Social media and finance

FIRN Masterclass – Day 2

J. Anthony Cookson
CU Boulder

Our plan

1. Social media landscape
   Definitions, growth over time, data sources
   Some key examples

2. **Social media as a lens**
   Examples and approach, opportunities with new formats and features

3. Social transmission bias and social signals
   Examples and approach, blossoming area of research

4. Effects of social media
   Subtopic: production, consumption, and distribution of information
   Opportunities and challenges

Social media as a lens

What does it mean to study social media "as a lens"?

- We reveal ourselves in social media interactions
  - Users publicly declare optimism or pessimism – like a real time survey
  - Likes, follows, comments, retweets – drawing connections between users
  - User join groups, like particular content, discuss politics

- "As a lens" treats this abundant information as data that are informative of classical theories. Some examples:
  - Peer effects (Bailey et al 2018 JPE)
  - Social capital and economic connections (Chetty et al 2023 Nature)
  - Disagreement (Cookson and Niessner 2020 JF)

Using social media connections to study in-person social networks

Journal of Political Economy > Volume 126, Number 6

PREVIOUS ARTICLE NEXT ARTICLE

The Economic Effects of Social Networks: Evidence from the Housing Market

Michael Bailey, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel

PDF | PDF PLUS | Abstract | Full Text | Supplemental Material

Abstract

We show how data from online social networking services can help researchers better understand the effects of social interactions on economic decision making. We combine anonymized data from Facebook, the largest online social network, with housing transaction data and explore both the structure and the effects of social networks. Individuals whose geographically distant friends experienced larger recent house price increases are more likely to transition from renting to owning. They also buy larger houses and pay more for a given house. Survey data show that these relationships are driven by the effects of social interactions on individuals' housing market expectations.

Bailey et al JPE paper had individual data, but county aggregates data have enjoyed wider application

Social Connectedness: Measurement, Determinants, and Effects

Michael Bailey

Rachel Cao

Theresa Kuchler

Johannes Stroebel

Arlene Wong

JOURNAL OF ECONOMIC PERSPECTIVES
VOL. 32, NO. 3, SUMMER 2018
(pp. 259-80)

Download Full Text PDF (Complimentary)

Google Scholar (3/19/25): 690 > 514

data.humdata.org/dataset/social-connectedness-index?fbclid=IwZXh0bgNhZW0CMTEAAR22sHAcR9jal5bS-F4VeOVNPedV8e__09FKbMyHjpE-HErTkbKzY...

OCHA Services | FAQ | Log in | Sign up | Switch to HDX Lite

HDX | Search Datasets | DATA | LOCATIONS | ORGANISATIONS | PRODUCTS | ADD DATA

HOME / DATASETS / **FACEBOOK SOCIAL CONNECTEDNESS INDEX**

Facebook Social Connectedness Index

We use an anonymized snapshot of all active Facebook users and their friendship networks to measure the intensity of connectedness between locations. The Social Connectedness Index (SCI) is a measure of the social connectedness between different geographies. Specifically, it measures the relative probability that two individuals across two locations are friends with ... More

Meta

26000+ Downloads | This dataset updates: As needed

Contact the contributor

Bailey et al JPE paper had individual data, but county aggregates data have enjoyed wider application

Social Connectedness: Measurement, Determinants, and Effects

Michael Bailey

Rachel Cao

Theresa Kuchler

Johannes Stroebel

Arlene Wong

JOURNAL OF ECONOMIC PERSPECTIVES
VOL. 32, NO. 3, SUMMER 2018
(pp. 259-80)

Download Full Text PDF (Complimentary)

Google Scholar (3/19/25): 690 > 514

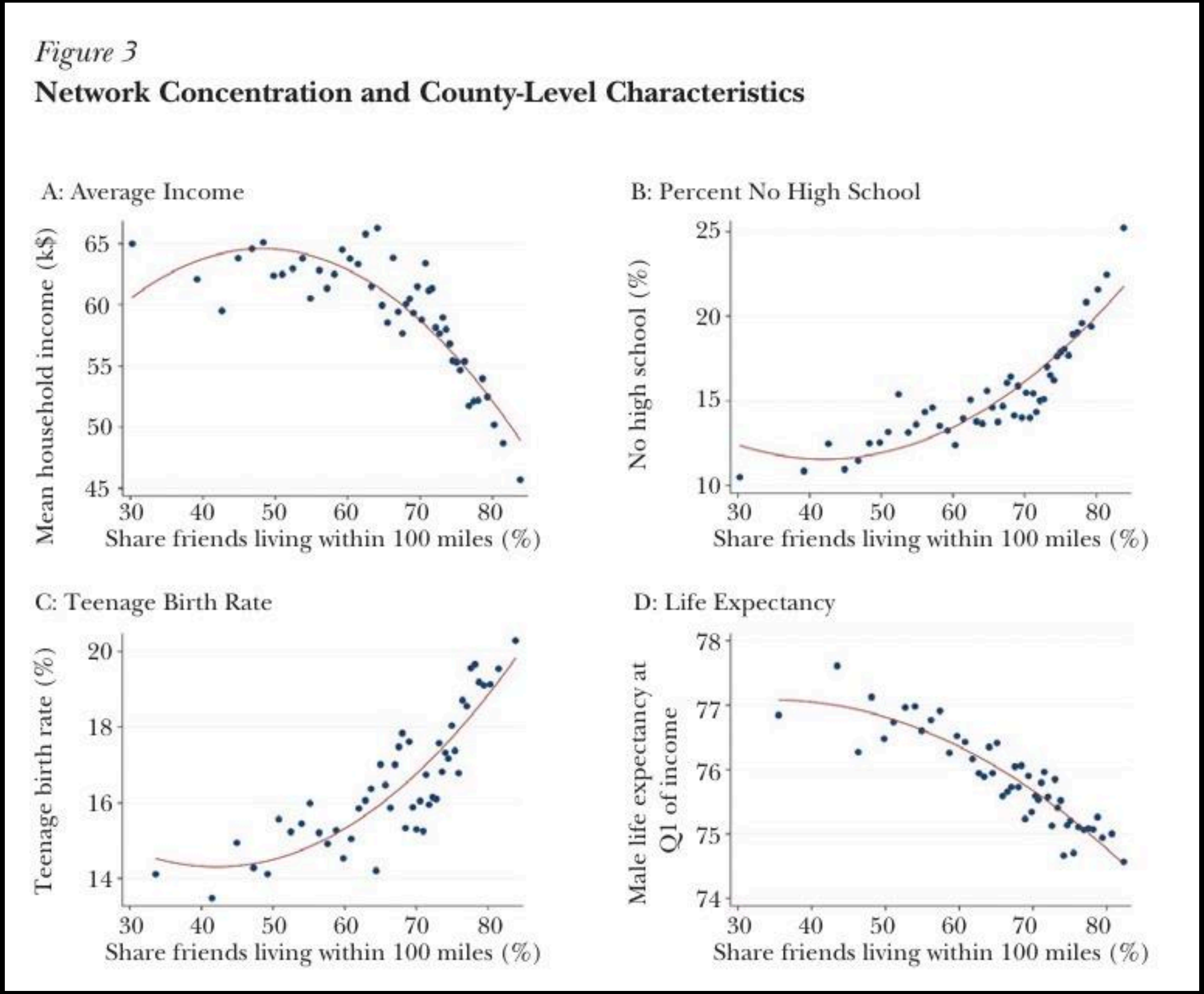
Me from 2016 when discussing the original Facebook paper

Individual-level analysis or county aggregates?

- Facebook contributes to disagreement about housing price expectations.
  - Don’t need to take a stand on social networks or experience for this.
- The county connections could proxy for…
  - Individual connections, added up.
  - Second order connections, added up.
- Could aggregates be more empirically satisfying than individual-level outcomes?
  - Play this up.

Some interesting facts in the county-level data

The geographic breadth of social interactions has a demographic tilt



Some interesting facts in the county-level data

Stronger social connections predict more trade between states, more patent citations and more migration.

In finance, people have used these data to connect firms by HQ locations, etc.

Table 3

Social Connectedness and Across-Region Economic Interactions

	(1)	(2)	(3)	(4)
Panel A: Dependent Variable: log(State-Level Trade Flows)				
log(Distance)	−1.057***		−0.531***	−0.533***
	(0.071)		(0.084)	(0.085)
log(SCI)		0.999***	0.643***	0.637***
		(0.051)	(0.071)	(0.060)
State Fixed Effects	Y	Y	Y	Y
Other State Differences	N	N	N	Y
Observations	2,219	2,220	2,219	2,219
R^2	0.912	0.918	0.926	0.930
Panel B: Dependent Variable: Indicator for Patent Citation				
log(Distance)	−0.048***		−0.011**	−0.021**
	(0.002)		(0.005)	(0.009)
log(SCI)		0.063***	0.049***	0.066***
		(0.003)	(0.006)	(0.012)
Technological Category + County Fixed Effects	Y	Y	Y	Y
Cited + Issued Patent Fixed Effects, Other County Differences	N	N	N	Y
Observations	2,171,754	2,171,754	2,171,754	2,168,285
R^2	0.056	0.059	0.059	0.101

(Investor) social media as a lens

The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

ARTICLE

Why Don't We Agree? Evidence from a Social Network of Investors

J. ANTHONY COOKSON, MARINA NIESSNER

First published: 18 October 2019 | https://doi.org/10.1111/jofi.12852 | Citations: 189

Volume 75, Issue 1
February 2020
Pages 173-228

References Related Information

Disagreement literature has mostly relied on analyst forecast dispersion as a proxy for disagreement (Diether, Malloy and Scherbina 2002).

But, analysts aren't investors, and analyst forecasts are updated infrequently.

By contrast, StockTwits has investors---albeit selected--- and frequent bullish versus bearish declarations → Could we build a measure of disagreement?

That's why we started the project.

Motivation

- Investor disagreement is central toward understanding prices (bubbles) and trading volume in financial markets.

- Clearly, market participants disagree
  - see analyst dispersion, survey evidence.

Motivation

- Investor disagreement is central toward understanding prices (bubbles) and trading volume in financial markets.

- Clearly, market participants disagree
  - see analyst dispersion, survey evidence.

- But, much less is known about **why investors disagree?**
  - Theory proposes two main reasons:
    - Different information sets (e.g., Hong and Stein)
    - Different investment models (e.g., Harris and Raviv 1993)

  - Empirical work is scarce
    - Data limitation, since disagreement is defined as *a difference of opinions*, which must be inferred.
    - Differences in information environment are more natural to observe.

      e.g., Bailey et al (2016, Facebook paper), Chang et al (2015, language diversity), and Knyazeva, Knyazeva and Kostovetsky (2015, mutual fund managers)

What we do

- New data on investor **opinions** and **investment approaches**
  - StockTwits, an investment-specific social network
- Construct a daily measure of disagreement
- Compare disagreement **within** and **across** approaches
  - Proxies for information sets versus investment models

What we find

- Cross-group disagreement on StockTwits strongly predicts aggregate trading volume.
    - Daily frequency, disagreement before trading.

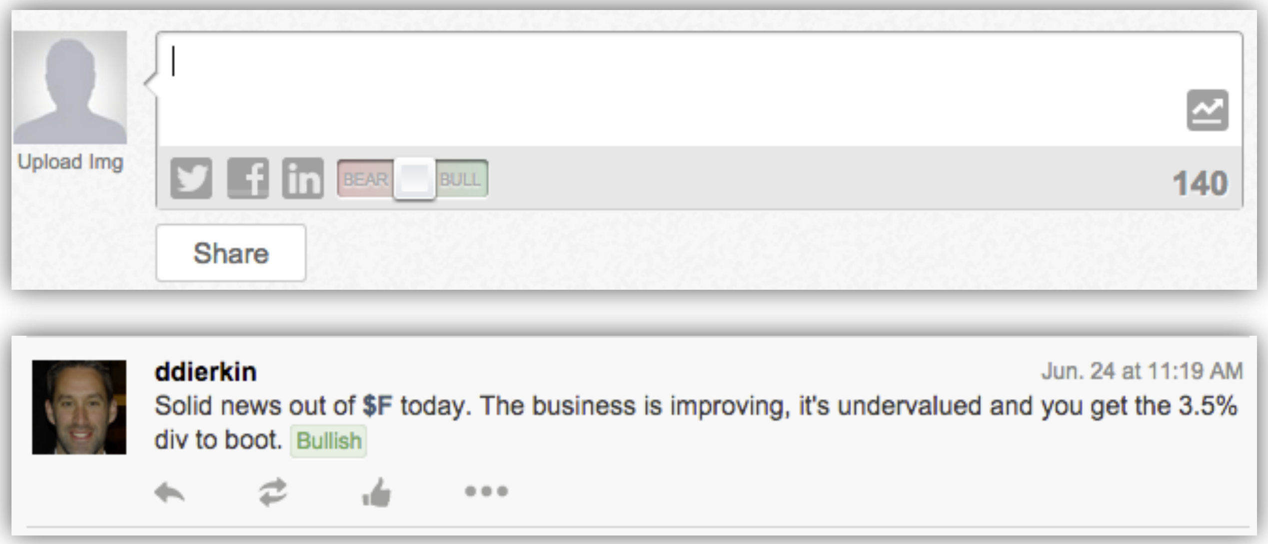
- Differing investment philosophies matter.
    - Contribute 34% to 65% of overall disagreement.
    - Strongly associated with abnormal trading volume.

- Daily disagreement is useful more broadly.
    - Explains 1/3 of volume spikes around earnings announcements.
    - Distinct from attention, news, sentiment.

Data

- Data are from StockTwits.com:
  - "a social media platform designed for sharing ideas between investors, traders, and entrepreneurs" (Wikipedia)
  - Founded in 2009, has more than 500K registered users

Example of a post:



Our Sample

- Sample period: January 2013 - September 2014
- 100 most discussed firms
  - our results are less susceptible to users posting false opinions
- Concentrate on messages that talk about 1 ticker
- Final sample has 1,460,349 messages by 11,874 users

Key advantage of our setting

- Self-stamped sentiment – **bullish** versus **bearish** tags on individual posts

- Two nice features:
  - This avoids any classification choices, making our approach transparent (recall the discussion of sentiment classification in day 1)

  - This gives us a large training sample, so we can classify all messages

Example of Differing Investment Approaches

<u>*Fundamental*</u>

<u>*Technical*</u>

We validate these groups are distinct

Contextually… we examine which words are “overused” (following the saliency definition of Goldmith-Pinkham, Hirtle and Lucca (2014))

Panel A: Most Salient Words Used by Approach

Approach	Most Common Unique Words
Fundamental	eps, sales, growth, sentiment, read, revenue, earnings, million, quarter, consensus, billion, share, cash, results, analysts
Technical	chart, support, nice, break, looking, looks, gap, move, day, stop, calls, daily, close, resistance, bounce
Momentum	play, calls, time, via, week, day, news, squeeze, hod (high of day), hit, shares, cover, highs, run, money
Value	view, attempts, bulls, rising, aboard, stair, intraday, correction overextended, breakdown, fresh, mayb, steak, moved, rollout
Growth	news, er (earnings report), hope, green, shares, plug, money, article, time, bears, waitings, ve, wait, board, share, future

We validate these groups are distinct

Not just salient words, but the **distribution of words is distinct** (using Kullback-Liebler Divergence, a measure of distance), and it matches off-platform writing.

Panel B: Kullback-Liebler Divergences of Word Distributions by Approach (Fundamental as the Baseline Approach)

	Growth	Momentum	Technical	Value
Divergence from Fundamental	0.0854	0.1146	0.1919	0.2336
Standard Error (100 bootstrap replications)	0.00008	0.00009	0.00009	0.00008

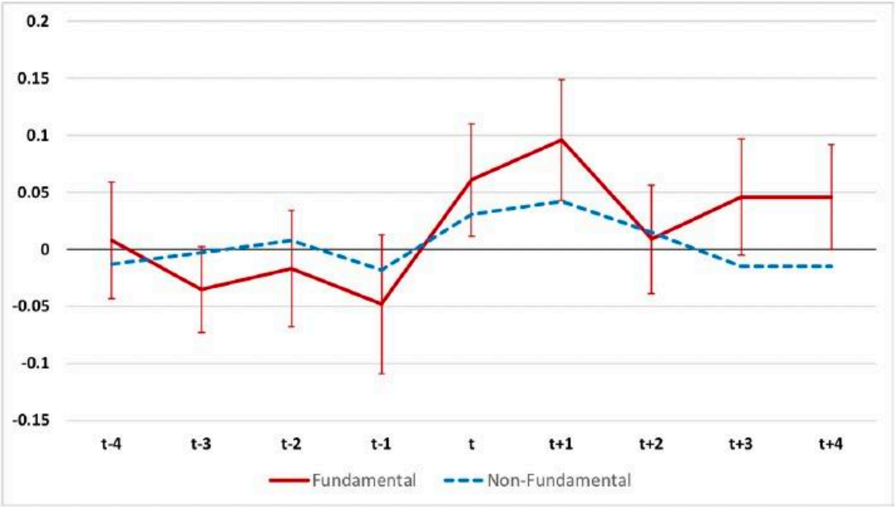
Panel C: Divergence from Writing by Focal Investors (Comparisons to Writing Outside of Stock-Twits)

	Fund, Value, Growth	Technical, Momentum	Difference	Standard Error	T-stat
Technical Focal Investor	1.186	1.113	0.073	0.028	2.624
Value Focal Investor	1.284	1.512	-0.228	0.039	-5.871

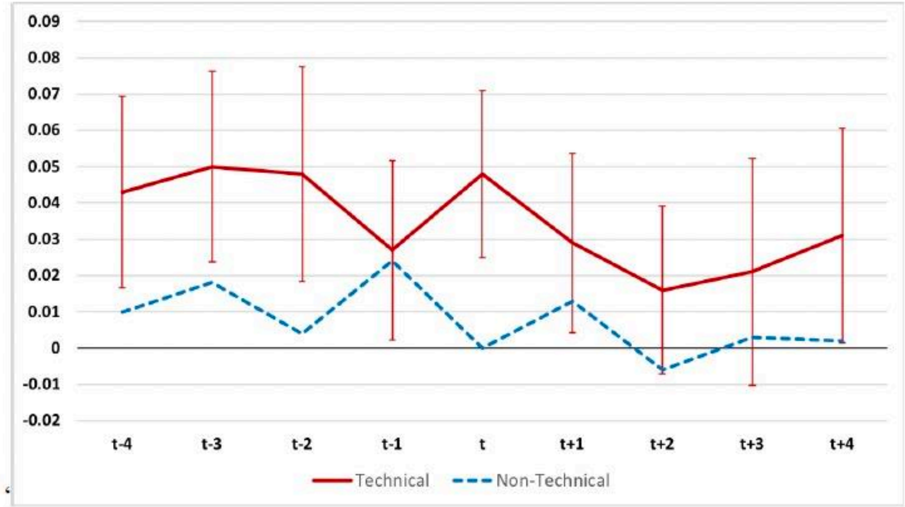
We validate these groups are consistent with their labels

Technical and fundamental events from RavenPack line up with on-platform sentiment updates

(a) Sentiment around Earnings Announcements



(b) Sentiment Around Technical Events

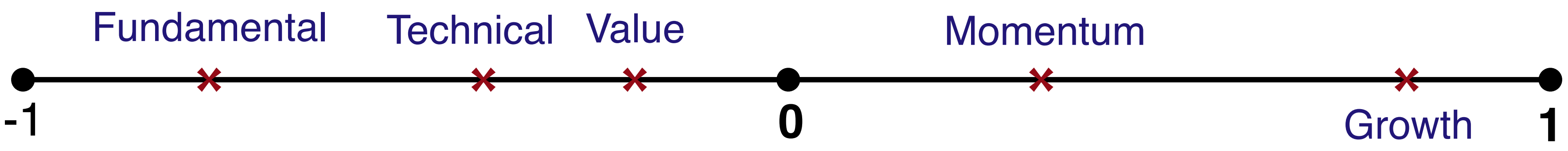


Roadmap

- Data
- **Disagreement and Trading Volume**
- Cross-group vs. Within-group Disagreement
- Disagreement around Earnings Announcements
- Other Checks

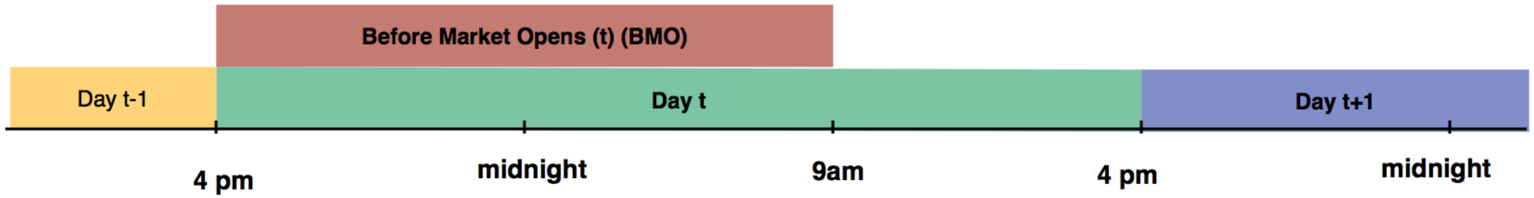
Cross-group Disagreement

$$AvgSentiment_{itg}$$



$$CrossDisagreement_{it} = WeightedSD(AvgSentiment_{itg})$$

Cross-group Disagreement and Trading Volume



$$AbLogVol_{i,t} = \alpha + \beta CrossDisagreement_{i,t} + \gamma AbLogVol_{i,t-1} + TimeFEs + FirmFEs + \epsilon_{i,t}$$

	(1)	(2)
Cross-group Disagreement (t)	0.034***	
	(0.008)	
Cross-group Disagreement (BMO, t)		0.038***
		(0.006)
Abnormal Log Volume (t-1)	0.738***	0.736***
	(0.013)	(0.014)
Observations	42,415	42,415
R-squared	0.601	0.601
Year, month, day-of-the-week FEs	X	X
Firm FEs	X	X

Roadmap

- Data
- Disagreement and Trading Volume
- **Cross-group vs. Within-group Disagreement**
- Disagreement around Earnings Announcements
- Other Checks

Disagreement Decomposition

Disagreement = (weighted) standard deviation of sentiment within group-day

	Mean	Stdev	p25	p50	p75
All Investors	0.467	0.446	0	0.628	0.932
W. Average within-group Disagreement	0.244	0.299	0	0	0.480

52%

Within vs. Cross-group Disagreement

$$AbLogVol_{it} = \alpha + \beta_1 CrossDisagreement_{it} + \beta_2 WithinDisagreement_{it} + AbLogVol_{it-1} + TimeFEs + FirmFEs + \epsilon_{it}$$

	(1)	(2)
Cross-Group Disagreement	0.034***	0.049***
	(0.008)	(0.008)
W. Average within-group Disagreement		0.195***
		(0.012)
Abnormal Log Volume (t-1)	0.738***	0.700***
	(0.013)	(0.016)
Observations	42,415	42,415
R-squared	0.601	0.623
Year, month, day-of-the-week FEs	X	X
Firm FEs	X	X

Other Checks

- The measure is externally valid.
  - Results are similar when…
    - dropping technical investors.
    - using different subsets of firms (top 50, top 150, 51-150).

Takeaways on “Why don’t we agree?”

- Generate a high-frequency measure of investor disagreement based on investors’ opinions.
  - Measure explains and forecasts volume, and volume spikes around earnings announcements.

- Insight into the *source* of investor disagreement.
  - Differences in investors’ models matter for investor disagreement.

Research in this StockTwits data ecosystem on different topics

- Echo Chambers with Joey Engelberg and Will Mullins (*tomorrow*)
- Does partisanship shape investor beliefs? with Engelberg and Mullins
- Does disagreement facilitate informed trading? with Fos and Niessner
- Do I really want to hear the news? with Ben-Rephael and Izhakian

As with the Facebook data, follow-on research has been stimulated here because we have made our disagreement measures available.

The Review of Asset Pricing Studies

Society for Financial Studies

JOURNAL ARTICLE

Does Partisanship Shape Investor Beliefs? Evidence from the COVID-19 Pandemic FREE

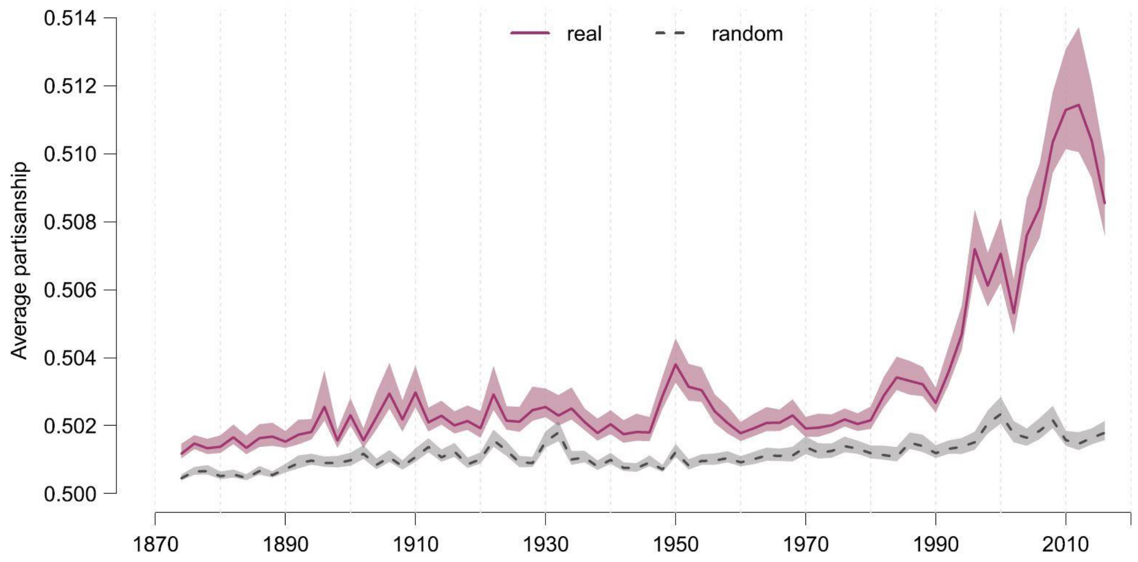
J Anthony Cookson ✉, Joseph E Engelberg, William Mullins

The Review of Asset Pricing Studies, Volume 10, Issue 4, December 2020, Pages 863–893, https://doi.org/10.1093/rapstu/raaa018

Published: 29 September 2020 **Article history** ▾

Polarization on the rise

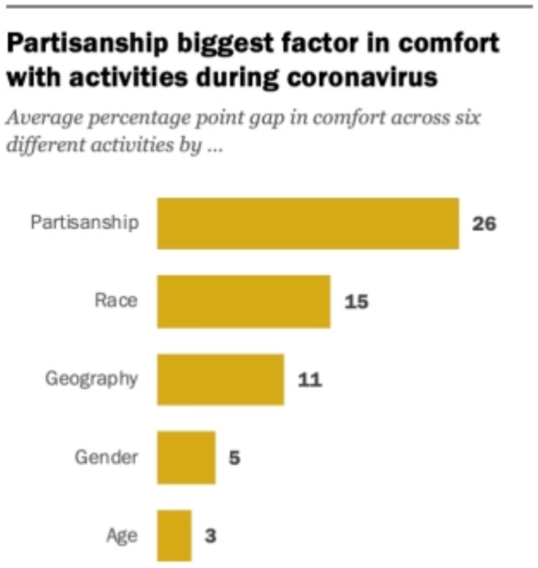
- Political polarization has increased substantially over the last 30 years.
  - News and other media (Gentzkow and Shapiro ‘10)
  - Speech of politicians (Gentzkow, Shapiro and Taddy ‘19)
  - Fewer cross-party interactions on Thanksgiving (Chen and Royla ‘18)



Polarization on the rise

- Political polarization has increased substantially over the last 30 years.
  - News and other media (Gentzkow and Shapiro ‘10)
  - Speech of politicians (Gentzkow, Shapiro and Taddy ‘19)
  - Fewer cross-party interactions on Thanksgiving (Chen and Royla ‘18)

- Not just the political domain.
  - Where to live (Bishop ’08)
  - Everyday activities
    - Social distancing, masks, etc.



Polarization on the rise

- Political polarization has increased substantially over the last 30 years.
    - News and other media (Gentzkow and Shapiro ‘10)
    - Speech of politicians (Gentzkow, Shapiro and Taddy ‘19)
    - Fewer cross-party interactions on Thanksgiving (Chen and Royla ‘18)
- Not just the political domain.
    - Where to live (Bishop ’08)
    - Everyday activities
        - Social distancing, masks, etc.
- However, unclear that partisan differences would matter in contexts with high personal stakes (e.g., investing).

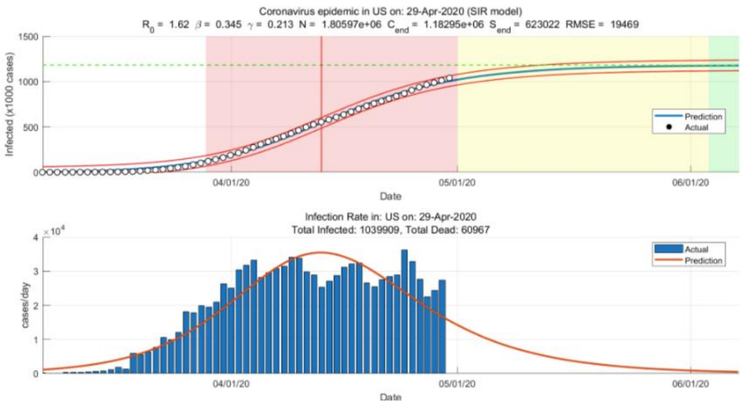
Question, setting & core findings

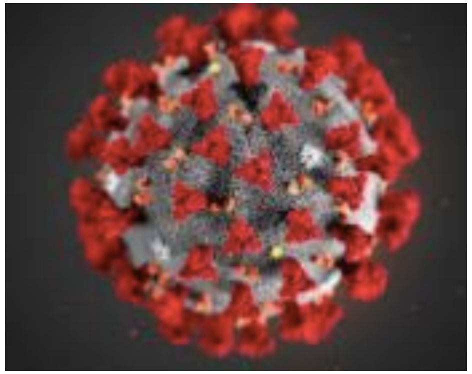
- **Question:** Does partisanship shape investment beliefs?
  - **Strong null**: information about cash flows should inform investment beliefs, not political identity.

Question, setting & core findings

- **Question:** Does partisanship shape investment beliefs?
    - **Strong null**: information about cash flows should inform investment beliefs, not political identity.

- **Setting**: The COVID-19 period.
    - Daily “belief updates” from StockTwits through COVID-19 period.
    - Identify partisan Republicans versus others from pre-COVID tweets.





Question, setting & core findings

- **Question:** Does partisanship shape investment beliefs?
  - **Strong null**: information about cash flows should inform investment beliefs, not political identity.

- **Setting**: The COVID-19 period.
  - Daily "belief updates" from StockTwits through COVID-19 period.
  - Identify partisan Republicans versus others from pre-COVID tweets.

- **Core findings**: During the pandemic,
  - Partisan Republicans remain substantially more optimistic.
  - … but become more pessimistic about Chinese stocks.
  - And, partisan disagreement explains 20% of the Covid-19 rise in turnover.

Measurement

Measurement using StockTwits

- **Investor Beliefs.**
  - StockTwits users self-declare as bullish/bearish about a particular symbol.
  - 5.3 million messages about 1,042 securities from Mar’19 to Apr’20

- **Partisan Identity.**
  - Party-identifying text of tweets from 2013-2019.
    - Contextual detail in coming slides.
  - Users make or like 3+ Republican messages ➔ ”Republican”
    - ~10% of messages are made by Republican users from Mar ‘19 – Apr ‘20.

Party classification of users

- Iterative keyword approach (Gentzkow and Shapiro ‘10):
  - Start with “#MAGA2020” and “#Trump2020”
    - Identify Republican users.
  - For those users, look for overused political terms (“The Liberals”)
    - Expand set of Republican users.
  - Repeat until new users not identified.
    - (we follow a similar process to obtain a keyword list for Democrats)

- We classify as Republican any users in the pre-2020 period who made or liked 3 more Republican messages than Democrat messages.

Partisan keyword list

(a) *Partisan keywords*

Republican keywords		Democrat keywords	
"#MAGA"	"The Liberals"	"Drumpf"	"Orange Colored"
"Russia Hoax"	"Russian Collusion"	"Trump Nationalism"	"Idiot in Chief"
"#TRUMP2020"	"Stupid Dems"	"Trumptard"	"Criminal POTUS"
"Hussein Obama"	"Leftists"	"Trump is a liar"	"Trump is an idiot"
"Obummer"	"Trump Derangement"	"Idiot Trump"	"Clown Trump"
"Fake News Media"	"The Socialist"	"Faux News"	Imbecile Trump
"Crooked Hillary"	"MAGA 2020"	"Clown Child"	"Trump is an Imbecile"
"Snowflake"	"The Commie"	"Stupid Trump"	Orange Scum"
"Liberal Media"	"Libtard"	"Pig Clown"	"Scumpig Clown"
"Libs"	"Stupid Democrats"	"Liar in Chief"	"Lying Trump"
"Trump Hater"	"Sleepy Joe"	"Liar Trump"	"#IMPEACHTRUMP"
"Typical Liberal"	"Liberal Democrat"	"#F***TRUMP"	
"Liberal Agenda"	"You Liberal"		
"Your Liberal"			

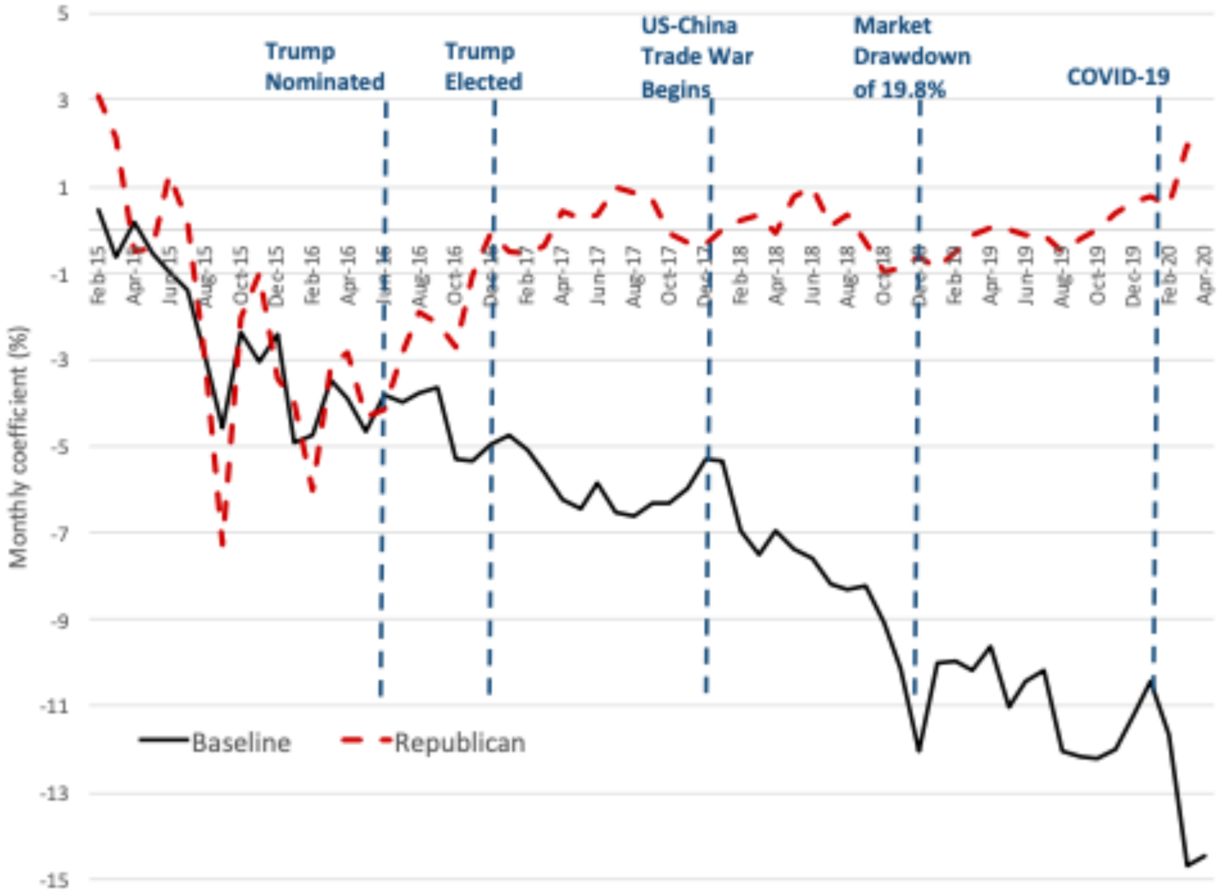
Examples

(b) Examples of partisan Tweets

Republican example messages	
October 10, 2018	"Fox News... This crash will teach those libtards!! $spy
October 27, 2019	"Therapy bro, Trump derangement syndrome is no joke. Get some meds"
July 8, 2019	"I probably won't be alive to see it but the US is a short step to being a socialistic country. Only one election away. Vote TRUMP 2020 or else"

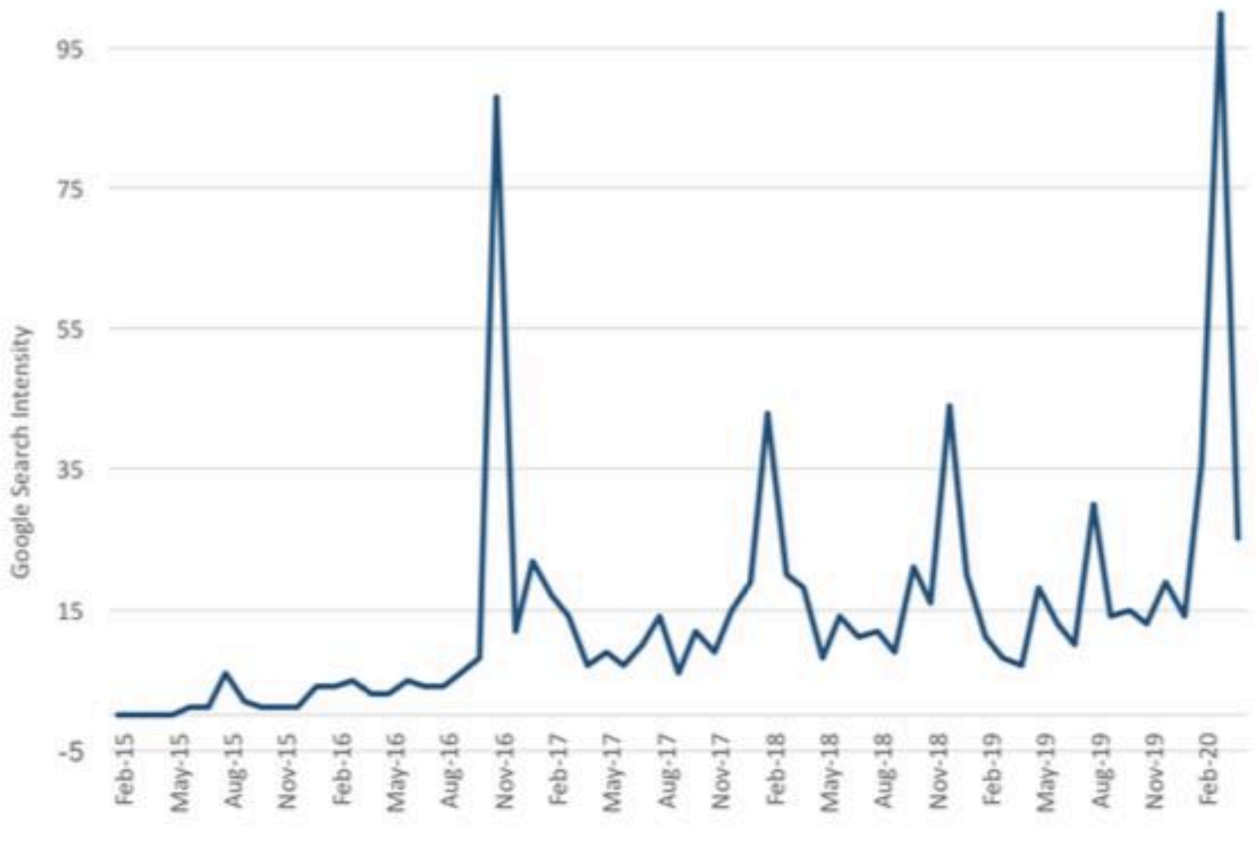
Democrat example messages	
January 29, 2018	"$WYNN the only one less popular than Wynn now is the orange colored scumpig clown child masquerading as potus. BEARISH"
July 20, 2018	$GM drumpf is killing this stock
November 2, 2018	"Glad to see the Manipulator in Chief saw Apple earnings on Faux News scrambling for a deal now to try to save markets; very stable genius!

Does it track notable events?



Beyond StockTwits

(b) Google Search Intensity for "Trump and Stock Market"

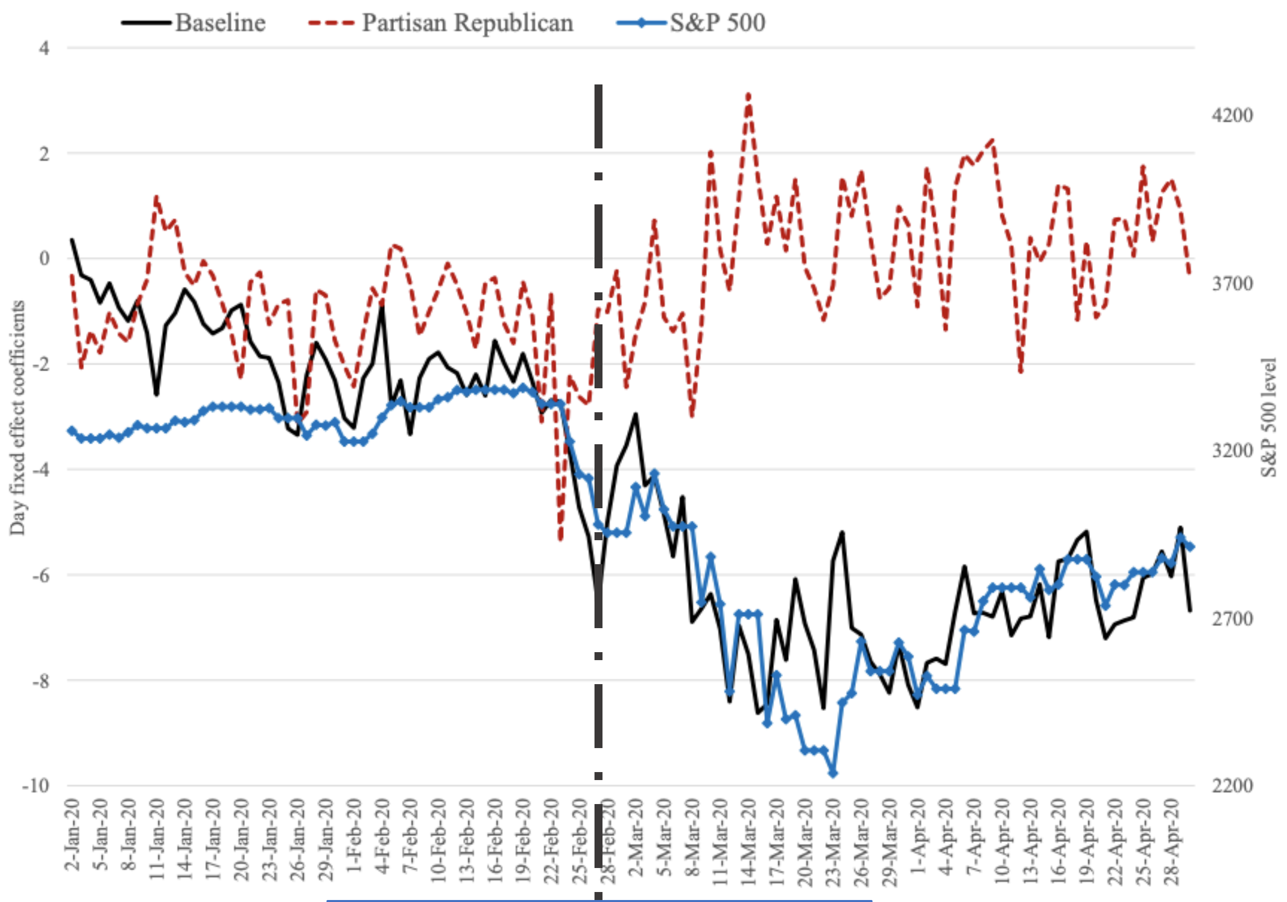


Empirical Tests of Sentiment

Timing and Graphical Evidence

- Timing can be daily, weekly or monthly.
  - C-19 event is 1st confirmed case of US community spread on Feb 26, 2020.
- We present main findings graphically.
  - Highlights the timing and persistence of the effect.

Daily evidence: January 2020 – April 2020

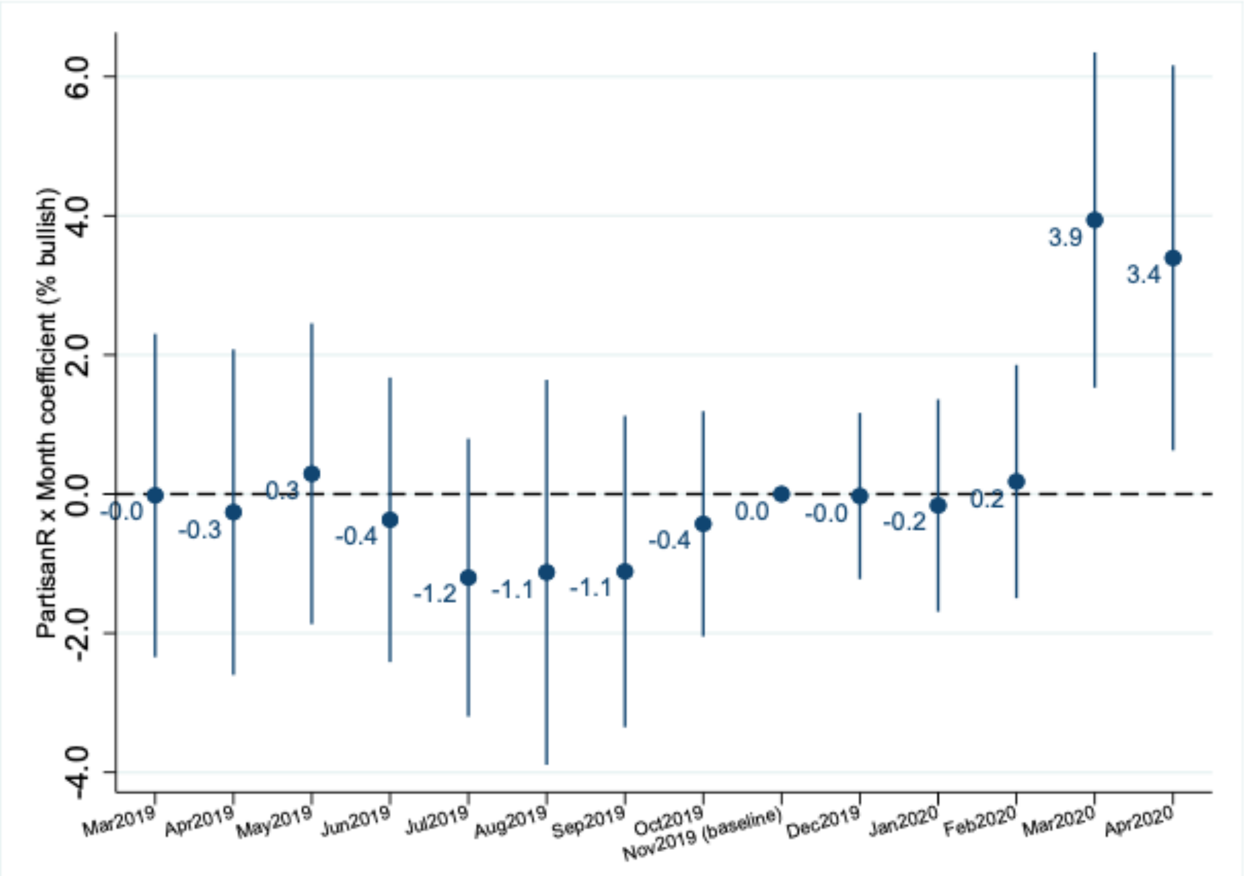


US Community Spread: Feb 26

Main specification

Monthly

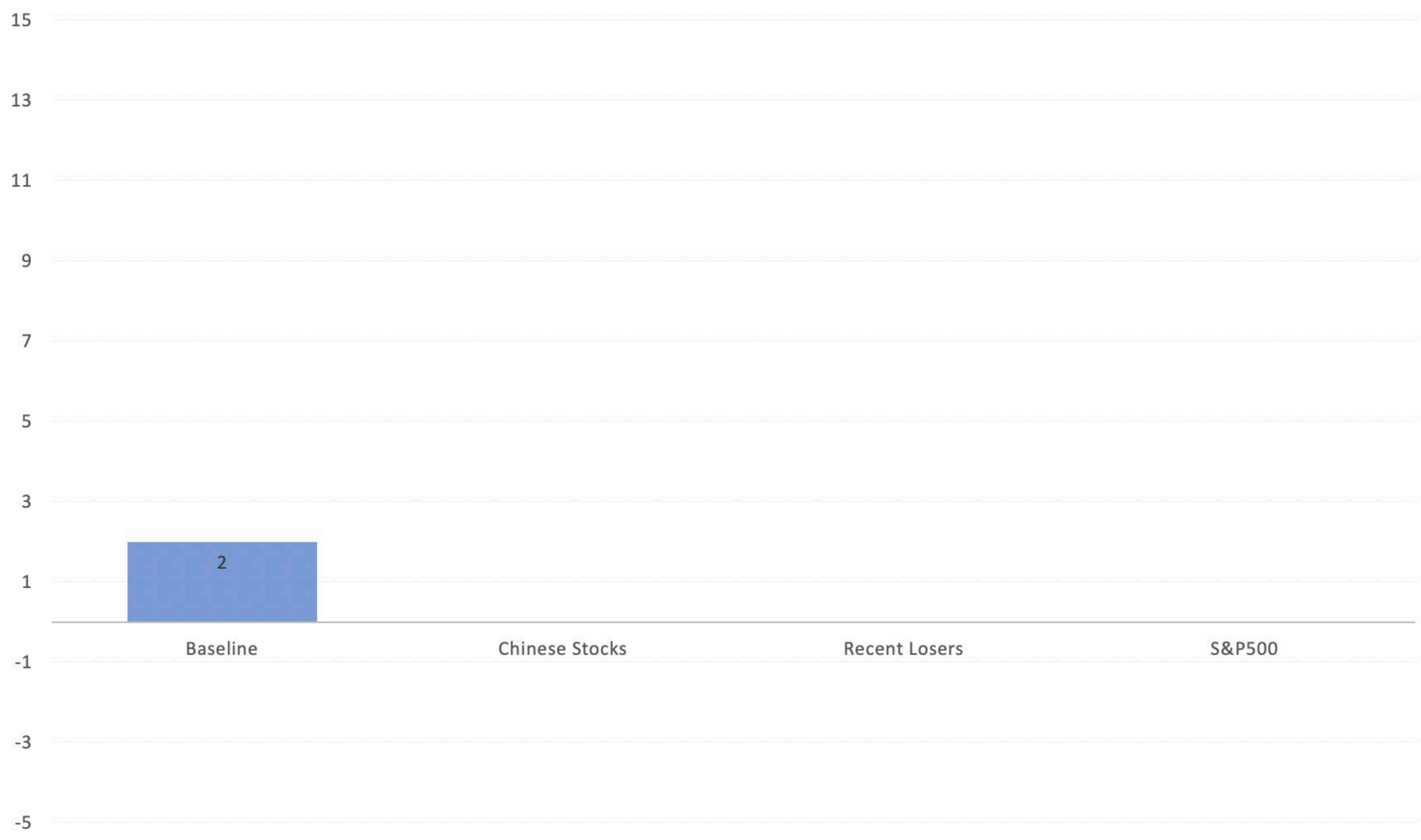
(a) All Stocks, excluding Small Cap Stocks



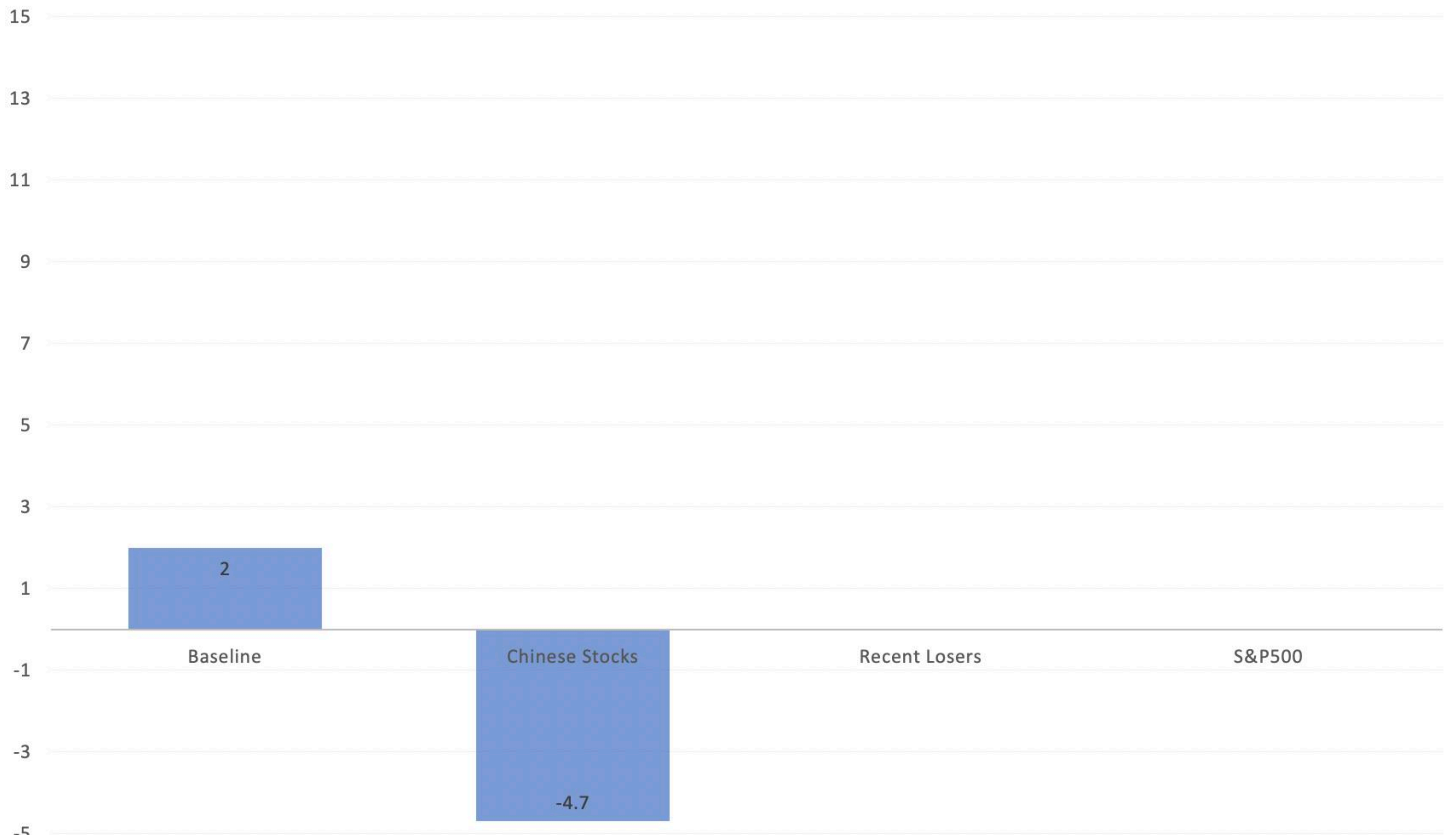
Heterogeneity

- We examine heterogeneity along the following dimensions:
  - Chinese stocks
  - Recent losers
  - Large, bell-weather stocks
- Is the partisan effect of the pandemic different for these groups of stocks?

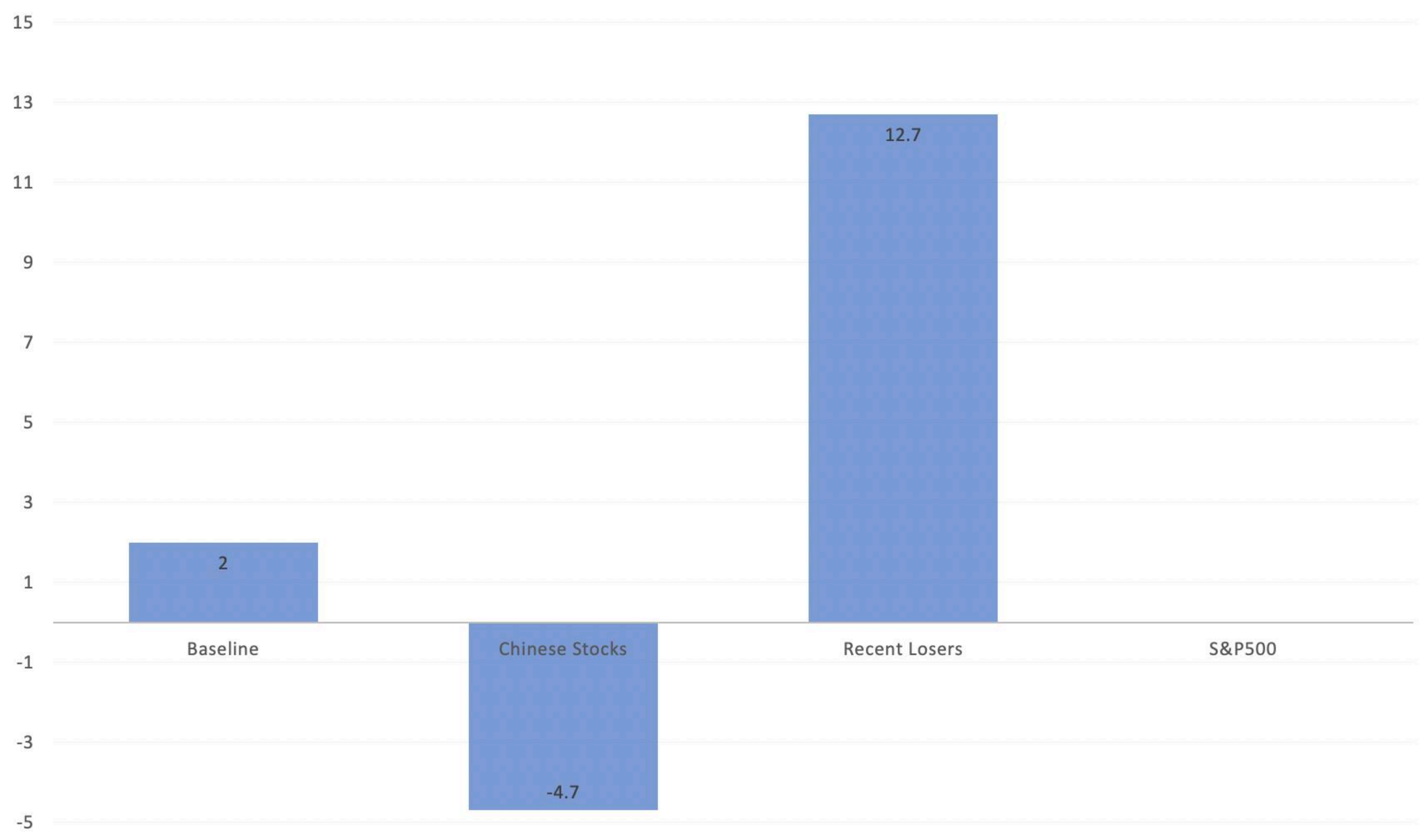
Heterogeneity in Partisan Optimism



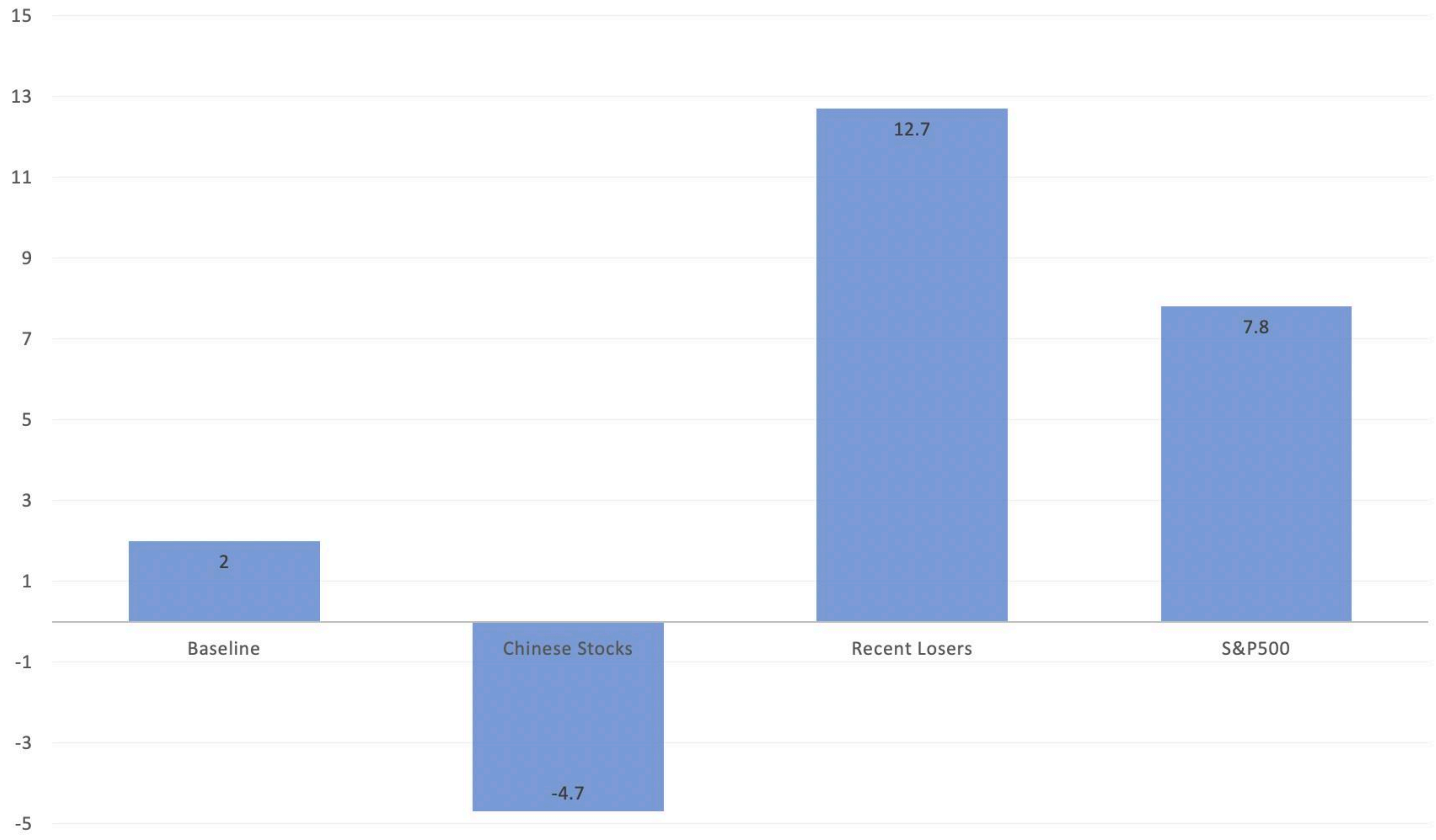
Heterogeneity in Partisan Optimism



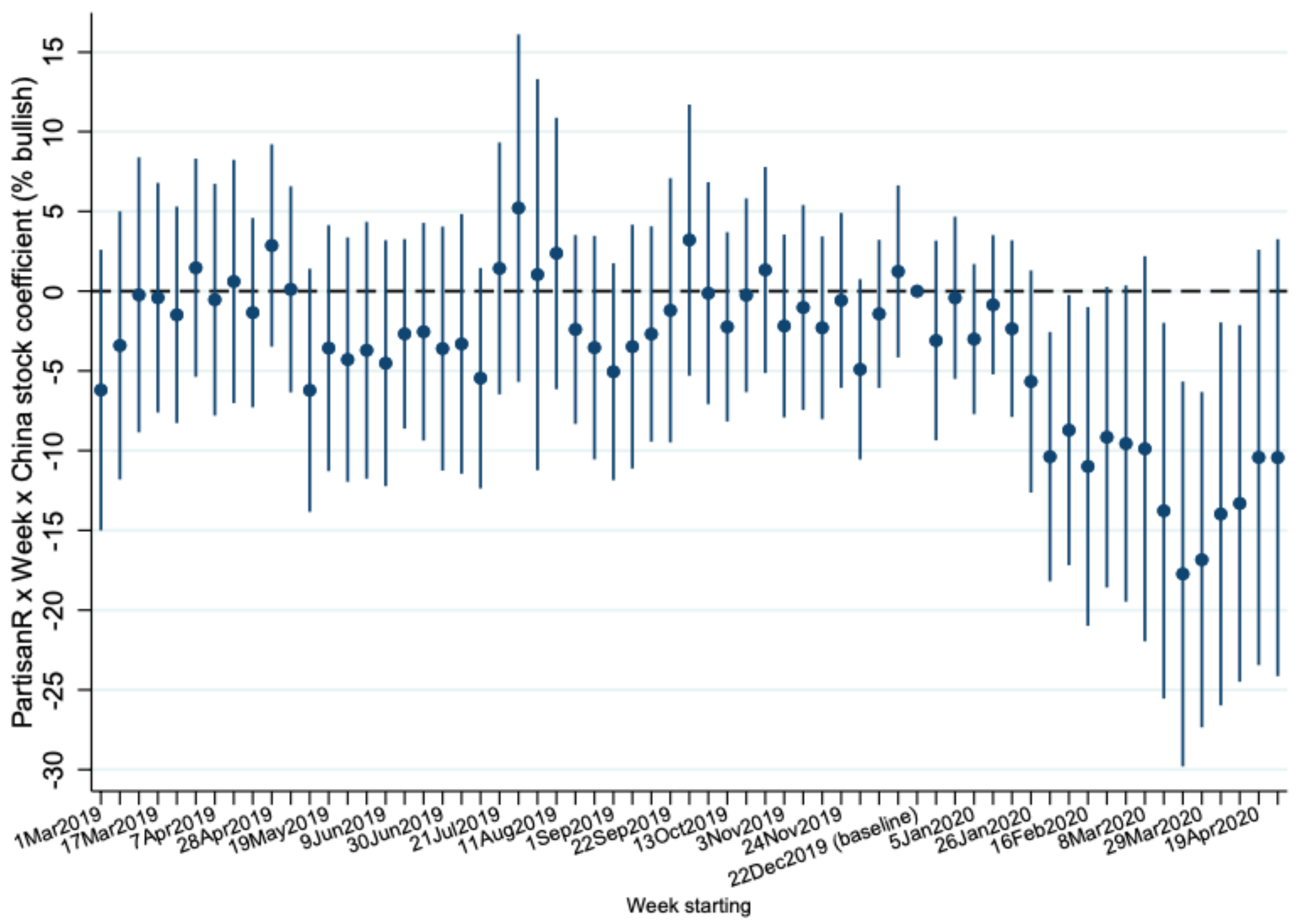
Heterogeneity in Partisan Optimism



Heterogeneity in Partisan Optimism

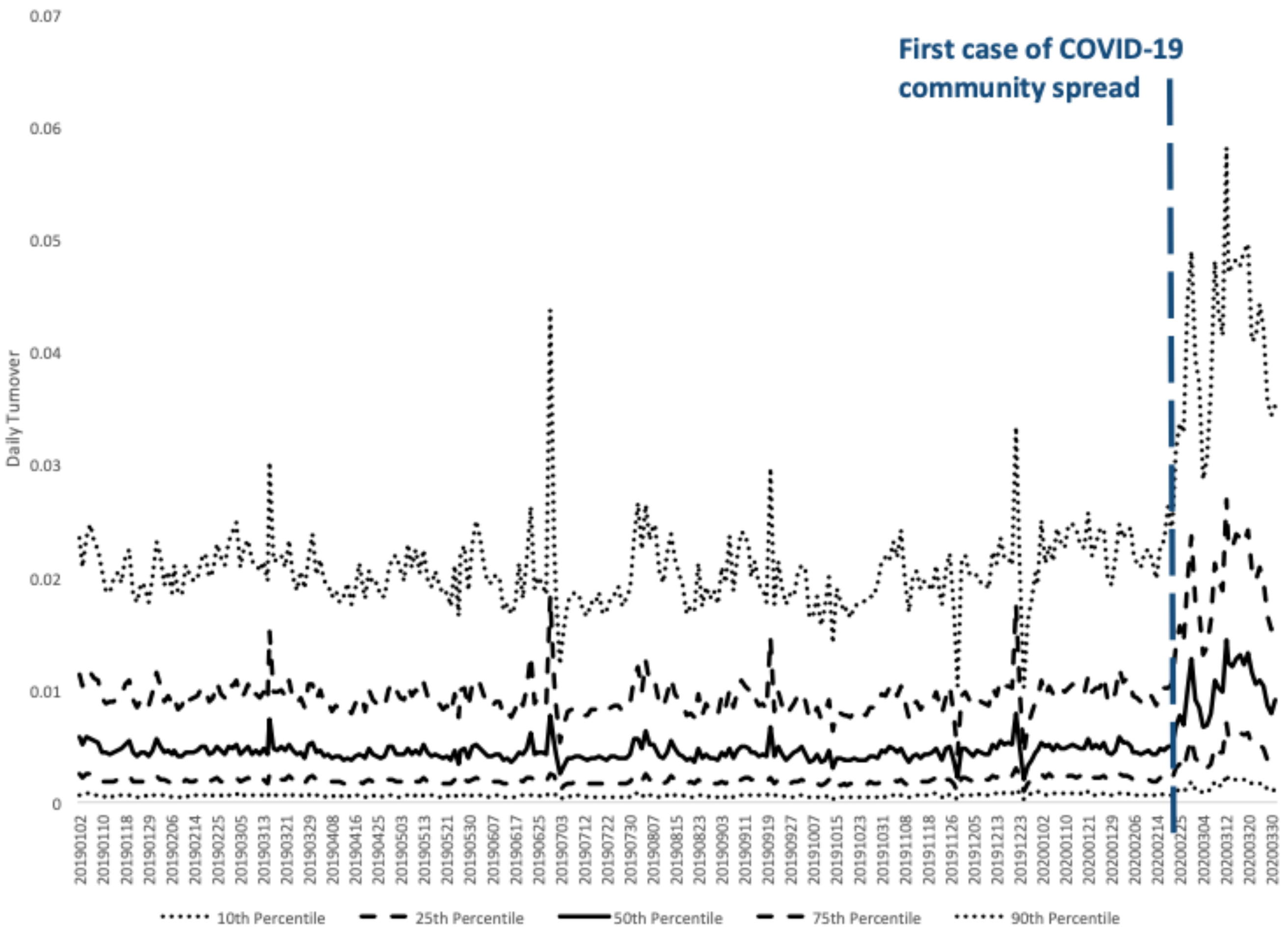


Chinese pessimism in mid March



Relating to Stock Turnover

Stock turnover increased ~40% during the pandemic

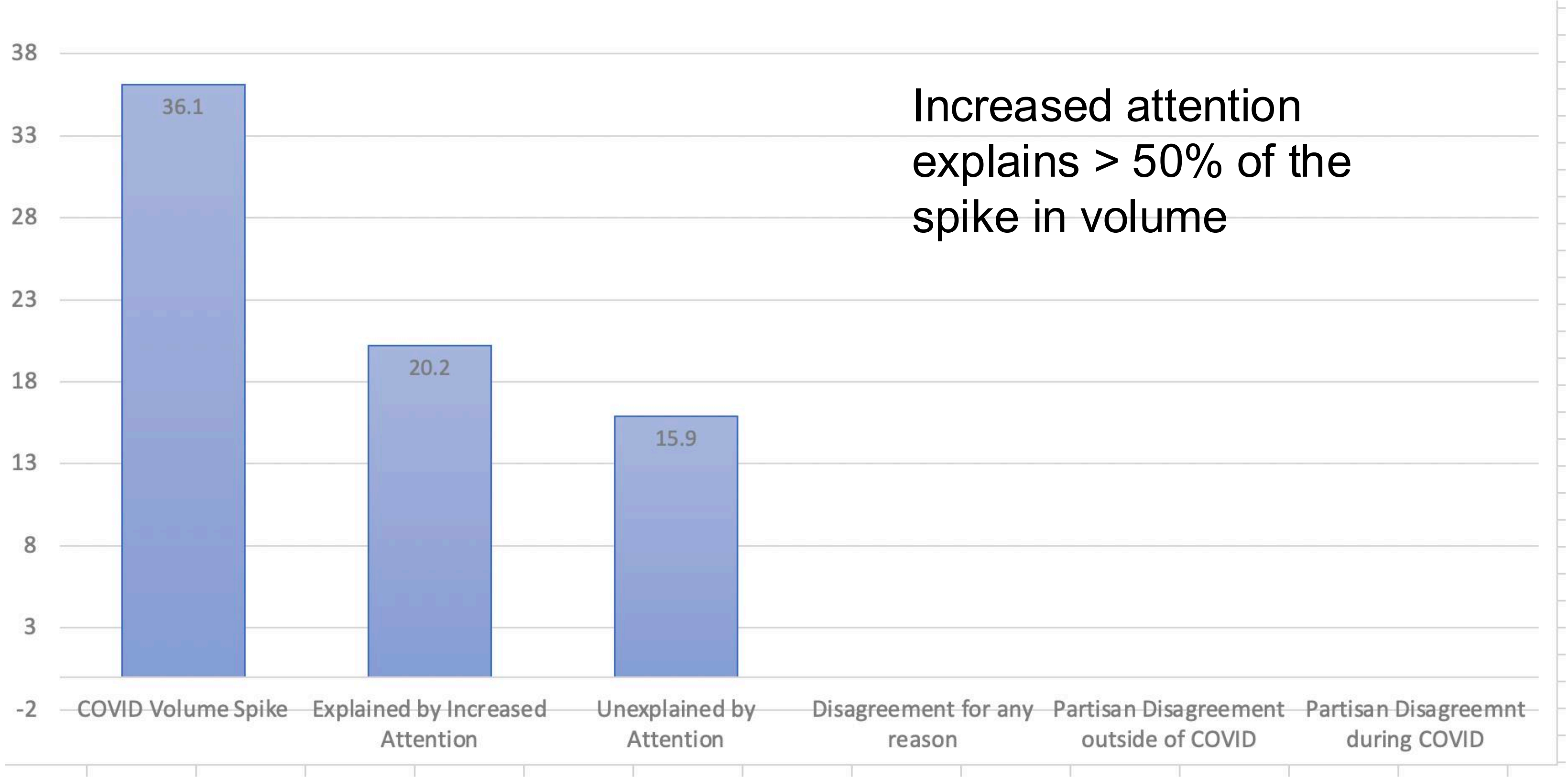


Partisan disagreement and trading

- At stock-day level, measure partisan disagreement.
  - Difference between R and D sentiment about a stock.

- Then, relate partisan disagreement to daily stock turnover.

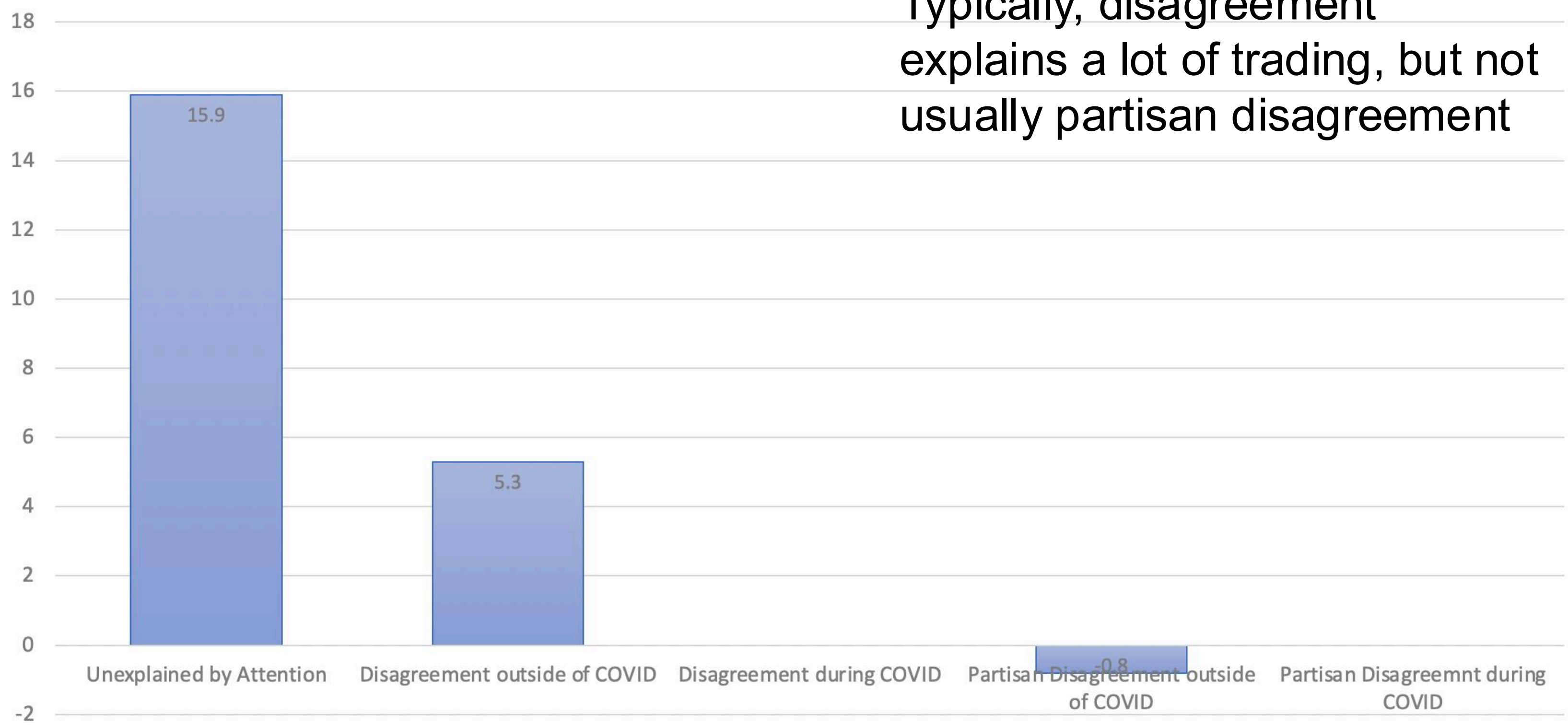
Stock turnover results



Increased attention explains > 50% of the spike in volume

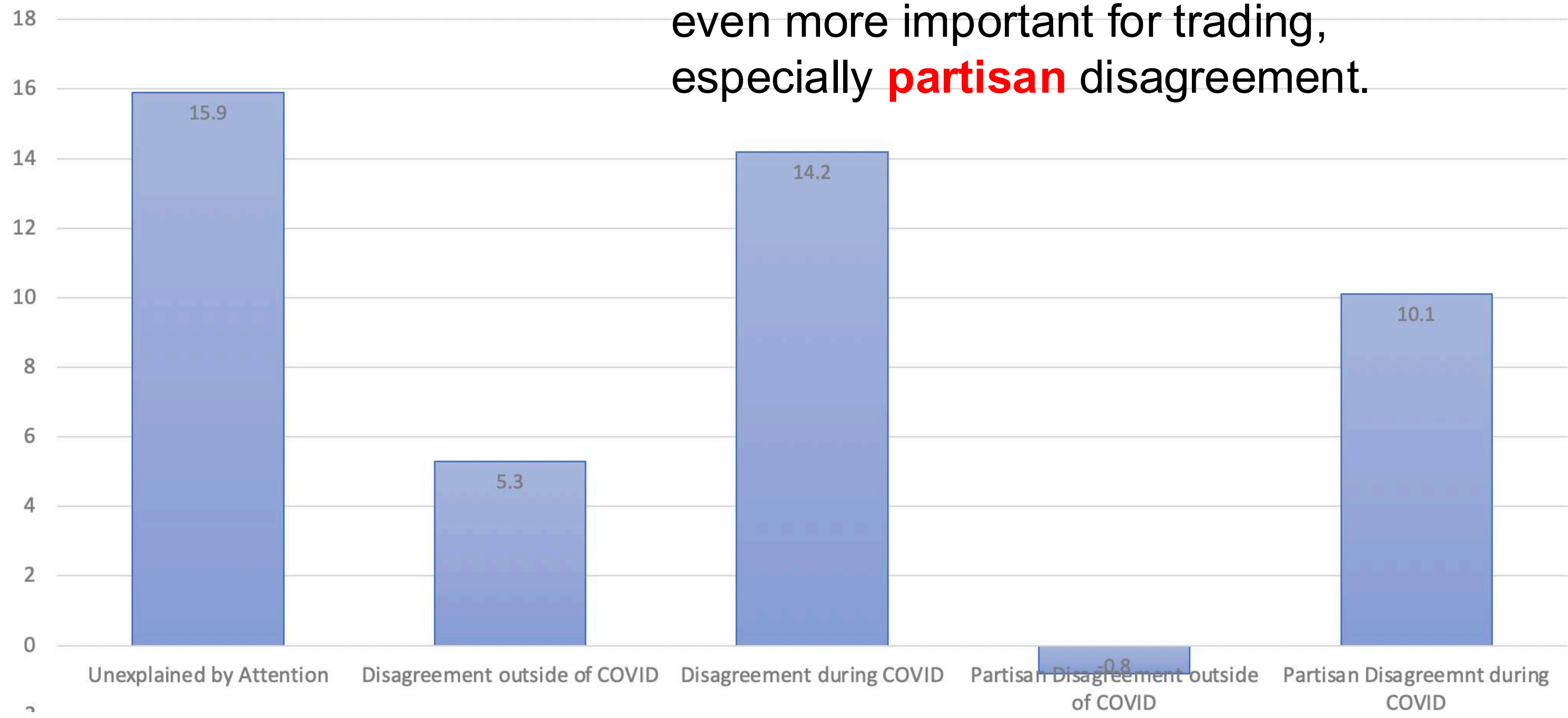
Stock turnover results

Typically, disagreement explains a lot of trading, but not usually partisan disagreement



Stock turnover results

During COVID, disagreement has been even more important for trading, especially **partisan** disagreement.



A standard deviation increase in partisan disagreement ≈ 28% of the COVID volume spike!

Conclusion

- Partisanship shapes investor beliefs during COVID.
  - Uncertainty drives investors to core identities.

- Part of a bigger picture linking politics and finance?
  - Timing of partisan disagreement lines up with politicization of stock market off-StockTwits.
  - Partisan disagreement is tightly linked to trading in the broader market.

- Of course, partisans disagree, but the partisan *stock investing* shows just how polarized society has become.

Social Media as a Lens on Other Platforms

Example: “Disagreement on the Horizon” J. Anthony Cookson, Chukwuma Dim and Marina Niessner, Working Paper

Disagreement on the Horizon

73 Pages • Posted: 11 Jul 2024 • Last revised: 13 Nov 2024

J. Anthony Cookson
University of Colorado at Boulder - Leeds School of Business

Chukwuma Dim
George Washington University

Marina Niessner
Indiana University - Kelley School of Business - Department of Finance

Date Written: July 03, 2024

Abstract

Using data from The Motley Fool's social prediction platform, CAPS, we document substantial differences in stock predictions across investment horizons. Short- and long-horizon investors respond differently to macroeconomic events and firm news announcements. At the onset of the Covid-19 pandemic, the sentiment of short-horizon predictions became sharply more negative while long-horizon predictions remained optimistic. Short-horizon investors also react more than twice as strongly as long-horizon investors to earnings surprises and technical view events. Around acquisition rumors, short- and long-horizon investors update in opposite directions about the target: short-term investors become more optimistic, while long-term investors become more pessimistic. Motivated by these findings, we develop a firm-day measure of *horizon disagreement*, spanning from 2006 to 2022, and find it relates significantly to abnormal trading. Additionally, the disagreement-trading relation strengthens on earnings announcement days, providing new evidence on the role of model disagreement.

Keywords: Social finance, social media, investment horizon, disagreement, trading

Investor horizon is important

Short-term versus long-term investors have impacts on a variety of corporate outcomes

- *ESG investment requires a long-term perspective (Starks et al 2017)*
- *R&D expenditures often take years to pay off (Bushee 1998)*
- *Firm lifespans often outlive the people and contracts (Cziraki and Groen-Xu 2020; Agarwal et al 2023)*

Different investor horizons are often posed as a reason why investors disagree (Hong and Stein 1999; Banerjee, Davis and Gondhi 2018).

- *A theory-empirics gap: no direct measure of investor horizons to study disagreement across horizons.*
- *Existing measures: portfolio turnover, churn, time left in a VC cycle, time left on CEO contract.*

Horizon as a form of model disagreement

Disagreement literature has focused extensively on the source of disagreement.

- *Informational Disagreement – i.e., gradual information diffusion (Hong and Stein 1999)*
- *Model Disagreement – i.e., differential interpretation of public signals (Kandel and Pearson 1995)*
- *Some perspectives are a mix between the two – i.e., echo chambers (Cookson, Engelberg and Mullins 2023)*

Horizon disagreement is an important form of model disagreement (*Kondor 2012*)

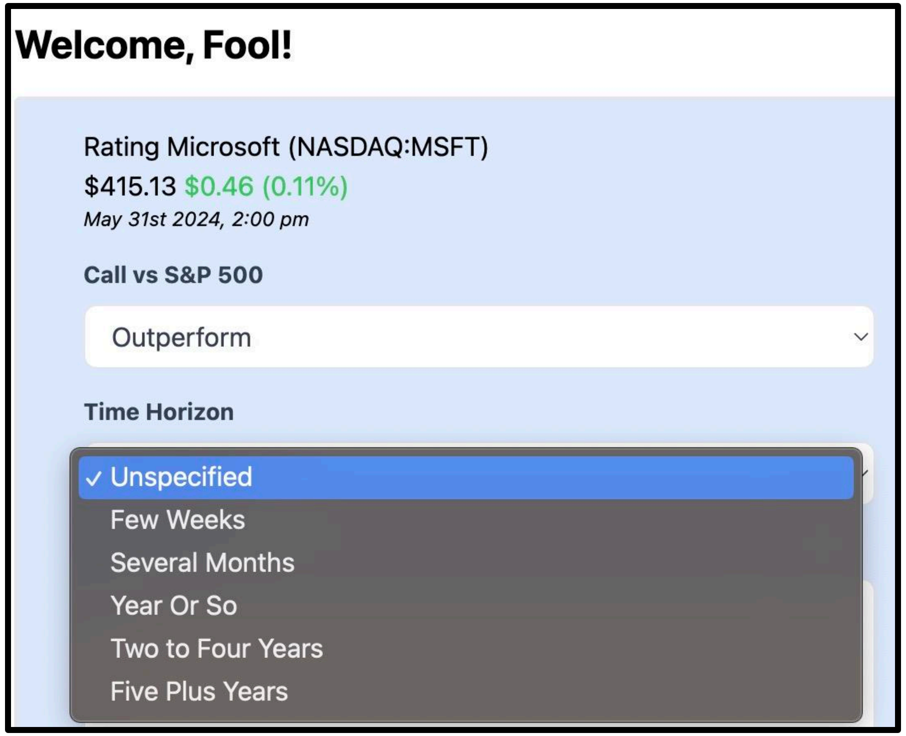
- *Yet, insight into it has been theoretical or indirect.*

What we do

- Data from the social investment platform **CAPS**, part of The Motley Fool.
  - Avery, Chevalier and Zeckhauser (2016) use CAPS data, but not much else in the literature.

- Universe of *stock predictions* made on CAPS.
  - 3.1 million predictions from 2006-2022
  - Each with a **sentiment** --- outperform versus underperform the S&P500
  - And, a **horizon** --- 3-week, 3-month, 1-year, 3-year, 5-year (& unspecified)

- We document horizon-based disagreement around financial events
  - And, we relate horizon disagreement to trading.



What we find

- Predictions at short-vs-long horizons use **distinctive language** in pitches, providing textual validation:
  - Short-horizon pitches use "short term" words and "technical" words more frequently
- *Short- vs long-horizon predictions disagree around major financial events:*
  - **Onset of Covid-19 Pandemic**: **short horizons become pessimistic**
  - **Earnings Announcements** and **Technical View Events**: **short horizons update more** in the direction of the news
  - **Acquisition Rumors**: short horizons become more **optimistic**, while long horizons become **pessimistic**
- Firm-day horizon disagreement relates strongly to abnormal trading
  - Especially on Earnings Announcement days.

Features of CAPS

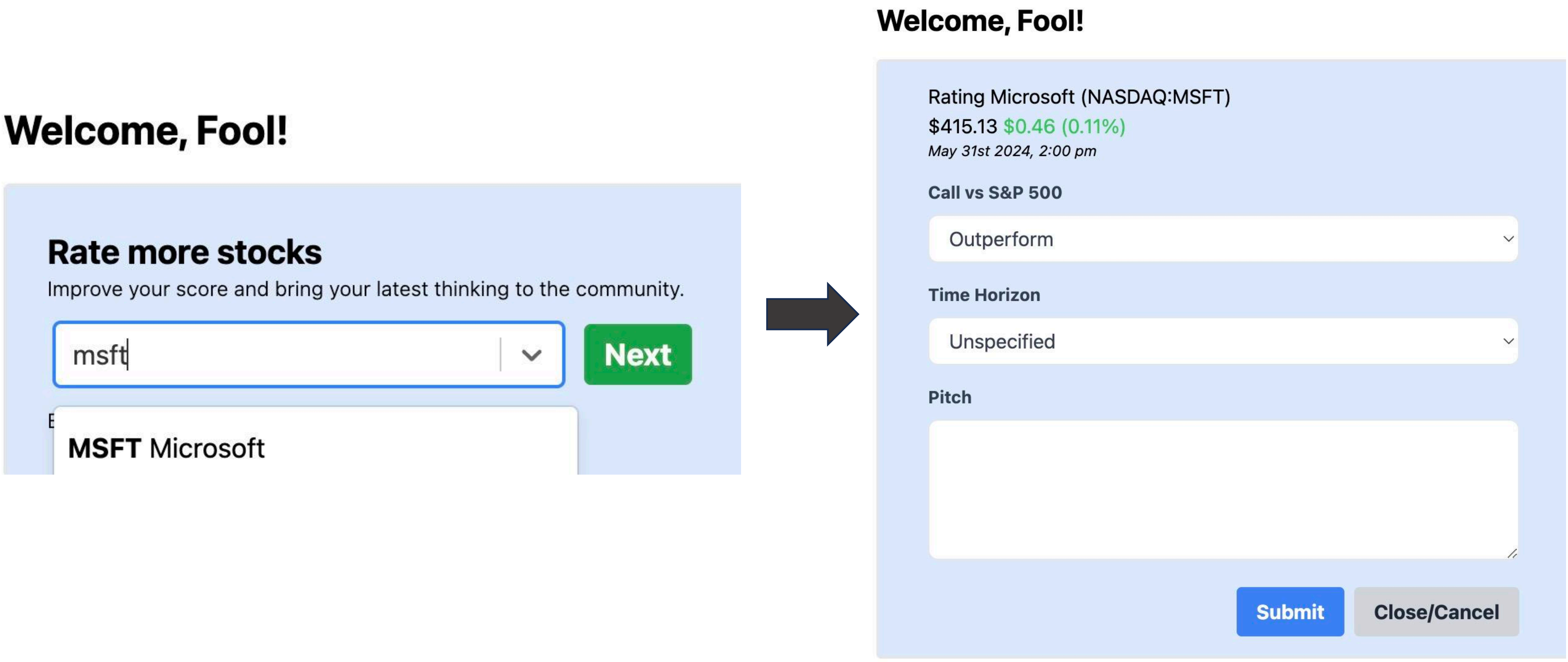
CAPS Sample

All CAPS predictions from 2006 to 2022

- **Sample filters**: Merged with financial data, restricted to users with at least two predictions, stocks in the top 20% of # of predictions
  - **3.1 million predictions** made by **137,750 users** about **1,333 stocks**

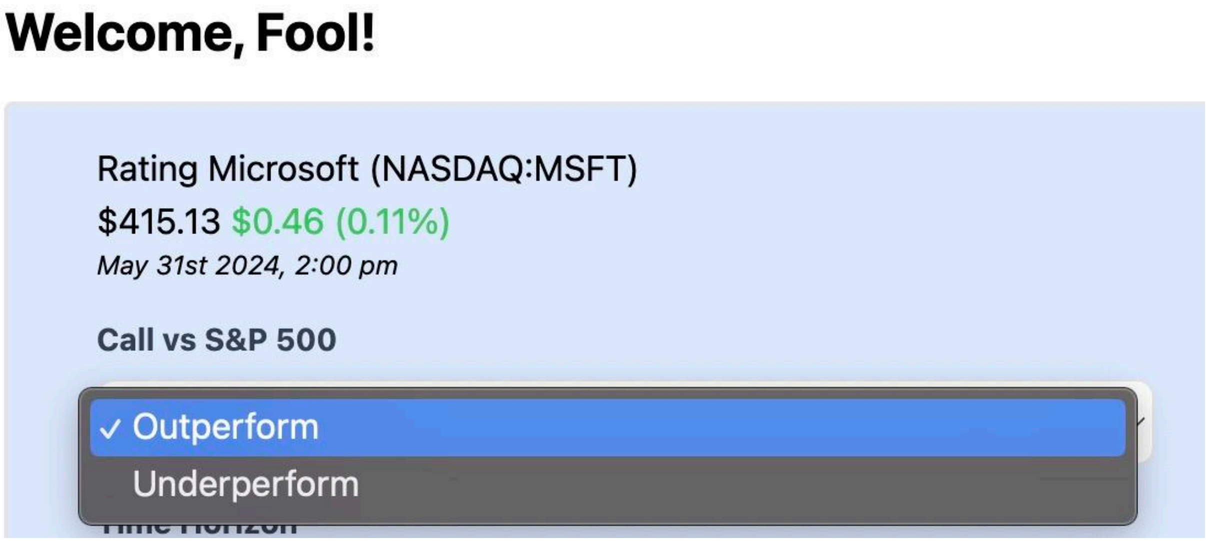
- CAPS users can have only one active prediction per stock at a time
  - <u>Identifiable horizon</u>: Users **must** choose a horizon (*contrast with analysts*)
  - <u>Fresh sentiment</u>: New predictions on date t are not mere reiterations of past sentiment.

CAPS Prediction Interface

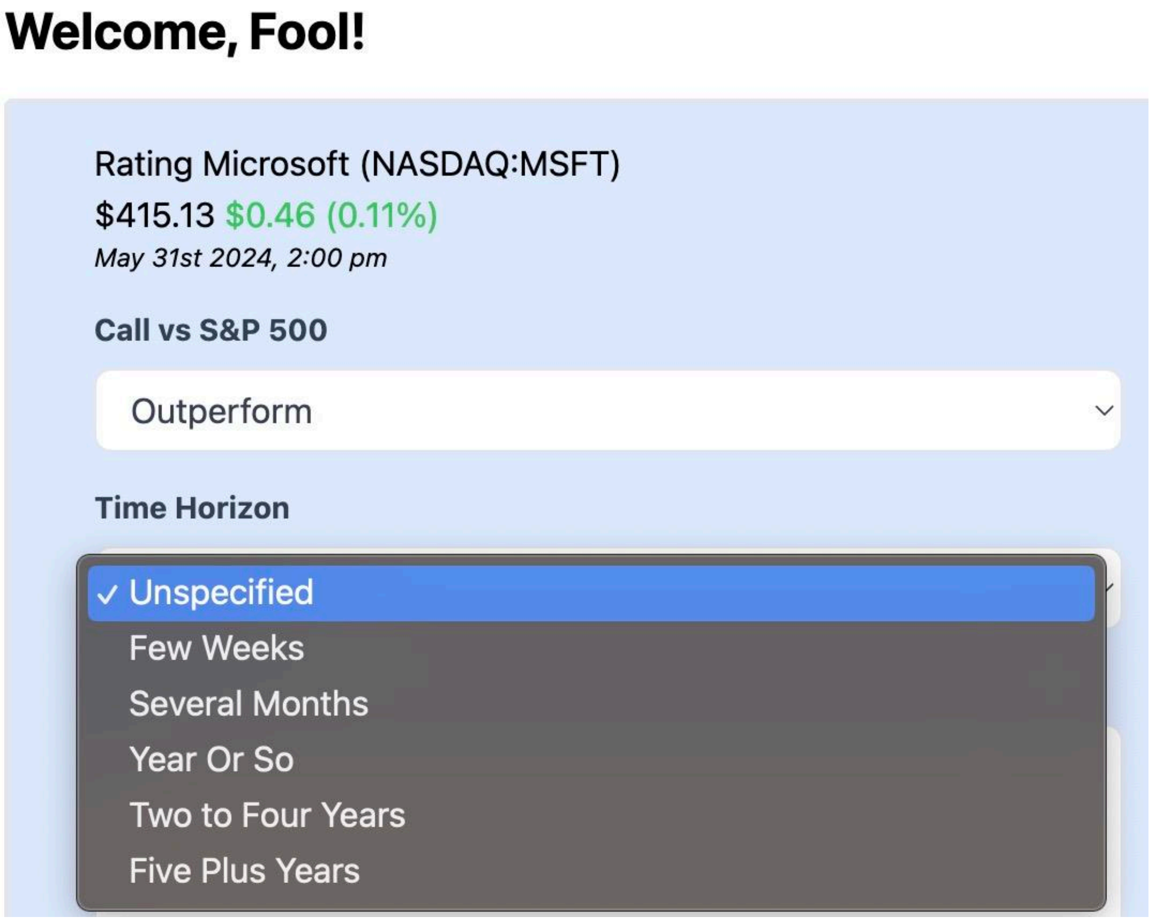


CAPS Prediction Interface

Choosing a sentiment



Choosing a horizon



CAPS Prediction Interface: *What we see*

User scorecards

Start Date	Ticker	Call	Time Frame	Start Price	Today (Change)	Pick Gain	Index Gain	Score	End Date
05-31-2024	CNX		Five Plus Years	$25.85	+$0.45 (+1.74%)	+1.22%	+0.91%	0.31	----
05-29-2024	TER		Five Plus Years	$145.23	$-0.33 (-0.23%)	-4.47%	-0.46%	-4.01	----
05-28-2024	AAPL		Two to Four Years	$189.98	+$0.96 (+0.50%)	+0.64%	-0.83%	1.46	----
05-28-2024	AMZN		Two to Four Years	$180.75	$-2.88 (-1.61%)	-3.36%	-0.83%	-2.53	----
05-28-2024	CRSP		Five Plus Years	$55.24	$-0.32 (-0.59%)	-2.99%	-0.83%	-2.16	----

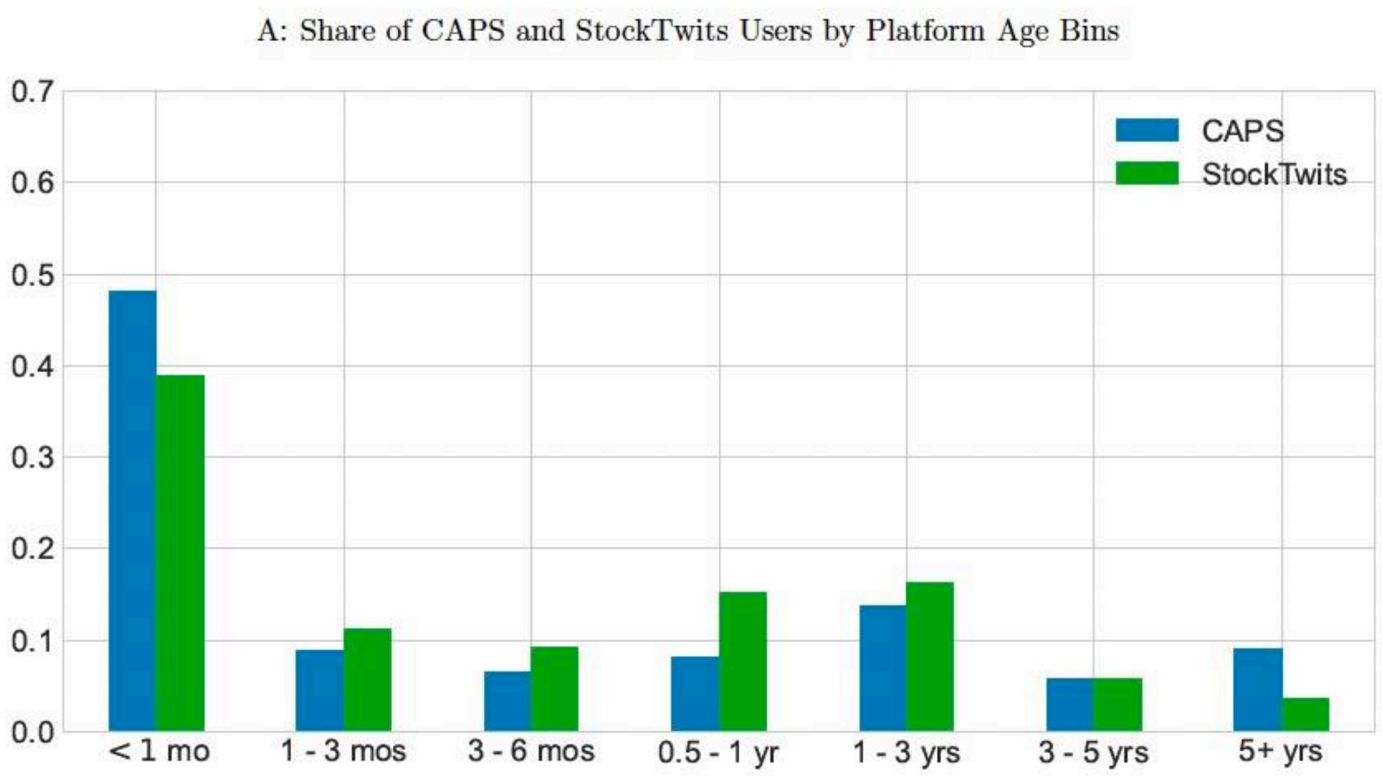
Pitch text

CNX Resources (CNX)

a day ago by Smoker6790**(76.70)** | Reply

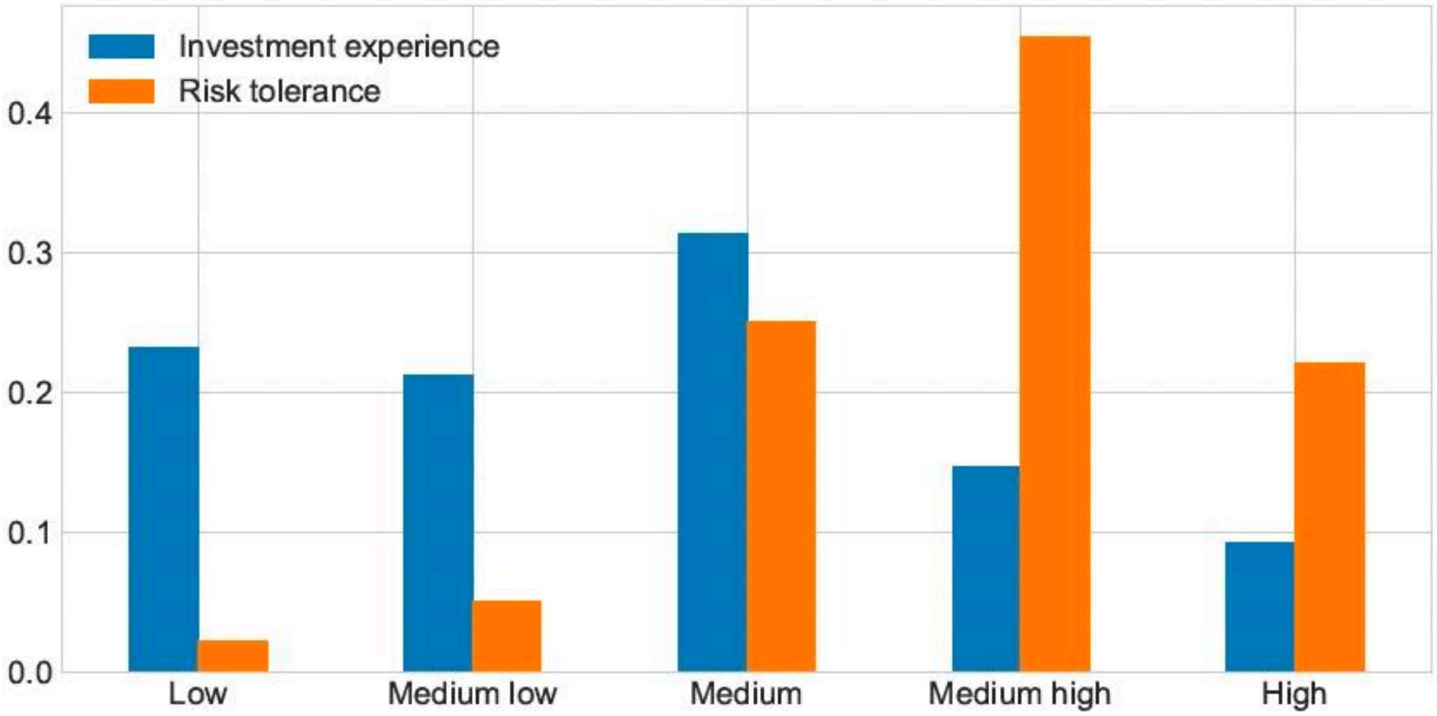
I work at Berkshire Hathaway Energy. We share a facility with CNX. I enjoy the company I work for very much. We always go above and beyond in all aspects. With just my first hand experience I can say that it is a joy to see other energy operators go above and beyond by protecting their assets, the public, environment, and most of all their future. The company has expanded their production of natural gas at a pace that didn't put them at risk. They have maintained a good position in the market an... Read More

Survival on CAPS – similar to StockTwits

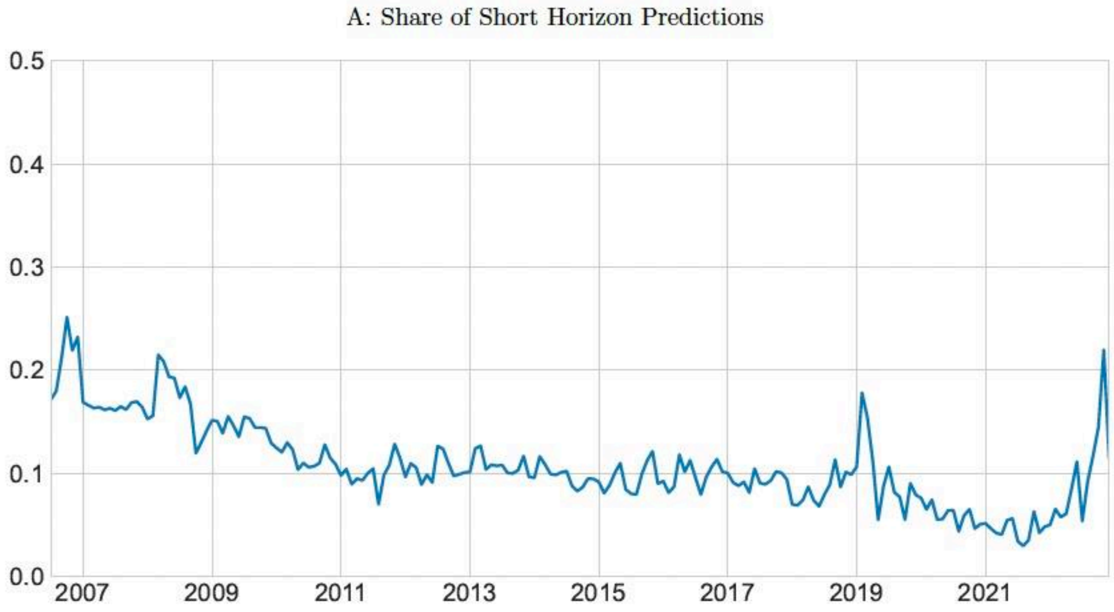


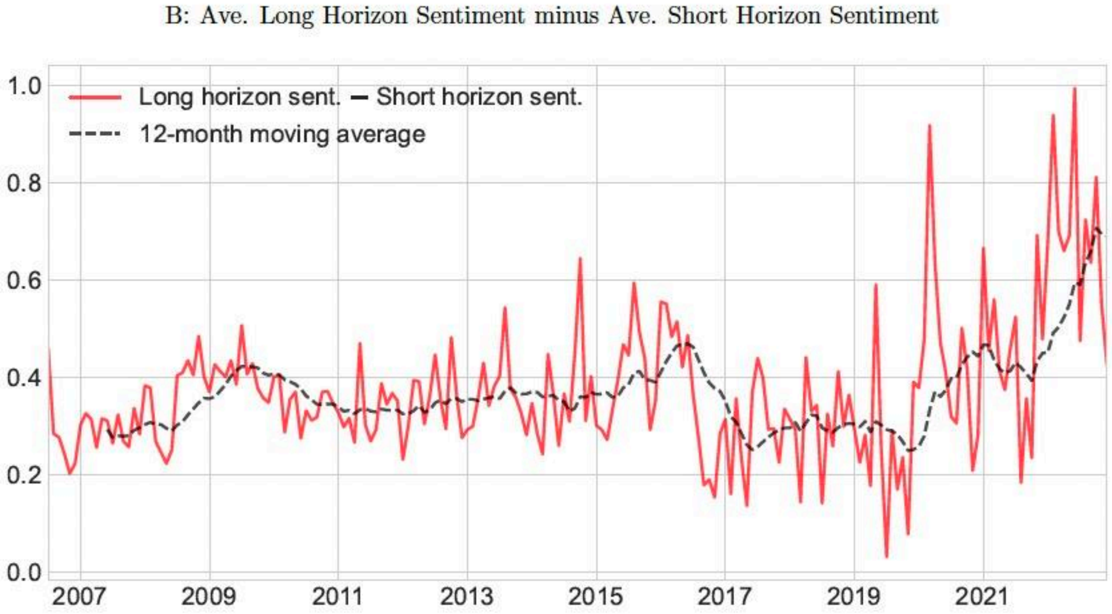
CAPS Users
a mix of experience, risk tolerance

B: Share of CAPS Users by Declared Experience and Risk Tolerance Levels



Time series patterns





Sample characteristics

420k **textual pitches**

Many predictions and pitches across all horizon choices

Lots of **within-user** and **within-stock variation** in horizon choice

Many **stocks** and **predictions per user** (median of 6).

Panel A: Summary of Prediction Horizons					
	# predictions	# pitches	# unique users	# unique stocks	Average sentiment
Total	3,115,187	420,287			
Three-week	163,228	21,250	26,655	1,333	0.300
Three-month	270,461	44,732	52,651	1,333	0.442
One-year	457,311	84,451	75,308	1,333	0.613
Three-year	344,684	69,940	63,903	1,333	0.819
Five-year	1,427,871	167,927	77,402	1,333	0.749
Unspecified	451,632	31,987	35,039	1,333	0.672

Panel B: Summary of Activity					
	Mean	SD	10%	50%	90%
# predictions per user	22.615	95.695	2	6	43
# unique stocks per user	18.143	43.201	2	6	40
# unique horizons per user	2.403	1.344	1	2	4
# predictions per stock	2336.974	2998.534	840	1396	4345

Sample characteristics

Like other social media platforms, users tend to issue **optimistic** ("outperform") **predictions**.

Pitches tend to be *a little longer than tweets* (66 words and 365 characters at the 90th percentile)

Panel C: Summary of Sentiment

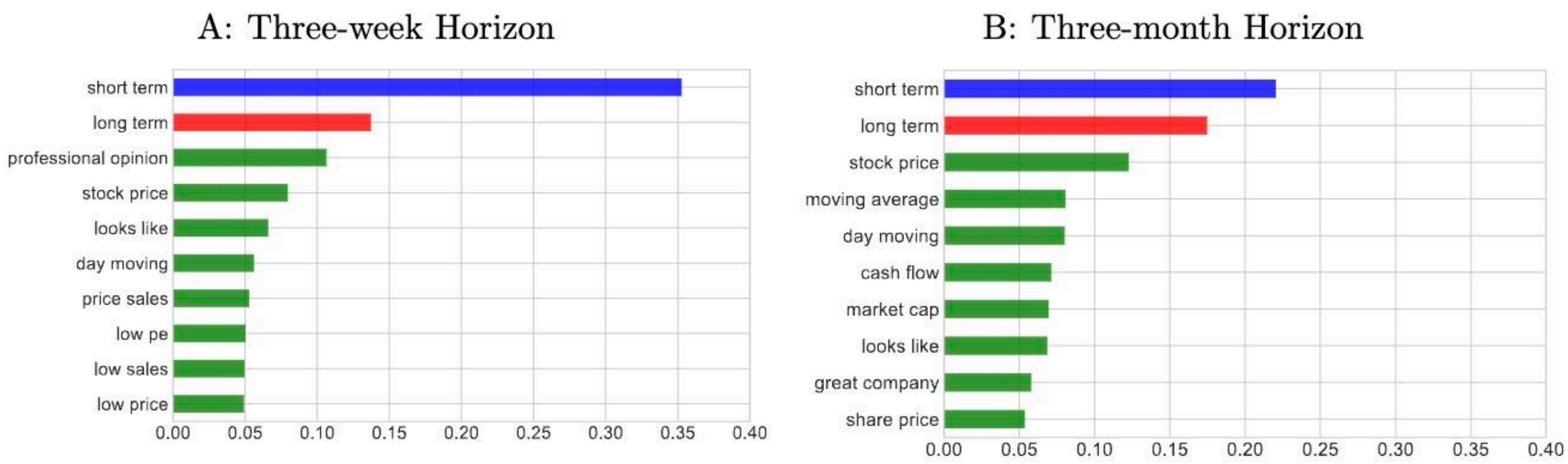
	Mean	SD	10%	50%	90%
All	0.676	0.737	-1.000	1.000	1.000
Per user	0.802	0.370	0.333	1.000	1.000
Per stock	0.661	0.312	0.195	0.783	0.897

Panel D: Summary of Pitches

	Mean	SD	10%	50%	90%
# characters	168	329	14	81	365
# words	30	57	2	14	66
Polarity	0.252	0.436	-0.307	0.226	0.844

Validating Horizons Using Pitch Text

Start with most common bigrams

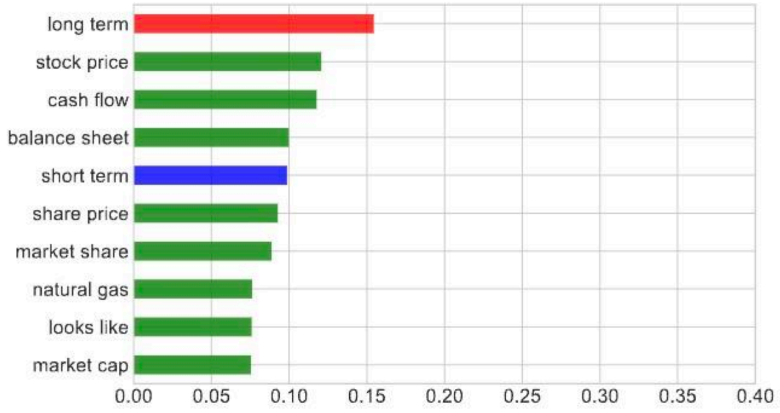


Short-horizon predictions most common bigram: “short term”

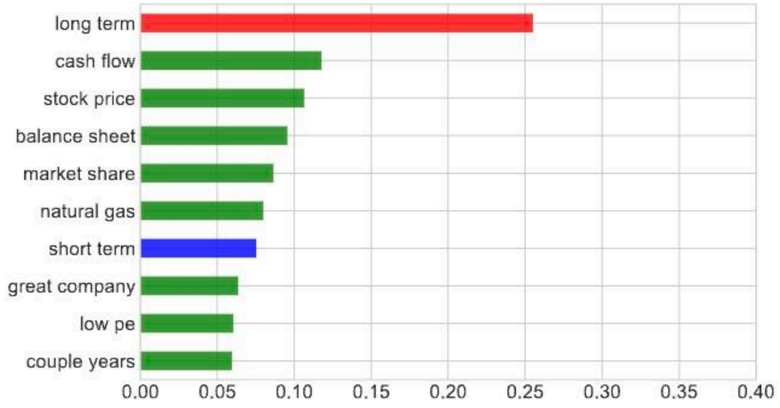
Also: “day moving” and “moving average”

Start with most common bigrams

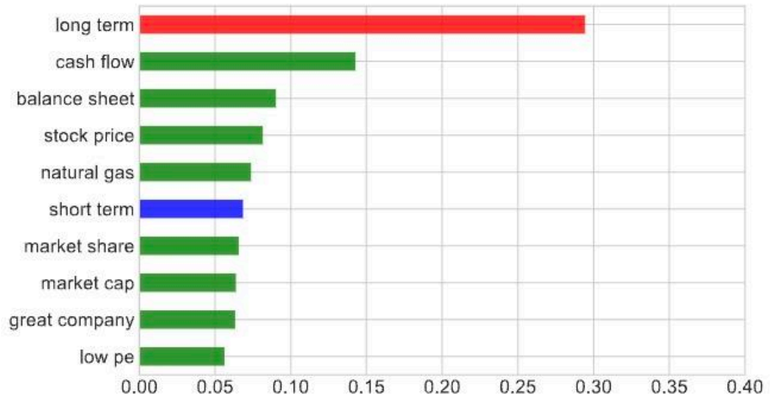
C: One-year Horizon



D: Three-year Horizon



E: Five-year Horizon



Long-horizon predictions most common bigram: “long term”

Also: “cash flow” and “balance sheet”

Are these patterns systematic?

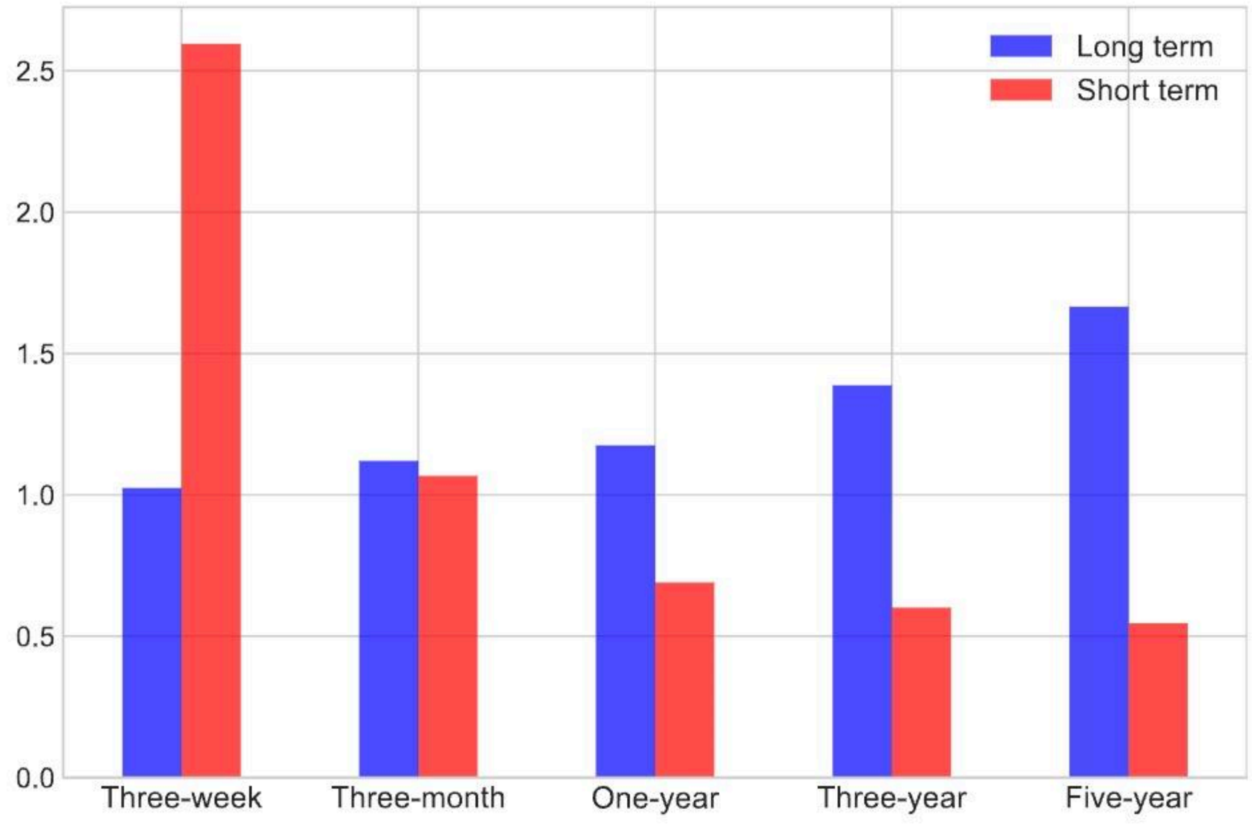
Construct topics using a supervised topic modeling procedure that feeds topic **seed words** to word2vec and uses this to identify the 100 most relevant words.

Topics:

- Short term (e.g., "short term")
- Long term (e.g., "long term")
- Fundamental (e.g., "cash flow")
- Technical (e.g., "moving average")

Then, each pitch has a topic loading for each topic. Regress this loading on pitch's horizon and relevant fixed effects.

Short-term language is more common for short horizons

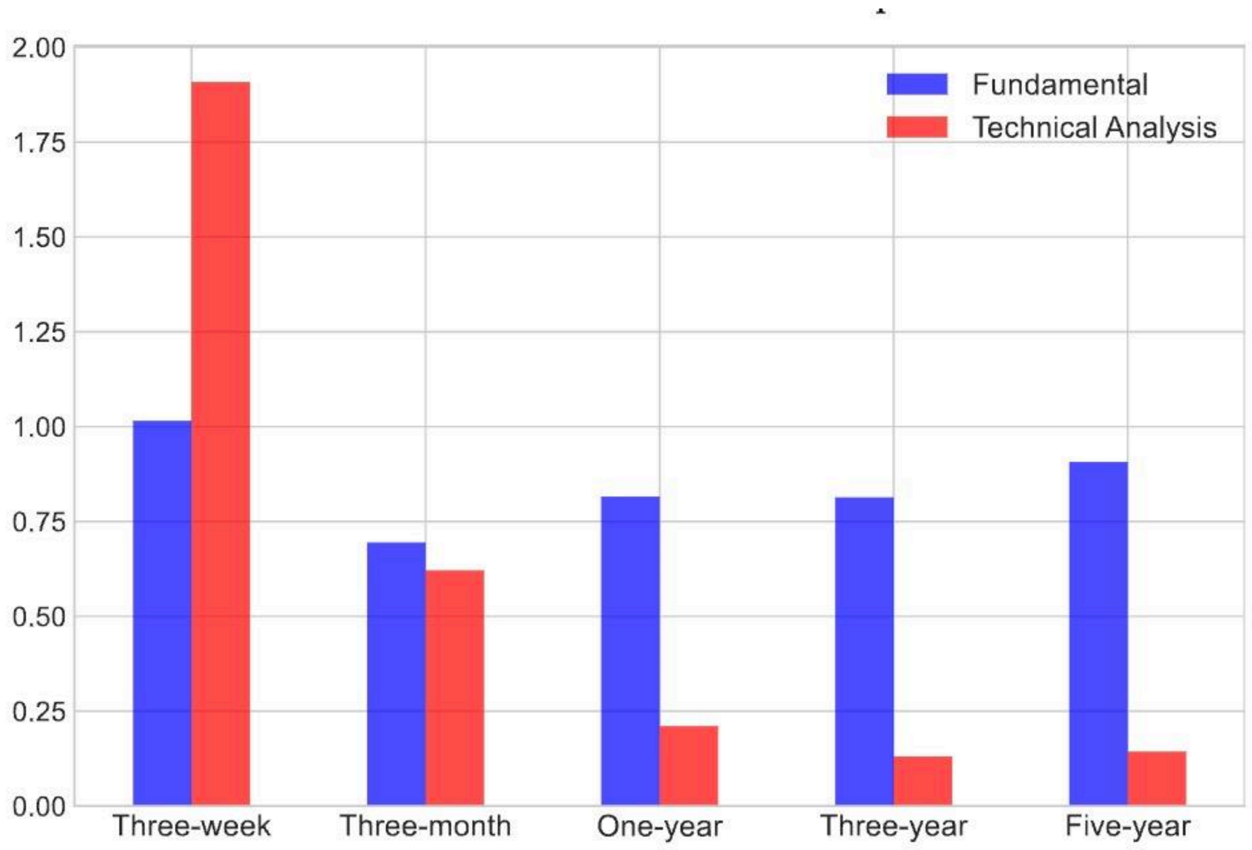


Short-term language in short-term predictions

Even within-user, within-date and within-stock

	Long term − Short term			
	(1)	(2)	(3)	(4)
Constant	-0.283**			
	(0.132)			
Three-month	0.292**	0.117***	0.117***	0.114***
	(0.132)	(0.020)	(0.019)	(0.019)
One-year	0.370***	0.193***	0.193***	0.185***
	(0.132)	(0.021)	(0.020)	(0.020)
Three-year	0.424***	0.256***	0.255***	0.244***
	(0.132)	(0.023)	(0.021)	(0.021)
Five-year	0.484***	0.318***	0.320***	0.298***
	(0.133)	(0.022)	(0.020)	(0.020)
Unspecified	0.418***	0.234***	0.231***	0.219***
	(0.133)	(0.024)	(0.023)	(0.022)
User FE	No	Yes	Yes	Yes
Date FE	No	No	Yes	Yes
Stock FE	No	No	No	Yes
R-squared Adj.	0.012	0.167	0.171	0.175
Obs.	419,579	419,579	419,579	419,579

“Technical” words: more common for short horizons



Also true using user, date and stock FE

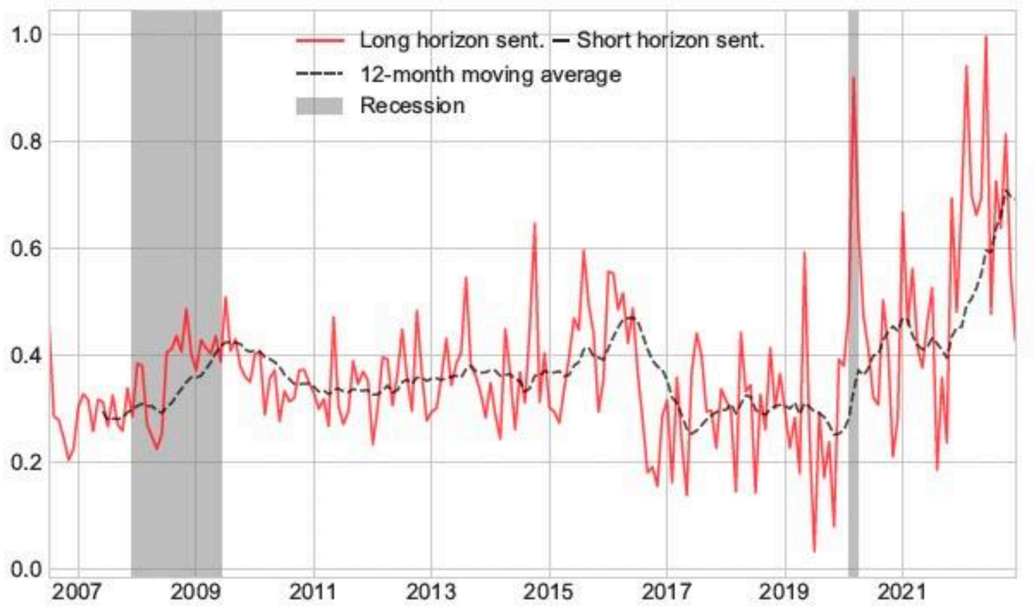
- Later, we control for StockTwits cross-group disagreement

Sentiment Differences Around Events

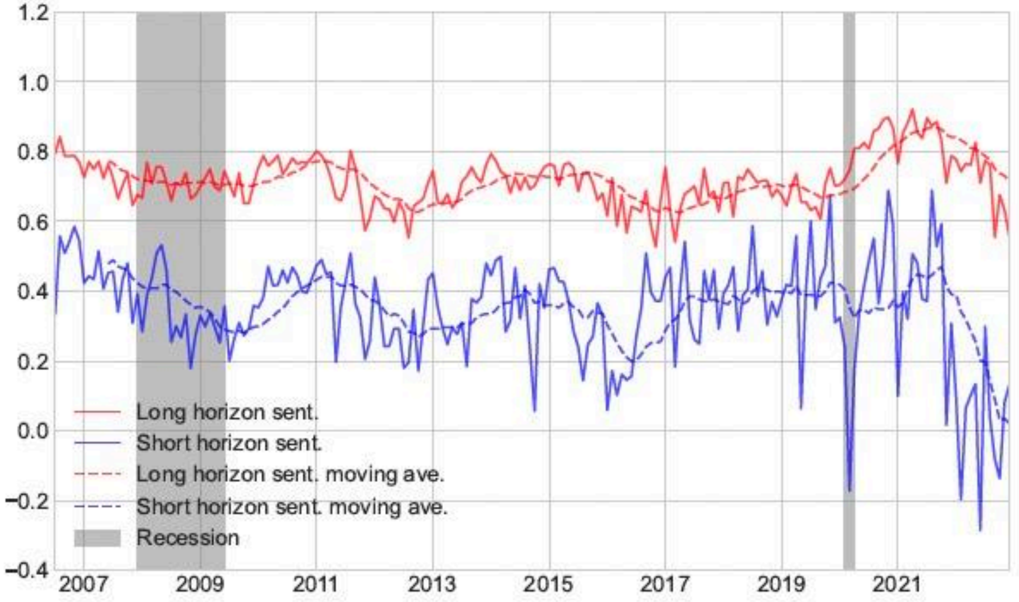
Sentiment updates around Covid-19

On average, short horizon predictions become more pessimistic, while long-horizon predictions maintain optimism.

B: Ave. Long Horizon Sentiment minus Ave. Short Horizon Sentiment



C: Ave. Long Horizon and Short Horizon Sentiment



Sentiment updates around Covid-19

	Outperform Gap			Long Horizon Outperform			Short Horizon Outperform		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Covid	11.489**	4.058		0.101	-4.896*		-11.388**	-8.953*	
	(4.685)	(5.014)		(2.526)	(2.961)		(4.434)	(5.185)	
Covid × High Covid shock ind.		25.058***	20.759***		16.518***	14.215***		-8.541	-6.543
		(5.077)	(4.851)		(4.925)	(5.018)		(6.790)	(7.339)
High Covid shock ind.		0.200	0.586		-0.645	-0.923		-0.845	-1.509
		(0.503)	(0.483)		(1.247)	(1.182)		(1.255)	(1.179)
Date FE	No	No	Yes	No	No	Yes	No	No	Yes
R-squared Adj.	0.007	0.007	0.032	0.055	0.055	0.079	0.026	0.026	0.060
Obs.	177,540	177,540	177,540	177,540	177,540	177,540	177,540	177,540	177,540

Heterogeneity in most affected industries:

- On average, short horizon becomes pessimistic
- In cross-section, long horizon sees opportunity (relatively more **optimistic**)

Sentiment Updates Around Firm Events

Do short-horizon versus long-horizon predictions update differently around financial events?

We estimate the **linear probability model**:

$$Outperform_{i(k,t)}(h) = b_1 \, Technical\ \ View\ \ Ann_{k,t} + \mathbf{\Gamma X} + \alpha_i + \nu_k + \gamma_t + \epsilon_{i,k,t}.$$

X – vector of controls for recent abnormal returns [(-5,-1) and (-25, -6)] and volatility [-5,-1] and dummies for previous week and next week.

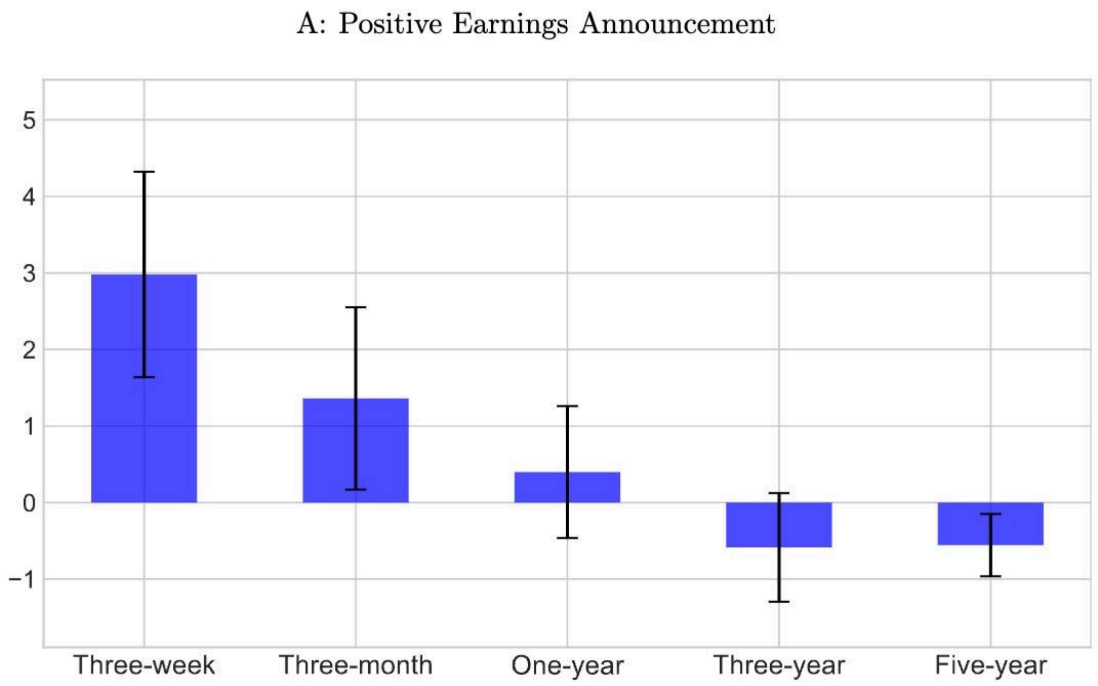
Consider 3 different kinds of *RavenPack* events: **earnings announcements** (SUE+, SUE-), **technical view events** (bullish, bearish) and **merger rumors**.

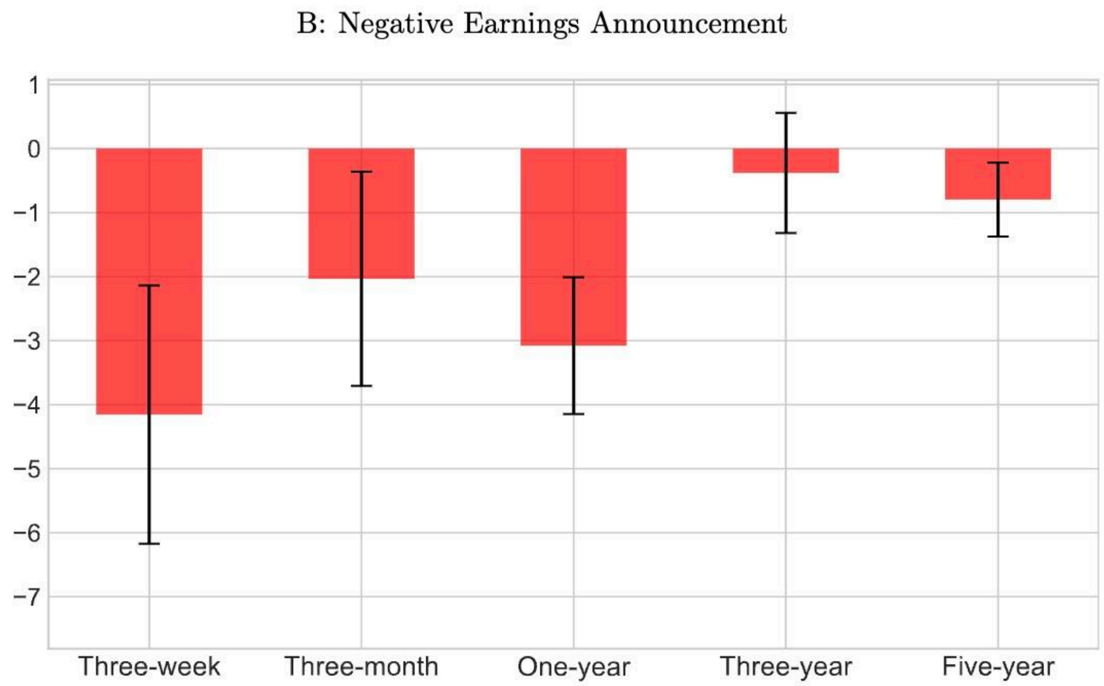
On Earnings Announcement Days

Short-horizon predictions update more in the direction of earnings news than long-horizon predictions.

	Outperform (h)				
	Three-week	Three-month	One-year	Three-year	Five-year
Positive EA	2.965***	1.362*	0.398	-0.561	-0.543**
	(0.818)	(0.723)	(0.524)	(0.433)	(0.248)
Negative EA	-4.178***	-2.025**	-3.057***	-0.378	-0.795**
	(1.226)	(1.016)	(0.651)	(0.573)	(0.351)
Controls	Yes	Yes	Yes	Yes	Yes
User FE	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes
R-squared Adj.	0.365	0.392	0.405	0.347	0.394
Obs.	159,879	264,752	447,208	335,753	1,402,142
% Outperform Predictions	64.98	72.18	80.79	91.03	87.66

On Earnings Announcement Days (graphically)

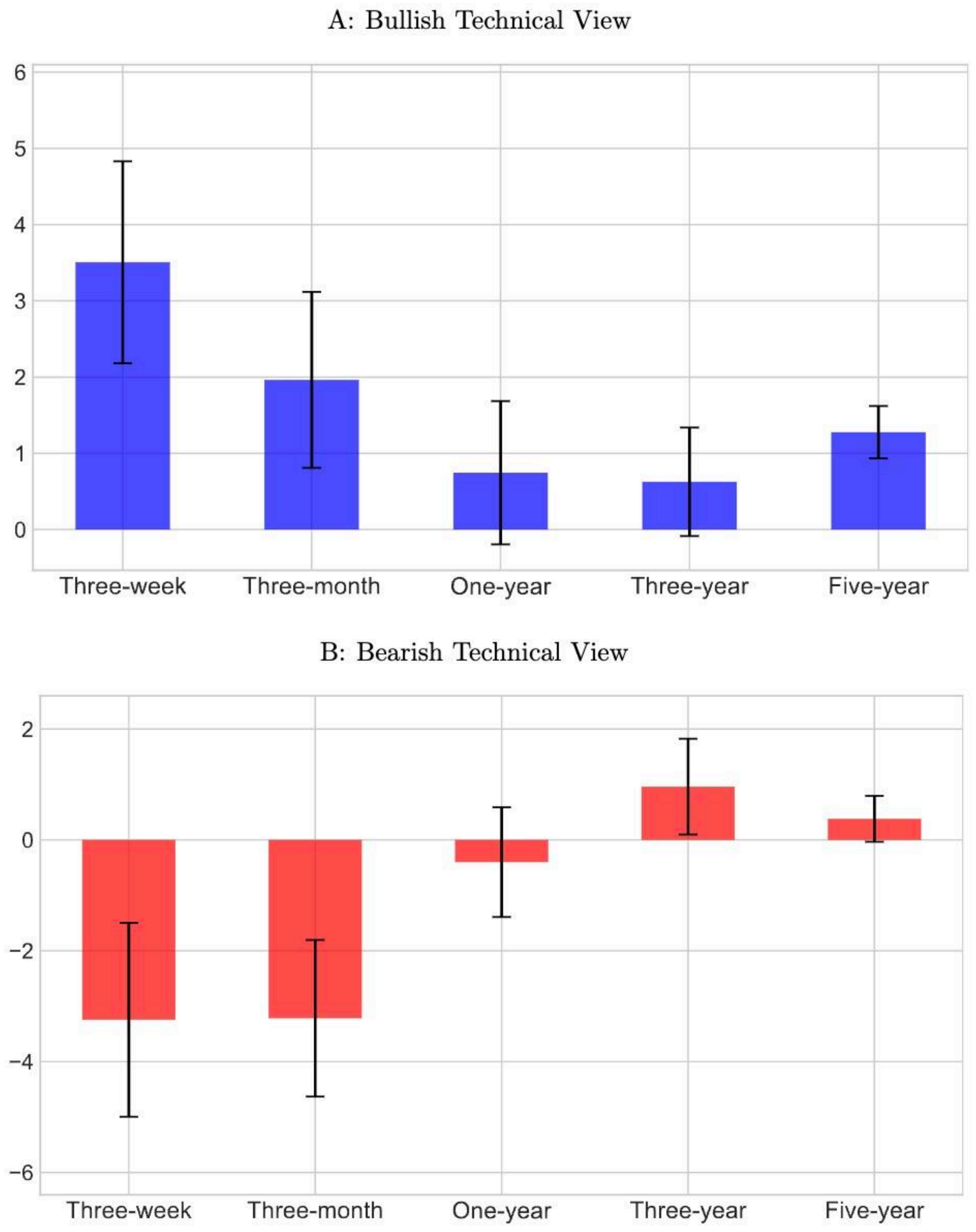




Technical View Days

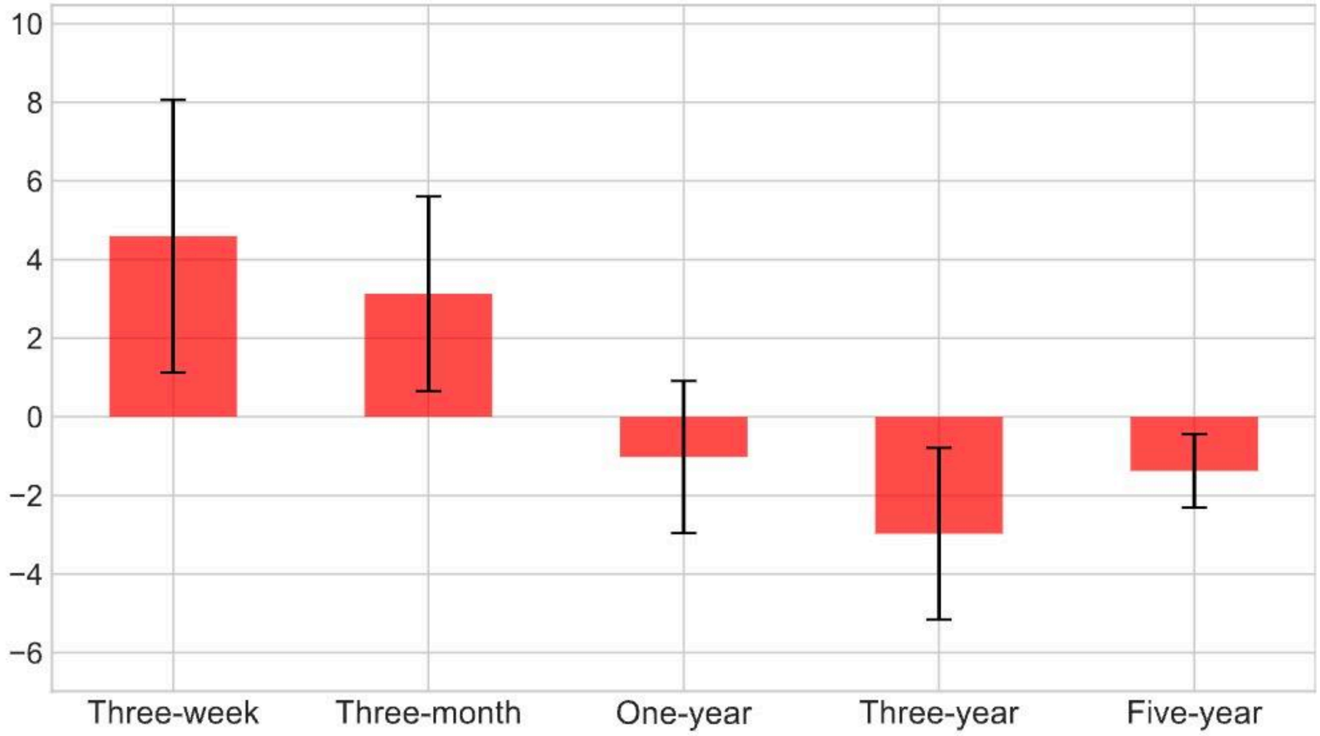
Technical "news" matches direction of updates

Short-horizon predictions (again) update more than long-horizon



Acquisition Rumors (updating about target)

Short-horizon predictions update positively while **long-horizon predictions update negatively** about the target



Horizon Disagreement

Building a horizon disagreement measure

Start with aggregating prediction (i)-level sentiment to stock (k)-day (t)-horizon (h) *average sentin*

$$AvgSent_{k,t,h} = \frac{1}{L}\sum_{l}^{L} Prediction\ Sentiment_{l,k,t,h}, \tag{3}$$

Then, compute **cross-horizon standard deviation** for each stock-day:

$$Horizon\ Disagreement_{k,t} = \sqrt{\frac{\sum_{h \in H} w_{k,t,h}\left(AvgSent_{k,t,h} - AvgSent_{k,t}\right)^2}{\frac{N_H - 1}{N_H}}},$$

Horizon disagreement vs other measures

Novel variation in horizon disagreement

Stronger correlation with StockTwits *within-group* disagreement (Cookson and Niessner 2020) than *cross-group*, but not much R-squared **(~2% of variation explained)**

Panel A: Disagreement Measures							
	CAPS Disagreement	Horizon Disagreement	Three-week	Three-month	One-year	Three-year	Five-year
Mean	0.0298	0.0215	0.0015	0.0029	0.0043	0.0017	0.0109
SD	0.1624	0.1377	0.0384	0.0524	0.0639	0.0396	0.1005

Panel B: Correlation of Horizon Disagreement with Philosophy-based Disagreement		
	CAPS Disagreement	Horizon Disagreement
ST Disagreement	0.1313	0.1019
ST Within-group Disagreement	0.1799	0.1385
ST Cross-group Disagreement	0.0275	0.0213

Horizon disagreement and trading volume

Horizon disagreement correlates with abnormal trading

Not explained by StockTwits measures

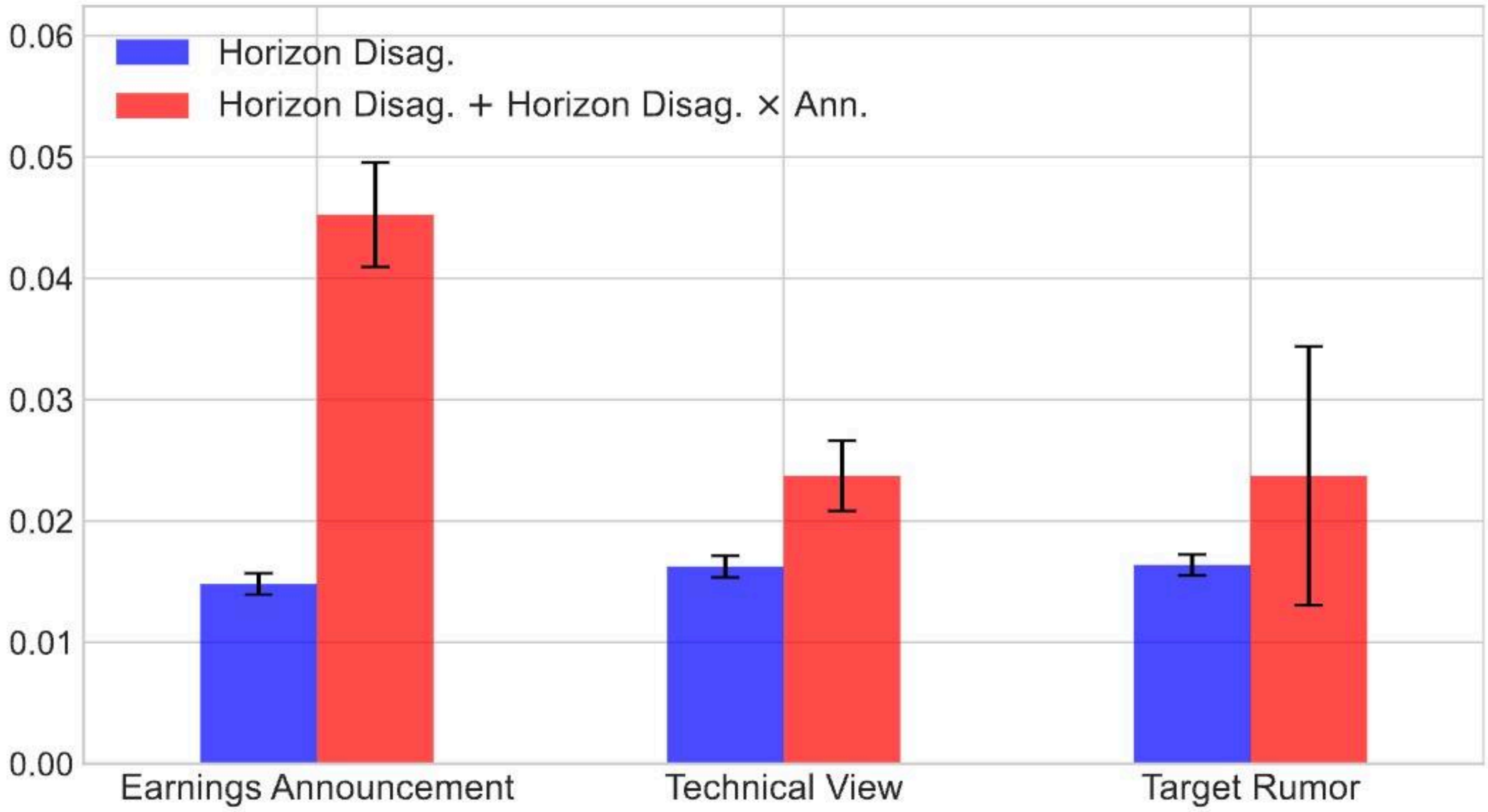
Similar magnitude to cross-group disagreement.

Also t-1 horizon disagreement predicts day t trading.

	Abnormal Log Volume (t)				
	(1)	(2)	(3)	(4)	(5)
Horizon Disagreement (t)	0.016***			0.013***	
	(0.001)			(0.001)	
ST Disagreement (t)		0.087***			
		(0.002)			
ST Within-group Disagreement (t)			0.117***	0.116***	
			(0.003)	(0.003)	
ST Cross-group Disagreement (t)			0.040***	0.040***	
			(0.001)	(0.001)	
Horizon Disagreement (t-1)					0.005***
					(0.000)
Abn. Log Volume (t-1)	0.556***	0.582***	0.573***	0.572***	0.557***
	(0.009)	(0.003)	(0.004)	(0.004)	(0.009)
Abn. Ret (t-25 to t-6)	0.031***	0.020**	0.021**	0.021**	0.031***
	(0.007)	(0.009)	(0.009)	(0.009)	(0.007)
Abn. Ret (t-5 to t-1)	-0.040***	-0.024**	-0.021	-0.019	-0.042***
	(0.008)	(0.012)	(0.013)	(0.013)	(0.008)
Volatility (t-5 to t-1)	0.380***	-0.490***	-0.730***	-0.745***	0.399***
	(0.085)	(0.046)	(0.050)	(0.050)	(0.085)
News Media (t)	0.108***	0.099***	0.087***	0.087***	0.108***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Market Beta	-0.099***	-0.146***	-0.161***	-0.160***	-0.095***
	(0.012)	(0.014)	(0.018)	(0.018)	(0.012)
Log(Mkt. cap.)	0.018	-0.531***	-0.681***	-0.680***	0.024
	(0.023)	(0.049)	(0.063)	(0.063)	(0.023)
Log(Book-to-market)	0.001	-0.002	-0.009	-0.009	0.001
	(0.008)	(0.013)	(0.017)	(0.017)	(0.008)
R-squared Adj.	0.476	0.517	0.528	0.529	0.476
Obs.	4,214,313	1,426,205	1,426,205	1,426,205	4,214,313

Differences in trading around events

Disagreement-trading relation strengthens around earnings announcements, not other events.



Conclusion

- We directly measure horizons and sentiment at the prediction level using CAPS from The Motley Fool.

Findings:

- Sentiment updates differ across horizons around important financial events: earnings announcements, technical view events, and acquisition rumors.

- Motivated by these differences, we measure horizon disagreement & find it has a tight connection to trading, especially around earnings announcements.

- We provide novel evidence of the role of horizons in generating model disagreement

Takeaways: **Social Media as a Lens**

Core argument is that social media **reflects** something deeper

- Does it connect broader economic outcomes?
  - Housing (Bailey et al) State-by-state trade (Bailey et al)
  - Economic opportunity (Chetty et al)
  - Trading (Cookson and Niessner 2020, Cookson, Dim, Niessner 2025)
  - Google searches (Cookson, Engelberg and Mullins 2020)

- Internal consistency and representativeness
  - Effects go back in time (Bailey et al)
  - Textual descriptions match strategies (Cookson and Niessner 2020)
  - Differences by Chinese stocks (Cookson, Engelberg and Mullins 2020)
  - Horizons use long-vs-short language (Cookson, Dim and Niessner 2025)