two sets of exercises
  - a **horse race** among survey-based measures of trading motives
  - a **comparison** between survey-based and transaction-based measures

Main Findings

- Two trading motives stand out in the horse race: *gambling preferences* and *belief in having information advantage*
  - gambling preference: 21%; perceived information advantage: 24% (s.d. of turnover is 126%)
- Additional evidence further supports these two trading motives
  - gambling preference: trade smaller, high-beta, more volatile, and more positively skewed stocks
  - information advantage: no better returns → *overconfidence* about information advantage
- Certain explanations are indeed subsumed by others
  - e.g., sensation seeking is significant in univariate regressions, but not in multivariate regressions
- For a given bias with multiple forms, they don't have the same explanatory power
  - e.g., out of the three forms of overconfidence, overconfidence about having information advantage works well while miscalibration of uncertainty works poorly
- Popular arguments such as neglect of trading cost, low financial literacy, and social influence do **not** contribute to higher volume

Main Findings, cont'd

- We construct an alternative measure for gambling preference based on transactions
  - called "gambling *behavior*", measured by the propensity to buy lottery-like stocks
  - *more* powerful in explaining turnover, but *correlated* with other trading motives

- Our analysis highlights the tradeoffs between survey-based and transaction-based approaches
  - survey-based approach:
    - pros: a direct measure for each trading motive, allowing for collecting many trading motives at once
    - cons: subject to measurement noise at the individual level and are thus less powerful
  - transaction-based approach:
    - pros: more powerful in explaining trading volume
    - cons: simultaneously capture multiple trading motives and less reliable in isolating a single mechanism

The Survey

- Investor Education Center at the Shenzhen Stock Exchange (SZSE)
  - time: September 2018
  - randomized across branches of China's 10 largest brokers
  - 500 branches in total, each with a target size of 20 investors

- To boost response rate
  - logos of SZSE and Shenzhen Finance Institute
  - confidentiality agreement
  - monetary rewards

- Four parts
  1. Financial literacy
  2. Trading motives
  3. Demographics
  4. "Nudge" experiment: see the paper

More on Part 2: Trading Motives

- For each motive, we phrase the questions to map closely to the underlying concept
  - by going back to the original paper proposing that particular motive
- A motive may have different forms of representation
  - in such cases, we include at least one question for each form
- To ensure the quality of survey responses, we design all questions to be multiple-choice
  - qualitative questions
    - statement: “strongly agree”, “agree”, “neutral”, “disagree”, and “strongly disagree”
    - frequency: “always”, “often”, “sometimes”, “rarely”, and “never”
  - quantitative questions
    - each option covers a fixed range of value
- To facilitate the horse race, we encode all survey-based trading motives to dummy variables

Overview of Survey-based Trading Motives

Trading Motive	Measures
Overconfidence	• over-placement (performance and financial literacy) • mis-calibration of uncertainty
Neglect of trading cost	• underestimation of transaction fee • frequency of considering cost • lack of consideration for bid-ask spread
Gambling preference	• with probability weighting • without probability weighting
Sensation seeking	• novelty seeking • volatility seeking
Realization utility	• selling winners • holding losers
Extrapolation	• upward trend • downward trend
Information	• belief in having information advantage (*overconfidence* about own information) • fear for having information disadvantage (*dismissiveness* of others' information)
Social interaction	• family and friends • investment advisors
Others	liquidity needs, portfolio rebalancing needs, risk aversion, optimism/pessimism

Sample Characteristics

- Initial sample size: 12,856
  - drop obs. who spent < 3 min on the survey → 11,268

Gender	Survey	Population
Male	54.00%	71.70%
Female	46.00%	28.30%

Education	Survey	Population
Middle School or below	8.60%	7.30%
High School	15.60%	24.70%
Professional School	21.90%	26.00%
College	44.90%	23.60%
Graduate school and above	9.20%	3.40%

Age	Survey	Population
20 to 30	27.80%	21.30%
30 to 40	29.10%	27.40%
40 to 50	19.90%	24.50%
50 to 60	14.80%	15.10%
>60	8.50%	11.70%

Annual Income	Survey
<20K	3.80%
20K to 100K	17.20%
100K to 200K	29.50%
200K to 500K	29.50%
500K to 1M	12.60%
>1M to 2M	7.50%

Wealth	Survey
<20K	4.80%
20K to 100K	12.30%
100K to 500K	27.50%
500K to 1M	22.30%
1M to 2M	21.90%
2M to 10M	6.50%
10M and above	4.80%

- Bottom-line: a relatively well-educated, wealthy sample

Merging Survey Responses with Trading Data

- Merging process
  - demographic variables: name, date of birth, broker name, and branch name
  - sample size: 11,268 → 6,013
  - positive stock holding in the two-year window before the survey: 6,013 → 4,671

- Summary statistics in the post-survey period (2018:10 to 2019:06)

Panel A: Summary Statistics (monthly)

	P25	Median	P75	Mean	S.D.
Turnover	12.1%	**46.6%**	121.6%	**94.20%**	**125.70%**
Raw returns	-1.8%	0.3%	2.2%	**-0.10%**	3.80%
Net returns	-2.1%	0.1%	2.0%	**-0.30%**	3.80%

Panel B: Correlation Matrix

	Turnover	Raw returns	Net returns
Turnover	1		
Raw returns	**-0.07*** **	1	
Net returns	**-0.16*** **	0.99***	1

Validation: Gambling Preference

- Gambling *behavior*
  - measured as the tendency to buy lottery-like stocks
  - lottery-like stocks: stocks that hit upper price limit in the previous month

Gambling Behavior (2018:01 to 2019:06)		
Gambling preference, with prob. weighting	0.112***	0.109***
	(3.875)	(3.768)
Male		-0.034
		(-1.164)
Controls	NO	YES
R2	0.004	0.023
N	4,145	4,145

  - results are robust to alternative specifications
- other validation tests
  - extrapolation, risk aversion, and return expectations (Giglio et al. 2020)

A Horse-race Among Various Trading Motives

Monthly Turnover in % (2018:10 to 2019:06)

	Univariate	Multivariate
Actual performance in 2017	4.104***	4.198***
	(5.332)	(5.219)
Over-placement, performance	15.695***	11.549**
	(2.760)	(2.063)
Financial literacy score	11.922***	7.065*
	(3.127)	(1.800)
Over-placement, financial literacy	1.729	-2.621
	(0.400)	(-0.625)
Miscalibration	1.116	-2.989
	(0.289)	(-0.764)
Underestimation of transaction cost	-3.549	-3.989
	(-0.980)	(-1.071)
Do not consider transaction cost	-2.143	-4.029
	(-0.548)	(-1.052)
Do not think bid-ask spread is a cost	-15.135***	-9.456***
	(-4.254)	(-2.650)
Extrapolation, up	4.379	-1.255
	(1.110)	(-0.254)
Extrapolation, down	3.810	-1.208
	(1.005)	(-0.262)
Gambling preference, with prob. weighting	10.924***	11.764***
	(2.878)	(2.920)
Gambling preference, without prob. weighting	2.750	-1.159
	(0.684)	(-0.263)

	Univariate	Multivariate
Realization utility, winner	7.188*	7.049*
	(1.874)	(1.848)
Realization utility, loser	0.409	-2.321
	(0.093)	(-0.538)
Sensation seeking, novelty	10.184**	6.598
	(2.270)	(1.360)
Sensation seeking, volatility	11.984***	3.632
	(2.885)	(0.824)
Perceived information advantage	21.747***	15.660***
	(4.254)	(2.988)
Dismissive of others' information	4.778	2.942
	(1.318)	(0.805)
Affected by family and friends	-15.647***	-7.839
	(-3.317)	(-1.616)
Affected by investment advisors	-16.469**	-12.089*
	(-2.708)	(-1.943)
Controls		YES
Male		21.488***
		(6.124)
R2		0.089
N		4,648

Gambling Preference: Magnitude

I aim to select those stocks whose prices would rise sharply in a short period time so that I can get rich quickly

	Panel A: Monthly Turnover (2018:10 to 2019:06)							Panel B: Monthly Raw Returns (2018:10 to 2019:06)	
Gambling preference	P10	P25	P75	P90	Median	Mean		Median	Mean
1. Strongly disagree	0%	4%	99%	206%	25%	74%		0.19%	0.15%
2. Disagree	0%	3%	100%	222%	31%	77%		0.00%	0.04%
3. Neutral	0%	5%	112%	238%	33%	84%		0.01%	0.11%
4. Agree	0%	7%	117%	248%	42%	90%		0.03%	-0.04%
5. Strongly agree	0%	5%	119%	274%	42%	95%		0.00%	-0.20%
DIFF (5−1)	0%	0%	20%	68%	**17%****	**21%****		**-0.19%**	**-0.35%**
Annual transaction fee	0.00%	0.00%	0.60%	1.96%	**0.51%**	**0.63%**	**Net returns**	**0.00%**	**-0.40%**

- trading behavior
  - trade smaller, high-beta, more volatile, and more positively skewed stocks

Information Advantage: Magnitude

How often do you believe that you know the stocks better than others?

	Panel A: Monthly Turnover (2018:10 to 2019:06)							Panel B: Monthly Raw Returns (2018:10 to 2019:06)	
Information Advantage	P10	P25	P75	P90	Median	Mean		Median	Mean
1. Never	0%	4%	102%	232%	30%	76%		0.10%	0.12%
2. Rarely	0%	3%	100%	218%	32%	76%		0.07%	0.06%
3. Sometimes	0%	5%	109%	244%	34%	86%		0.00%	0.08%
4. Often	0%	11%	139%	286%	46%	103%		0.00%	-0.13%
5. Always	0%	10%	139%	253%	44%	100%		0.00%	-0.01%
5−1	0%	6%	37%	21%	**14%****	**24%****		-0.10%	-0.13%
Annual transaction fee	0.00%	0.18%	1.11%	0.63%	**0.42%**	**0.72%**	**Net returns**	**0.00%**	**-0.19%**

- lack of better raw returns: *over*confidence about having information advantage

Remarks on Survey-based Approach

- So far, we have shown that gambling preferences and belief in information advantage are the main drivers for excess trading

- Still, there are concerns associated with survey responses
  - survey responses could be noisy

	Gambling Behavior	
	Around the survey (2018:01 to 2019:06)	
Gambling preference, with prob. weighting	0.112***	0.109***
	(3.875)	(3.768)
Male		-0.034
		(-1.164)
Controls	NO	YES
R2	0.004	0.023
N	4,145	4,145

  - what if we use transaction-based measures directly?

Sorting Investors Based on Gambling Behavior

Gambling *Behavior*

	Monthly Turnover	
	Mean	Median
1(lowest)	60%	29%
2	81%	39%
3	72%	29%
4	93%	44%
5(highest)	157%	98%
DIFF (5–1)	**97%*** **	**69%*** **

Gambling *Preference*

	Monthly Turnover	
	Mean	Median
1(lowest)	25%	74%
2	31%	77%
3	33%	84%
4	42%	90%
5(highest)	42%	95%
DIFF (5–1)	**17%** **	**21%** **

Regressing Gambling Behavior on Survey Responses

Volume-weighted Past One-month Count of Up-limit Hits Based on Initial Buys (2018:01-2018:09)			
Actual performance in 2017	-0.009**	Realization utility, winner	0.015
	(-2.533)		(0.843)
Over-placement, performance	0.002	Realization utility, loser	0.009
	(0.071)		-0.409
Financial literacy score	-0.031	Sensation seeking, novelty	-0.032
	(-1.478)		(-1.518)
Over-placement, financial literacy	-0.014	Sensation seeking, volatility	0.022
	(-0.633)		(1.030)
Miscalibration	0.017	Perceived information advantage	0.049**
	-0.942		(2.097)
Underestimation of transaction cost	-0.005	Dismissive of others' information	-0.001
	(-0.276)		(-0.031)
Do not consider transaction cost	0.040**	Affected by family and friends	-0.005
	-2.221		(-0.178)
Do not think bid-ask spread is a cost	-0.043**	Affected by investment advisors	0.025
	(-2.436)		-0.647
Extrapolation, up	0.003		
	-0.133	Controls	YES
Extrapolation, down	-0.001		
	(-0.045)		
Gambling preference, with prob. weighting	0.071***	Male	0.011
	(3.598)		(0.623)
Gambling preference, without prob. weighting	-0.011	R2	0.031
	(-0.482)	N	3,528

Conclusion

- We study why retail investors trade so much with a new approach
  - survey + transaction
- We show that survey responses capture trading behaviour in significant ways
  - by merging survey data with transaction data
- Our empirical analysis shows that
  - overconfidence (about information advantage) and gambling preferences have significant explanatory power on turnover
  - popular arguments such as neglect of trading cost, low financial literacy, and social influence do not explain volume
- Our study sheds light on the pros and cons of survey- and transaction-based approaches