

KEYSTONE

November 3, 2023

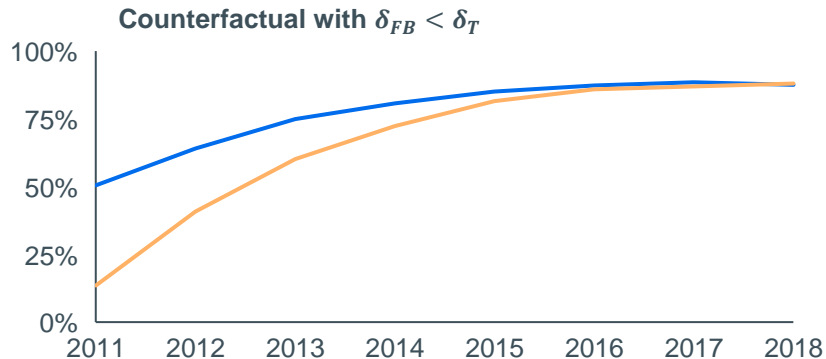
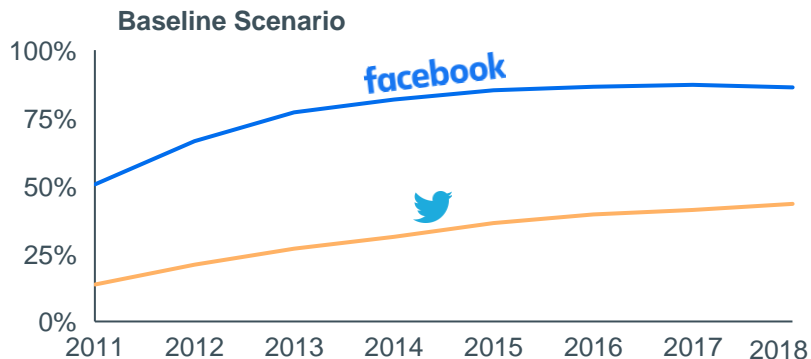
Engagement Substitution on Social Forums: The Case of Meme Stocks on Reddit

KS Research Team

Last year, we explored whether network effects were barriers to entry in markets for other social networks...

We used a **user adoption model** to understand the **drivers of platform growth** (i.e., **network effects**, platform **quality** and **multi-homing costs**) in the context of **social networks** – we found that:

- **Multi-homing costs** are roughly **zero** across the social networks studied (Facebook and Twitter)
- While **network effects** do matter and contribute to growth of leaders (FB) and competitors (Twitter), **platform quality** (δ) – often in the form of product differentiation – is a **key driving factor** behind platform growth

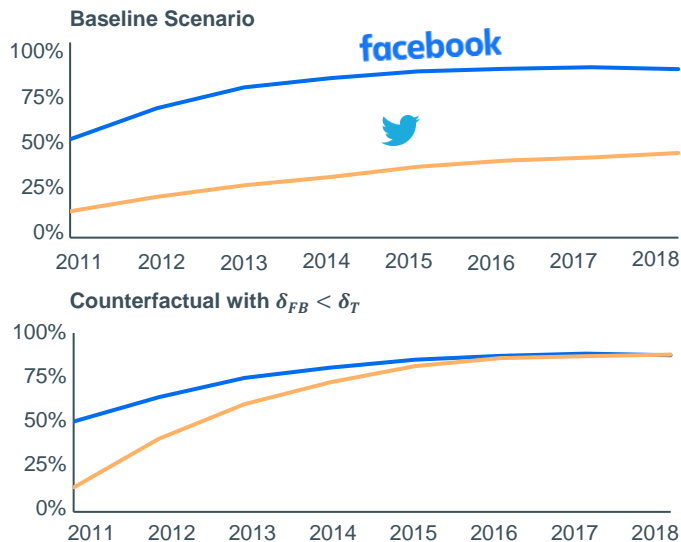


**Counterfactual in graphs: Twitter quality higher than FB quality $\delta_{FB} < \delta_T$, then Twitter can overtake FB when multi-homing costs are zero*

...this year, we're exploring the substitution of engagement on social forums, specifically around the rise of r/wallstreetbets on Reddit

Adoption model

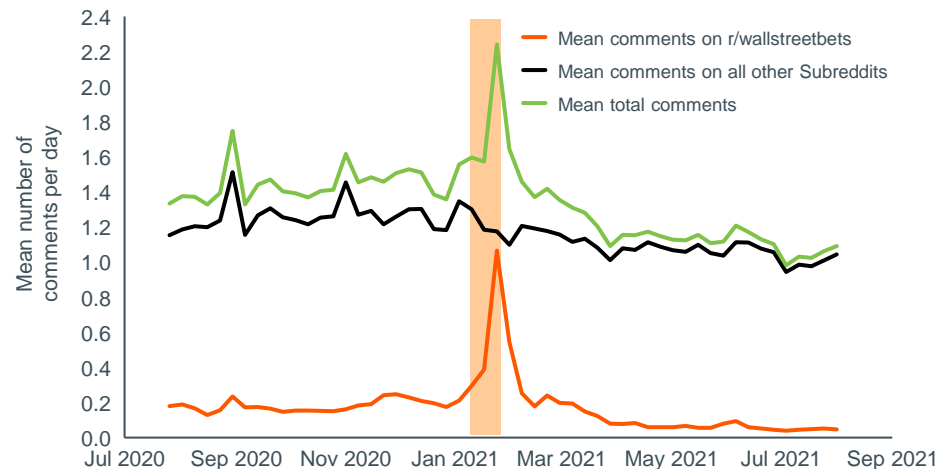
Last year we found that for **social networks**, network effects didn't present a barrier to entry in the context of Facebook



Engagement model

This year we examine how **user engagement** within Reddit changes following the spike in meme stocks and the subsequent increased popularity of WSB

Weekly mean comments among pre-shock commenters
Increase in overall Reddit engagement despite potential substitution



Understanding the drivers of engagement on social forums is important to attract and maintain users and subsequent advertising revenues

Literature shows that some social networks rely on crowdsourced activities to generate value that can be monetized via advertising

- Value on digital platforms (e.g., Reddit) often created by **communities of volunteers** contributing knowledge + **skills for free** (Fershtman & Gandal, 2011; Lakhani & Wolf, 2005; Lerner & Tirole, 2002; Burtch, He, Hong, & Lee, 2020)
 - Crowdsourced activities** include open-source knowledge / tools / software (e.g., Github, Wikipedia)
- Volunteer content **attracts audiences** that can be **monetized via advertising**
- Stronger platforms have more productive volunteer communities (Loh and Kretschmer, 2022)¹
- There's a weak positive relationship between competitive position and activity (Loh and Kretschmer, 2022)
 - Each contributor contributes more on a stronger platform**
 - This relationship is likely due to higher **social benefits** derived from being part of a **larger community**

"Trigger events" start waves of interest in a social network and can drive sustained engagement

- Literature emphasizes the significance of word of mouth (WOM) in social media, which can lead to the phenomenon of social media hype
- WOM is communication between consumers about a product, service, or company in which the sources are considered independent of commercial influence (Litvin et al., 2008)²
- Online WOM communication is known as **electronic word-of-mouth or eWOM** (Yang, 2017)³
- Social media hype** (Vasserman, 2005) is intense online WOM triggered by a key event and sustained by a self-reinforcing engagement.
 - It involves a **trigger event**, followed by **interest waves**, and **sustained interest** across social media platforms (Pang, 2013)⁴

¹ Johannes Loh and Tobias Kretschmer, "Online communities on competing platforms: Evidence from game wikis", 2022

² Litvin, S. W., Goldsmith, R. E., and Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management.

³ Yang, F. X. (2017). Effects of restaurant satisfaction and knowledge sharing motivation on eWOM intentions: the moderating role of technology acceptance factors.

⁴ Augustine Pang, "Social Media Hype in Times of Crises: Nature, Characteristics and Impact on Organizations", 2013

Reddit and the impact of trigger events

- Evidence suggests that platforms success is dependent on 2 dimensions that drive **advertising revenue: user adoption and engagement**.
 - Trigger events can affect both adoption and engagement dimensions.
 - In 2022, Reddit generated most of its revenue from online advertising¹
- Previous KS Meta research found indirect network effects are key to adoption. However, **when volunteer contributions are the main driver of engagement due to higher community social benefits, direct network effects can be key to success**
- Social network platforms like Reddit are organized around online communities whose members discuss various topics, and as such, **rely on crowdsourced activities** to generate value

Existing research offers evidence of how Reddit users behave:

- Users **rarely exhibit significant participation in more than one community**. (Buntain & Golbeck, 2014)²
- Reddit experienced **historic increases in activity during Covid-19** (Veselovsky & Anderson 2023)³
 - Some rises were short lived, as new users on the platform were relatively shorter-lived than previous cohorts
 - But, on the user level Reddit still experiences increases on the number of subreddits and comments per month from Covid-19 because of **older cohorts that became more active**

To explore the impact of trigger events on Reddit user adoption and engagement, we ask two questions:

1. What do trigger events/attention shocks tell us about substitution and complementarity in social engagement?
2. Can a trigger event to a subreddit drive adoption and retention of news users?

¹ <https://www.statista.com/statistics/1310172/reddit-revenues-by-segment/>

The GameStop short-squeeze was attributed to WSB users which triggered increased mainstream media and outsider attention for WSB

Unprecedented growth of meme stocks in Jan. 2021
Stock price of GME and AMC from 2020 to 2022



Google Trends interest over time for r/WSB
Massive spike starting the week of January 25, 2021



Leverage the r/WSB attention shock as a trigger event to examine extent of interest waves for r/WSB and other subreddits (spillover effects), and measure impact on cross and within platform engagement

Keith Gill, a former Director of Financial Wellness Education at MassMutual, had taken an interest in stock trading



Keith Gill

*Marketing and financial education at
MassMutual & CFA charter holder*

In September 2019, he started posting about his gains in GameStop call options on r/wallstreetbets under the username u/DeepF—ingValue



Keith Gill

Marketing and financial education at
MassMutual & CFA charter holder

1.4k

r/wallstreetbets · Posted by u/DeepF—ingValue · gamecock · 4 years ago

Hey Burry thanks a lot for jacking up my cost basis

YOLO

Symbol	Actions	Qty	Price Paid	Last Price	Bid	Ask	Day's Gain %	Day's Gain \$	Total Gain	Total Gain \$	Total Cost	Value \$
Game Jan 15 '21 \$8 Call		1,000	0.53	0.85	0.75	1.25	14.29%	12,500.00	86.69%	46,433.96	53,566.04	100,000.00
06/07/2019	Close Roll	50	0.60	0.85	0.75	1.25	14.29%	625.00	64.97%	1,969.23	3,030.81	5,000.00
06/29/2019	Close Roll	50	0.75	0.85	0.75	1.25	14.29%	625.00	32.25%	1,219.20	3,780.81	5,000.00
07/17/2019	Close Roll	50	0.55	0.85	0.75	1.25	14.29%	625.00	79.80%	2,219.20	2,780.81	5,000.00
07/18/2019	Close Roll	50	0.50	0.85	0.75	1.25	14.29%	625.00	97.57%	2,469.20	2,530.80	5,000.00
08/14/2019	Close Roll	200	0.30	0.85	0.75	1.25	14.29%	2,500.00	227.43%	13,891.77	6,108.23	20,000.00
08/29/2019	Close Roll	100	0.30	0.85	0.75	1.25	14.29%	1,250.00	227.16%	6,943.41	3,056.59	10,000.00
08/22/2019	Close Roll	100	0.55	0.85	0.75	1.25	14.29%	1,250.00	79.97%	4,443.41	5,556.59	10,000.00
08/22/2019	Close Roll	200	0.65	0.85	0.75	1.25	14.29%	2,500.00	52.58%	6,891.77	13,108.23	20,000.00
08/27/2019	Close Roll	100	0.60	0.85	0.75	1.25	14.29%	1,250.00	65.11%	3,943.41	6,056.59	10,000.00
09/03/2019	Close Roll	100	0.75	0.85	0.75	1.25	14.29%	1,250.00	32.33%	2,443.41	7,556.59	10,000.00
Cash												\$13,962.21
Total			\$53,566.04				14.29%	\$12,500.00	86.69%	\$46,433.96		\$113,962.21

397 Comments Share Save Hide Report 100% Upvoted


Total Gain \$	Total Cost	Value \$
46,433.96	53,566.04	100,000.00

Reddit was initially unimpressed, with other retail investors calling him crazy...



Keith Gill

*Marketing and financial education at
MassMutual & CFA charter holder*

 j1187064 · 4 yr. ago 🏆 ❤️ 🐱 🗨️ 2 & 13 More

hahaha

RemindMe! if this stupid f██k lost it all in 1 week



LincolnAtTheTheatre · 4 yr. ago

You're gonna have a hard time dumping these when the stock drops below \$3 this week..



20



love2fukmarriedwoman · 4 yr. ago 🐱 🗨️ 📄 3 & 5 More

Why would anyone play GME IS BEYOND ME

I mean it's a f██k in 3rd party middle man game Seller ok

What F██k N TECH OR INNOVATIONS THEY F██k N HAVE ??? That other retailers don't !!!!

56



Reply Share ...

...but over time, as he held through wild fluctuations in his position, he started to build a following...

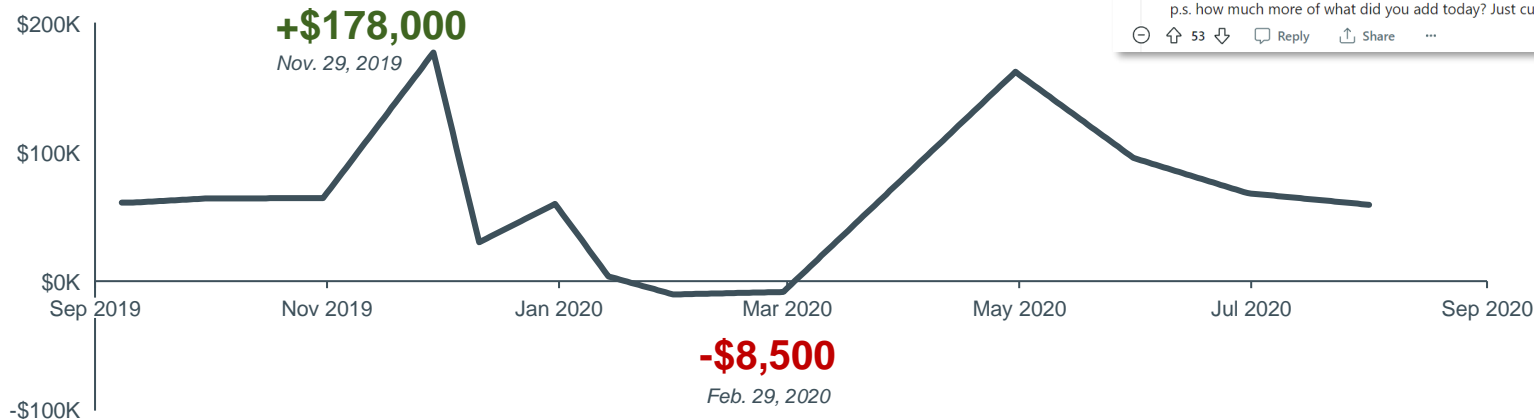


Keith Gill

Marketing and financial education at MassMutual & CFA charter holder

P&L on Keith's GameStop position

Massive fluctuation in value based on GME share price



ArcticGold • 4 yr. ago

I dunno if I should throw in \$2k to just follow u. This is some strong conviction if u are throwing money at it like this. Either a complete degenerate or another legend like u/WSBGod



27



Reply



Share



brutalpancake • 4 yr. ago

To have an opinion on whether OP is right or not is to miss the point entirely. He is what the future is made of. One way or another, he will bury all of us

He is a man of focus, commitment, sheer f[] will

(sorry I couldn't find the gif)

p.s. how much more of what did you add today? Just curious



53



Reply



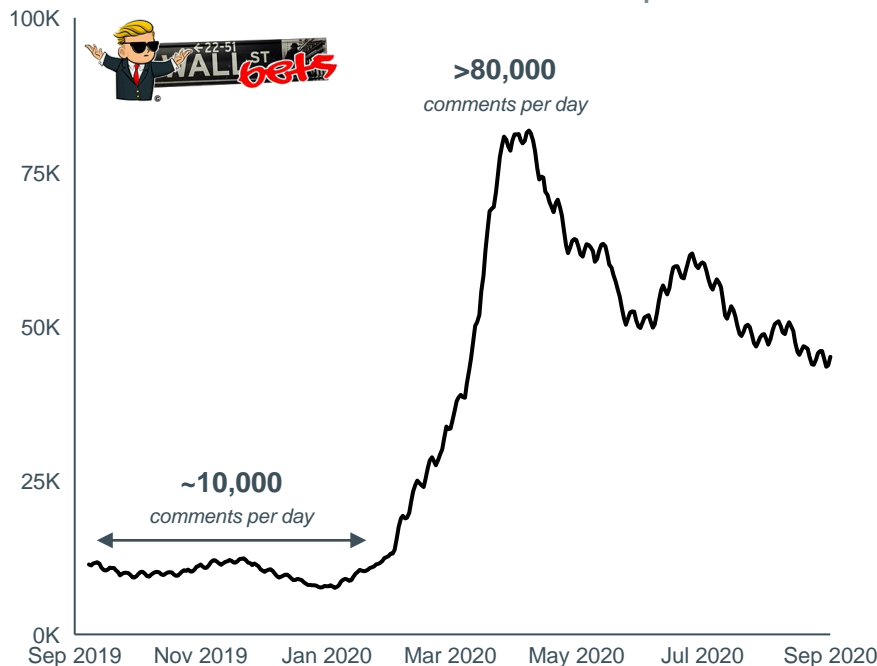
Share



Meanwhile, at the start of COVID, wallstreetbets had started to gain popularity as individual investors started piling into shutdown companies

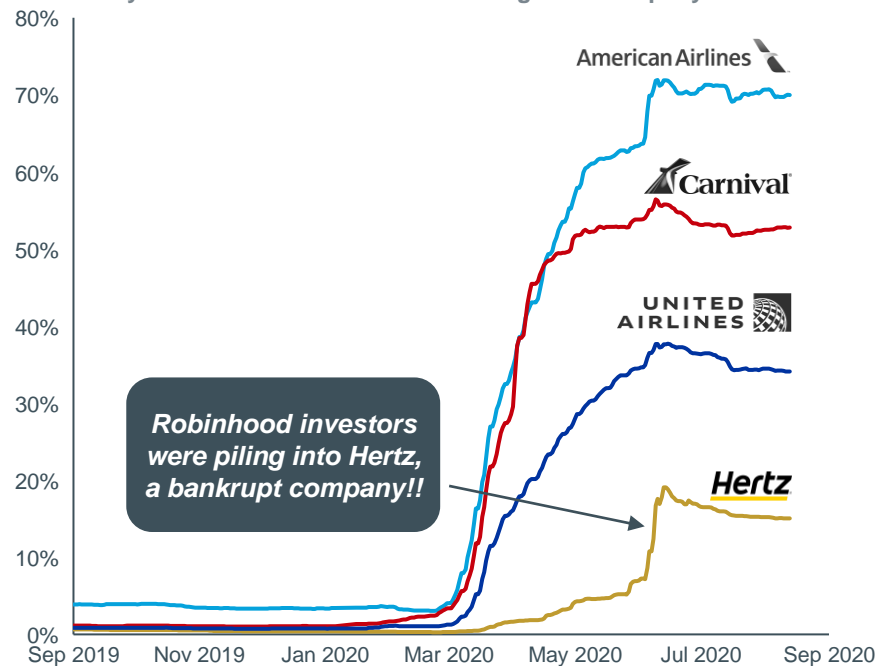
Growth of r/wallstreetbets Reddit comments¹

Massive fluctuation in value based on GME share price



Rampant gambling companies with no operations

Proxy for % of Robinhood users owning each company²



Sources: [Reddit](#).

¹Keystone Analysis, Pushshift API.

²Metric is the % of Robinhood users who own each stock as a percent of the holders of Ford. For was used as it was typically the number one most held stock on Robinhood and was, at times, given out as a free stock.

CONFIDENTIAL

KEYSTONE

By January 2021, the height of the meme stock frenzy had arrived, creating massive Reddit engagement, especially around Keith's posts

January 22, 2021

245%

GameStop up 245% YTD and hits \$65 from \$19 at the start of the year



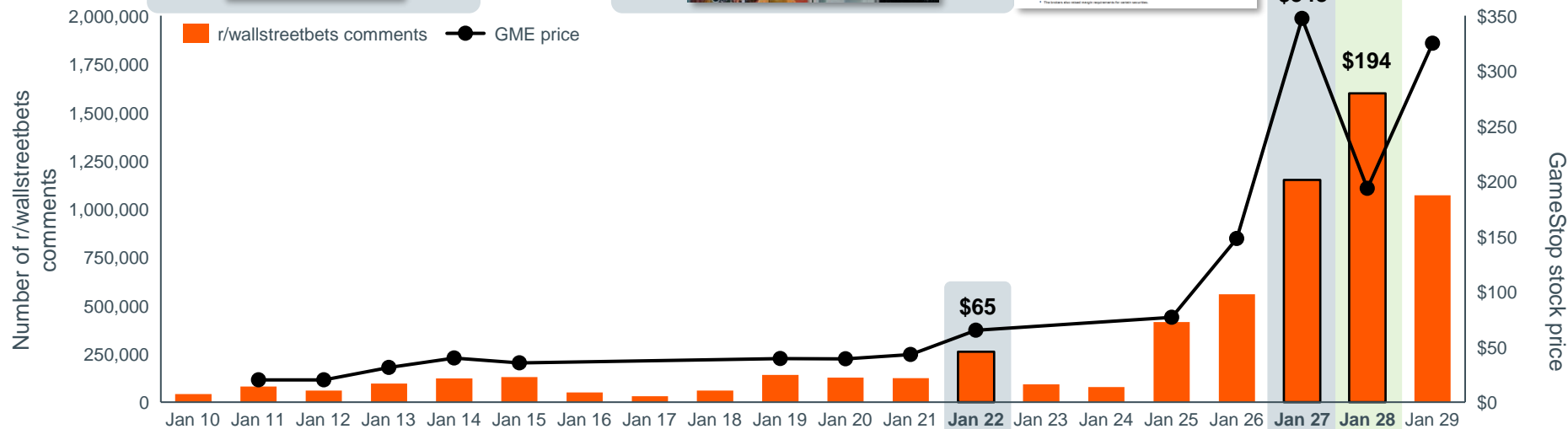
1.2M comments

r/wallstreetbets sets a single day record for Subreddit engagement while GME hits \$348



1.6M

r/wallstreetbets sets a single day record for subreddit engagement as Robinhood restricts trading; degenerates are furious



¹Bloomberg

²Keystone Analysis, Pushshift API

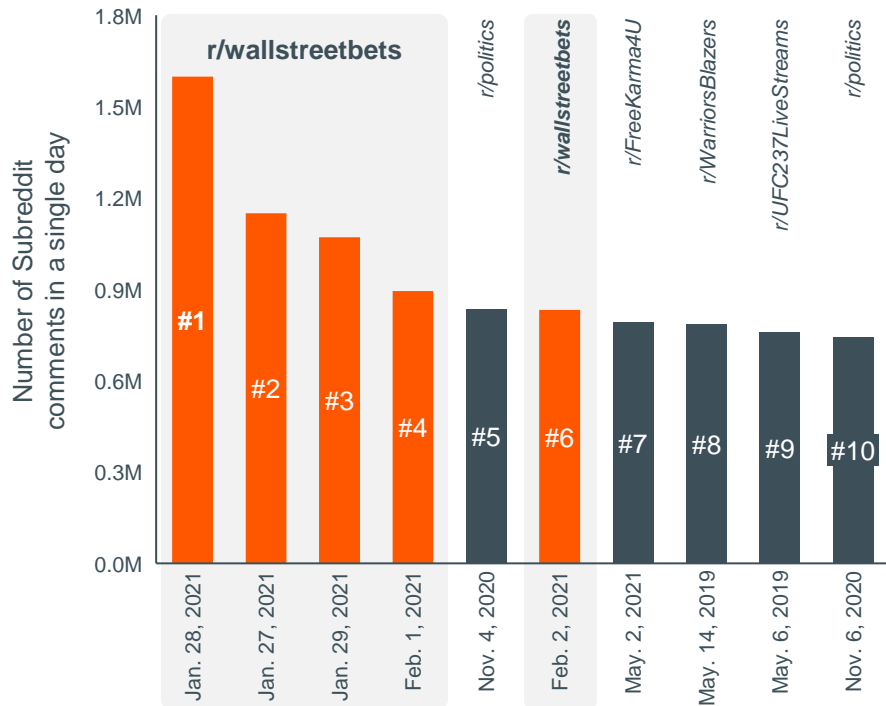
CONFIDENTIAL

KEYSTONE

Overall, the meme stock mania drove unprecedented engagement on Reddit

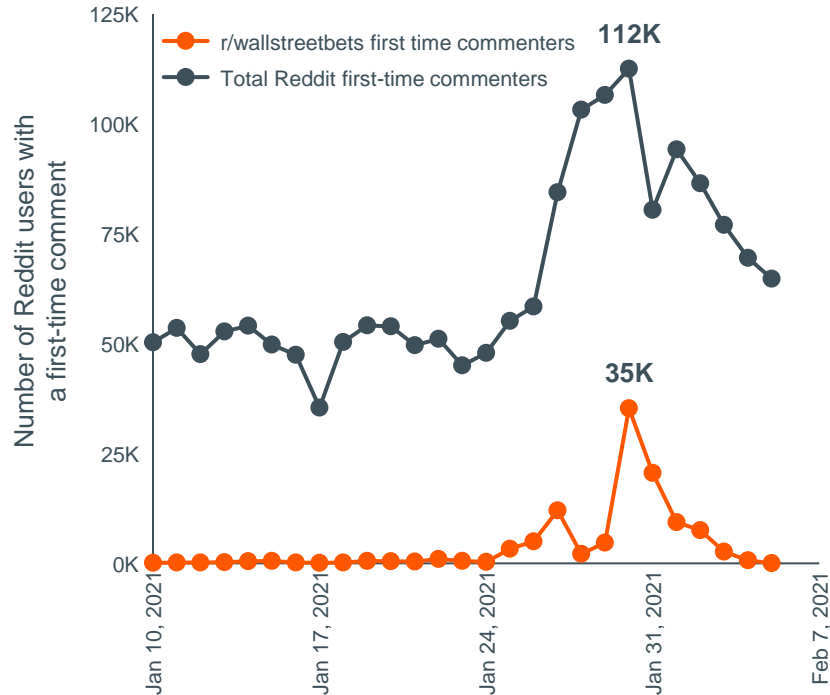
r/wallstreetbets holds the record for most engagement

Ranking of comments on a Subreddit in a single day



Largest number of users joining Reddit in a day

r/wallstreetbets likely drove new users to the platform



CONFIDENTIAL

KEYSTONE

Examine potential changes in adoption and engagement for different groups of users induced by the spike in WSB's popularity

Questions

1. **Effect on changes in engagement:** what are the effects of attention shocks, driving increases in engagement in a component of the platform (i.e., substitutability and complementarity between subreddits)?
2. **Effect on adoption:** can attention shocks to a component of the platform (e.g., subreddit) drive adoption and retention of new users?

Approach

- Capture changes in users' Reddit engagement¹ in the intensive and extensive margins following the WSB attention shock:
 - Measure the extent of engagement substitution from other subreddits to WSB vs WSB's virality complementing engagement to the rest of Reddit
 - Examine the persistence of changes in engagement over time by leveraging the time dimension of the dataset
- Analyze engagement patterns for different user groups that provide distinct insights into the intensive and extensive margins of Reddit engagement, where users are grouped based on
 - Whether they were engaging with Reddit before the shock
 - Whether they were engaging with WSB before the shock

¹We can only capture a user after they have commented at least once. Therefore, a new user refers to users who comment for the first time.

Data used for the analysis effectively cover the entire post and comment history of Reddit, from Pushshift.io



Pushshift.io archives all Reddit comment and post history



Submissions database

- All Reddit submissions
- Includes core fields for analysis, such as
 - Author username
 - Created time
 - Subreddit
 - Title
 - Post text
 - Number of comments
 - Score (net upvotes and downvotes)



Comments database

- All Reddit comments
- Includes core fields for analysis, such as
 - Author username
 - Created time
 - Subreddit
 - Comment text
 - Score (net upvotes and downvotes)

Main dataset used for analysis

Thanks to Data Shared Services, we have easy access to the entire Reddit archive!

In terms of credibility, Pushshift has been used for numerous academic publications across disciplines

The Pushshift Reddit Dataset

Jason Baumgartner,^{1,*} Savvas Zannettou,^{2,©} Brian Keegan,³ Megan Squire,⁴ Jeremy Blackburn^{5,©}
¹Pushshift.io, ²Max-Planck-Institut für Informatik, ³University of Colorado Boulder, ⁴Elon University, ⁵Binghamton University

^{*}Network Contagion Research Institute, [©]iDRAMA Lab
jason@pushshift.io, szannett@mpi-inf.mpg.de, brian.keegan@colorado.edu, msquire@elon.edu, blackburn@cs.binghamton.edu

Abstract

Social media data has become crucial to the advancement of scientific understanding. However, even though it has become ubiquitous, just collecting large-scale social media data involves a high degree of engineering skill set and computational resources. In fact, research is often times gated by data engineering problems that must be overcome before analysis can proceed. This has resulted in recognition of datasets as meaningful research contributions in and of themselves. Reddit, the so called “front page of the Internet,” in particular has been the subject of numerous scientific studies. Although Reddit is relatively open to data acquisition compared to social media platforms like Facebook and Twitter, the technical barriers to acquisition still remain. Thus, Reddit’s millions of subreddits, hundreds of millions of users, and billions of comments are at the same time relatively accessible, but time consuming to collect and analyze systematically. In this paper, we present the Pushshift Reddit dataset. Pushshift is a social media data collection, analysis, and archiving platform that since 2015 has collected Reddit data and made it available to researchers. Pushshift’s Reddit dataset is updated in real-time, and includes historical data back to Reddit’s inception. In addition to monthly dumps, Pushshift provides computational tools to aid in searching, aggregating, and performing exploratory analysis on the entirety of the dataset. The Pushshift Reddit dataset makes it possible for social media researchers to reduce time spent in the data collection, cleaning, and storage phases of their projects.

1 Introduction

Understanding complex socio-technical phenomena requires data-driven research based on large-scale, reliable, relevant data sets. Web data, particularly data from application programming interfaces (APIs), has been an enormous boon for researchers using online social platforms’

crisis informatics (Paley). Major scandals on media platforms like Facebook (Walker, Mercea, et al.) have shown the ability for researchers to instruct students, and re-

This “post-API age” is a data resources used for research (Puschmann 2019), increasingly based on social, technical (Crawford 2012; Manov et al. 2016) around violating research (Halavais 2016) had a profoundly chilling API-derived data to it. tion, harassment, radicalization. Furthermore, re- atically studying the role design affordances, and taining these forms of “i- Keegan 2018). Faced i protecting their users’ d commitments to values are exploring alternative third party” models that reputational risks (Brun Mervis 2019; Puschmar

Even if the “golden age” of social science and social in the shadow of privacy characterized by enormous tion and inequalities in : 2019), ethically-suspect 2016; Tufekci 2014; O

The Migrants of Reddit: An Analysis of User Migration Effects of Subreddit Bans

Bryce Cal
Stanford University
Stanford, CA 94305
bcal@stanford.edu

Sean Decker
Stanford University
Stanford, CA 94305
skdecker@stanford.edu

Crystal Zheng
Stanford University
Stanford, CA 94305
czheng11@stanford.edu

Abstract

As “the front page of the internet,” Reddit in recent years has started to crack down on “toxic” behaviors, including hate speech, bullying, etc. The primary method by which Reddit is combating this is by the banning of subreddits whose users exhibit these behaviors. However, differentiating subreddits that should be banned from others is a hard problem, and with new subreddits often immediately replacing banned subreddits, there is a need to make this process of identifying subreddits that should be banned more automated. In this paper, we propose using not only subreddit features for classifying if it is to be banned, but also the network structure of Reddit. We compare a simple logistic regression model of classification with a local classifier and discuss the useful network characteristics of the Reddit cross-link network for this classification problem.

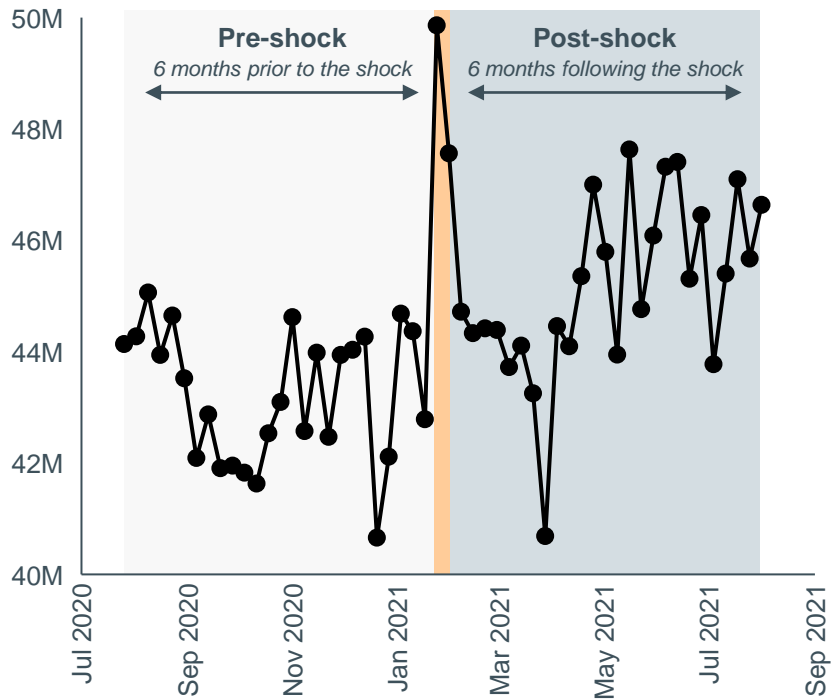
1 Introduction

In 2012, then-Reddit CEO Yishan Wong famously said, “We stand for free speech. This means we are not going to ban distasteful subreddits. We will not ban legal content even if we find it odious or if we personally condemn it.” Since then, however, Reddit.com has grown into the “front page of the internet.” As of 2018, Reddit had approximately 330 million monthly active users and more than 138,000 active subreddits, making it one of the most popular online social platforms in the world (1). Previous work has largely focused on characterizing inter-community interactions, individual user behaviors, and conflicts of Reddit as a whole or individual subreddits. However, investigating banned vs. non-banned subreddits in terms of network characteristics has not been previously researched. With the expansion of Reddit’s reach and power has come renewed controversy around how Reddit moderates content. In 2019, to fight against hate speech, bullying, and many other behaviors that are seen as “toxic,” Reddit’s main tool is now banning subreddits. Given the nature of Reddit, oftentimes it is difficult to identify which subreddits may or may not be banned currently or in the future.

To examine changes to user behavior over time, we examine three core periods, pre-shock, shock, and post-shock

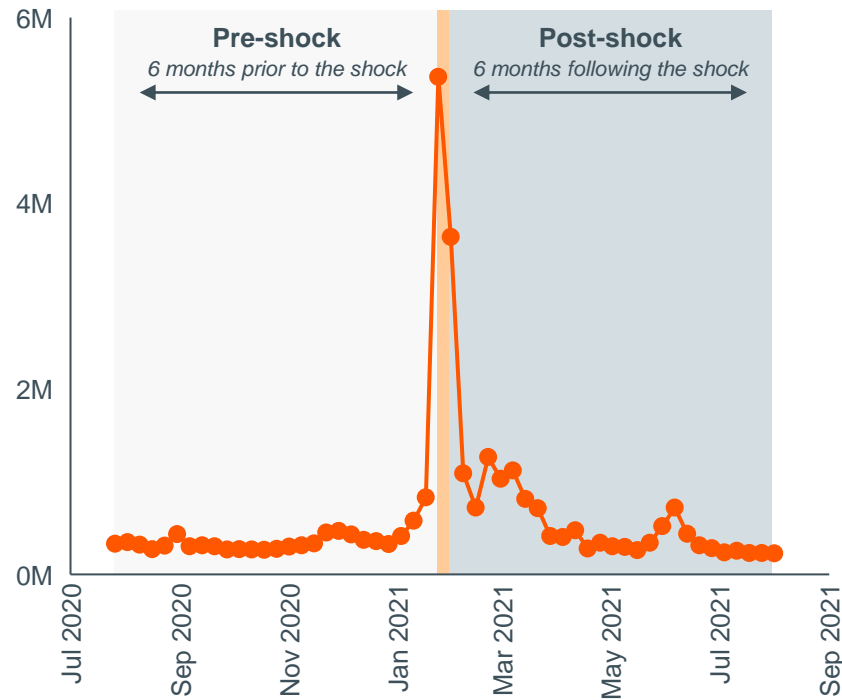
Total Reddit comments over sample period

Spike of comments during the shocked period



Total r/wallstreetbets comments over sample

Major spike during the shocked period



CONFIDENTIAL

Shock period

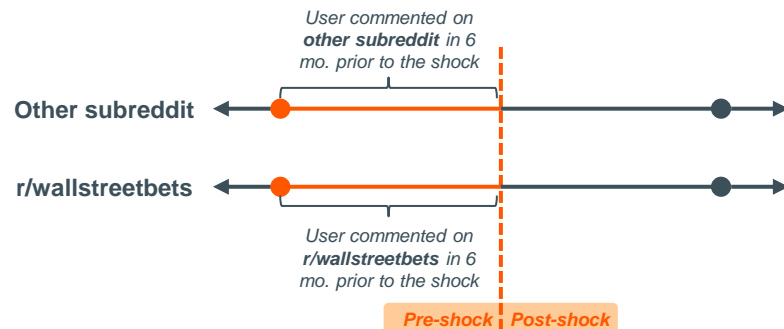
KEYSTONE

We can observe two sets of existing Reddit users of interest based on when they commented on r/WSB and other Subreddits

Group A: Pre-shock r/wallstreetbets commenters

Two conditions:

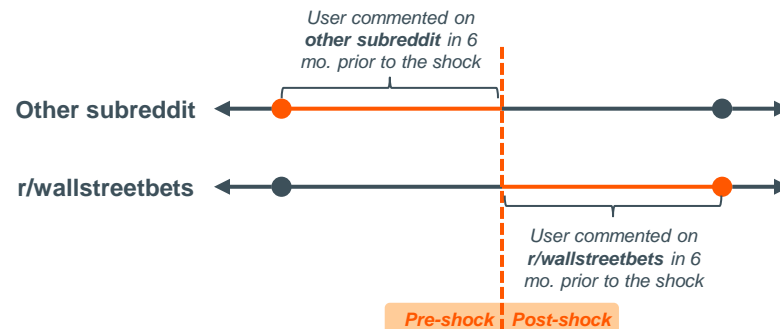
1. Pre-shock other subreddit comments: the user must have commented on any subreddit other than r/wallstreetbets in the 6 months **prior to the shock**
2. Pre-shock r/wallstreetbets comments: the user must have commented on r/wallstreetbets in the 6 months **prior to the shock**



Group B: Post-shock r/wallstreetbets commenters

Two conditions:

1. Pre-shock other subreddit comments: the user must have commented on any Subreddit other than r/wallstreetbets in the 6 months **prior to the shock**
2. Post-shock r/wallstreetbets comments: the user must have commented on r/wallstreetbets in the 6 months **after to the shock**



Capturing expected engagement and complementarity/substitutability

Regression discontinuity (RD) in time: Compare user i engagement pre shock ($t = 0$) to the engagement during/after the shock ($t = 1$)

$$Y_{i,R,t} = \alpha + \beta \cdot NonWSB + \gamma \cdot (1_{t=1}) + \delta \cdot NonWSB \cdot (1_{t=1}) + \varepsilon_{i,R,t}$$

where $Y_{i,R,t}$ is user i engagement in period t in subreddit R and

$$NonWSB = \begin{cases} 0 & \text{if } WSB \\ 1 & \text{if Rest of Reddit} \end{cases}$$

Identifying assumption: Only exogenous shock in WSB's popularity changes during the sample period

Parameter	Description
α	Pre-shock average WSB comments
$\alpha + \beta$	Pre-shock average non-WSB comments
γ	Change in average WSB comments
$\gamma + \delta$	Change in average non-WSB comments

Expected user engagement

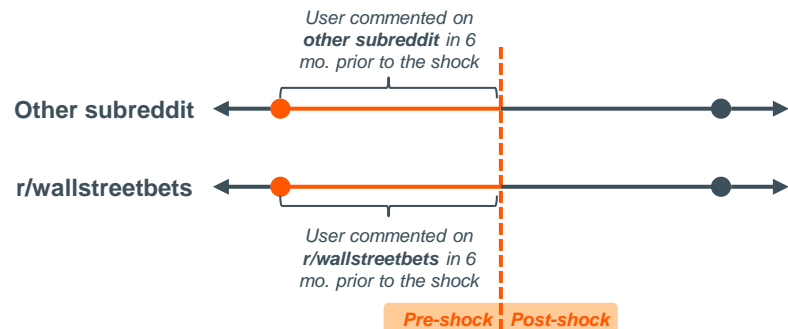
	WSB ($NonWSB = 0$)	Rest of Reddit ($NonWSB = 1$)
Pre shock ($t = 0$)	α	$\alpha + \beta$
Post shock ($t = 1$)	$\alpha + \gamma$ Change in WSB engagement post-shock	$\alpha + \beta + \gamma + \delta$ Change in non-WSB engagement post-shock

We first look at the results by pooling both groups together, looking at both pre-shock and post-shock r/wallstreetbets commenters

Group A: Pre-shock r/wallstreetbets commenters

Two conditions:

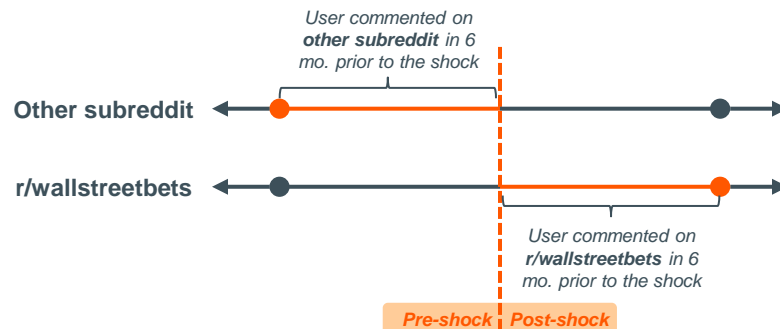
1. Pre-shock other subreddit comments: the user must have commented on any subreddit other than r/wallstreetbets in the 6 months **prior to the shock**
2. Pre-shock r/wallstreetbets comments: the user must have commented on r/wallstreetbets in the 6 months **prior to the shock**



Group B: Post-shock r/wallstreetbets commenters

Two conditions:

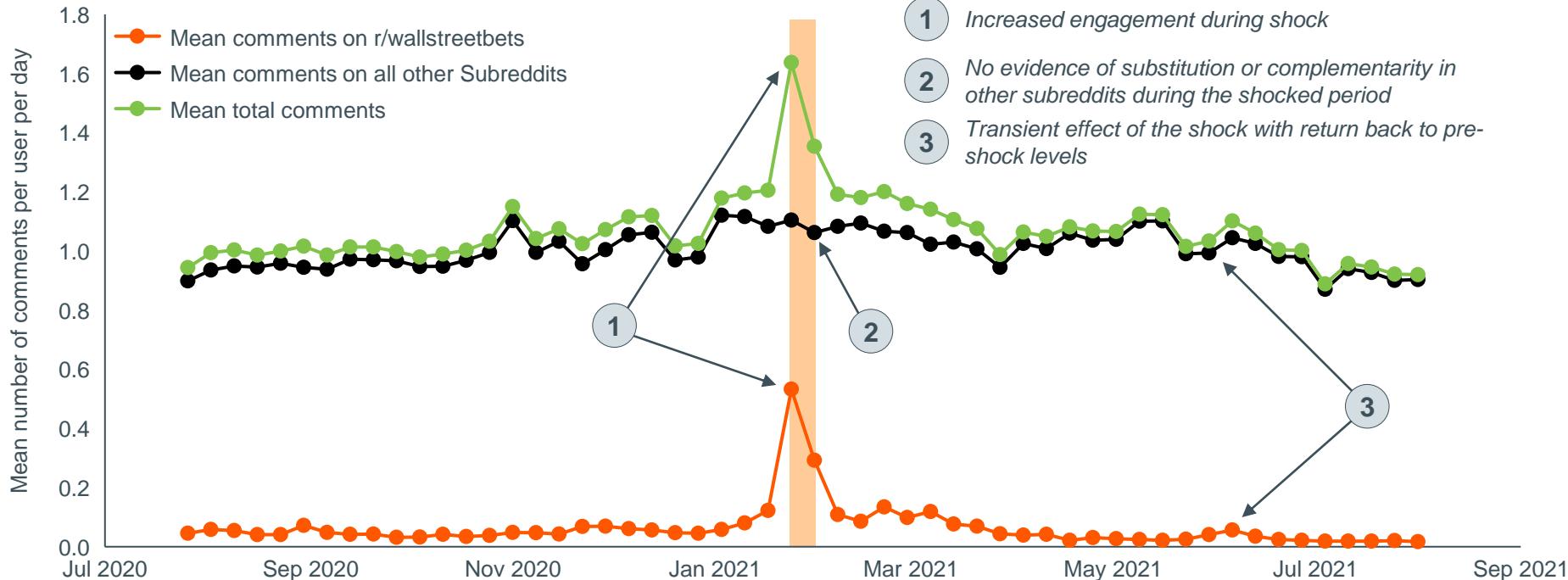
1. Pre-shock other subreddit comments: the user must have commented on any Subreddit other than r/wallstreetbets in the 6 months **prior to the shock**
2. Post-shock r/wallstreetbets comments: the user must have commented on r/wallstreetbets in the 6 months **after to the shock**



The pooled results show that the shock on r/wallstreetbets was largely transient: engagement increased rapidly but died down quickly after

Weekly mean comments among pooled sample

Increased effect of engagement during the shocked period was transient

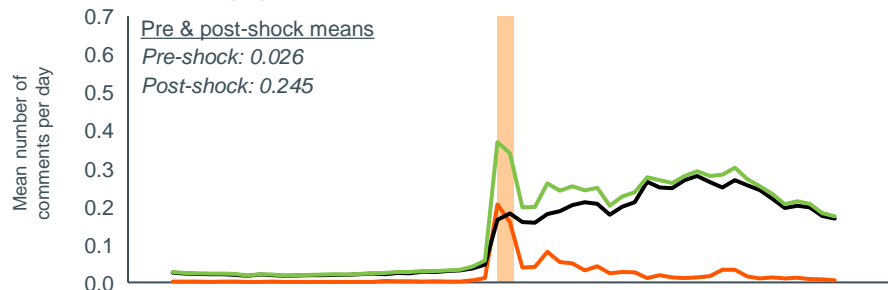


Three takeaways

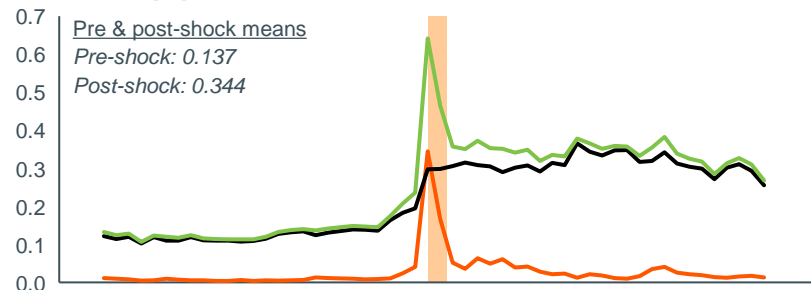
- 1 Increased engagement during shock
- 2 No evidence of substitution or complementarity in other subreddits during the shocked period
- 3 Transient effect of the shock with return back to pre-shock levels

Although an interesting pattern emerges when splitting users by their pre-shock levels of engagement...

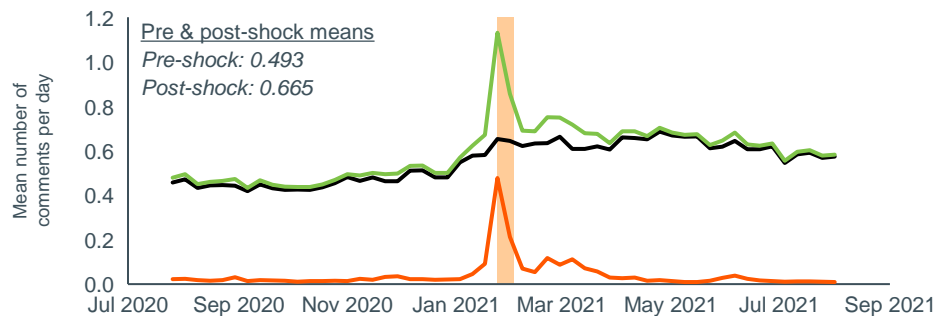
Quartile #1: Significant increase in engagement post-shock
Least-engaged pre-shock users



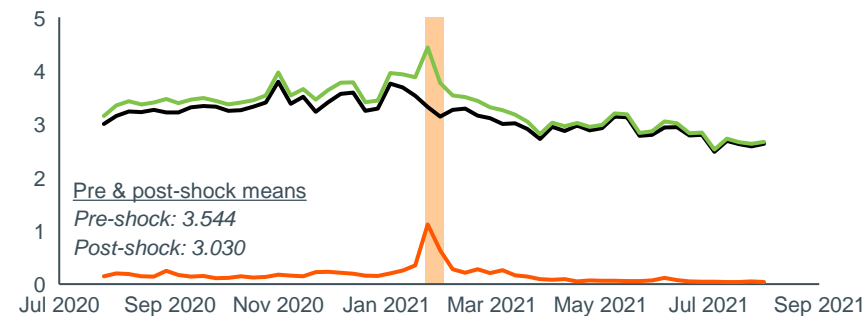
Quartile #2: Increase and complementary engagement
Less engaged pre-shock users



Quartile #3: Complemented engagement during shock
Engaged pre-shock users



Quartile #4: Substitution of engagement during shock
Most engaged pre-shock users

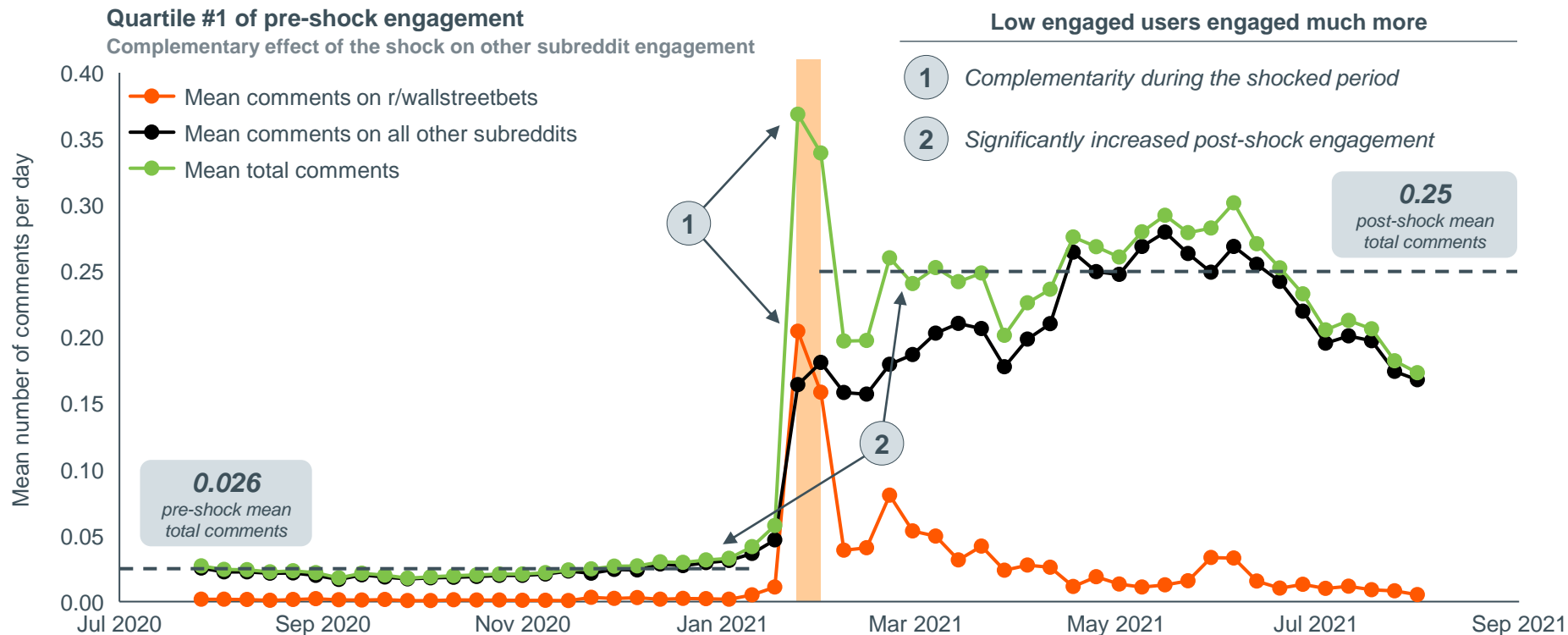


CONFIDENTIAL

— Mean comments on r/wallstreetbets
— Mean comments on all other subreddits
— Mean total comments

KEYSTONE

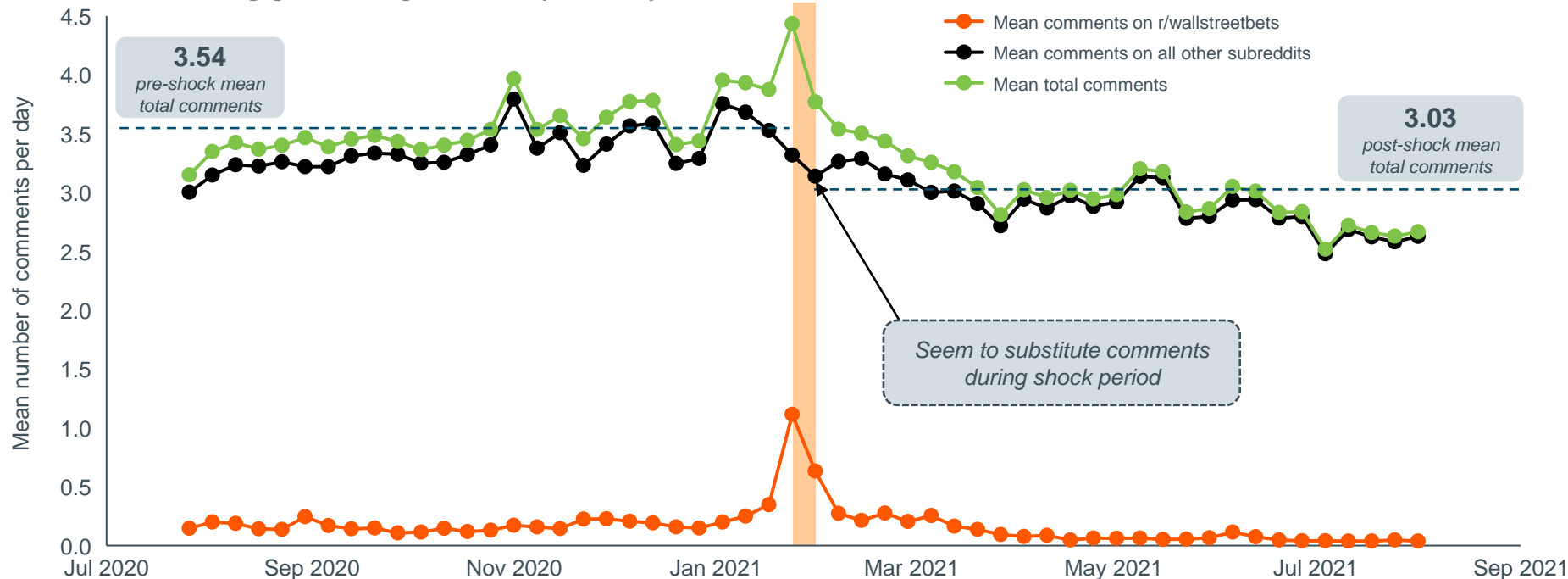
Low engagement users seemed to get a wake-up call and started engaging *nearly 10x* more with Reddit following the shock



However, the opposite was true for the engaged users who substituted during the shock and engaged less following the shock

Quartile #4 of pre-shock engagement of pre-shock commenters

Substituted engagement during the shocked period away from other subreddits



Regression results confirm our initial findings from the plots...

Regression	Full Sample	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Other subreddits flag (not r/wallstreetbets)	0.9412*** (0.0191)	0.0216*** (0.0003)	0.1186*** (0.0010)	0.4465*** (0.0034)	3.191*** (0.0667)
Shocked period (relative to pre-period)	0.4876*** (0.0160)	0.2270*** (0.0112)	0.3356*** (0.0195)	0.4465*** (0.0275)	0.9493*** (0.0530)
Post-shocked period (relative to pre-period)	0.0018 (0.0035)	0.0262*** (0.0023)	0.0221*** (0.0033)	0.0163*** (0.0026)	-0.0579*** (0.0133)
Other subreddits flag * shocked period	-0.4094*** (0.0268)	-0.0664*** (0.0154)	-0.1518*** (0.0230)	-0.2833*** (0.0306)	-1.144*** (0.0987)
Other subreddits flag * post period	0.0178 (0.0280)	0.1666*** (0.0126)	0.1625*** (0.0118)	0.1393*** (0.0183)	-0.3982*** (0.1091)
Author Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	14,960,000	3,938,220	3,603,116	3,682,404	3,736,260
Number of authors	20,000	5,265	4,817	4,923	4,995

non-r/wallstreetbets averages

Pre-shock average	0.9938	0.0238	0.1278	0.4698	3.3677
Shock average	1.0720	0.1844	0.3116	0.6330	3.1734
Post-shock average	1.0134	0.2166	0.3124	0.6255	2.9116
Change from pre-shock to shock	8%	676%	144%	35%	-6%
Change from pre-shock to post-shock	2%	811%	144%	33%	-14%

r/wallstreetbets averages

Pre-shock average	0.0526	0.0022	0.0092	0.0233	0.1763
Shock average	0.5402	0.2292	0.3448	0.4699	1.1256
Post-shock average	0.0543	0.0284	0.0313	0.0397	0.1184
Change from pre-shock to shock	928%	10,451%	3,638%	1,914%	539%
Change from pre-shock to post-shock	3%	1,208%	239%	70%	-33%

Specifically, less engaged users complemented their engagement on r/wallstreetbets with engagement on other subreddits during the shock

Regression	Full Sample	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Other subreddits flag (not r/wallstreetbets)	0.9412*** (0.0191)	0.0216*** (0.0003)	0.1186*** (0.0010)	0.4465*** (0.0034)	3.191*** (0.0667)
Shocked period (relative to pre-period)	0.4876*** (0.0160)	0.2270*** (0.0112)	0.3356*** (0.0195)	0.4465*** (0.0275)	0.9493*** (0.0530)
Post-shocked period (relative to pre-period)	0.0018 (0.0035)	0.0262*** (0.0023)	0.0221*** (0.0033)	0.0163*** (0.0026)	-0.0579*** (0.0133)
Other subreddits flag * shocked period	-0.4094*** (0.0268)	-0.0664*** (0.0154)	-0.1518*** (0.0230)	-0.2833*** (0.0306)	-1.144*** (0.0987)
Other subreddits flag * post period	0.0178 (0.0280)	0.1666*** (0.0126)	0.1625*** (0.0118)	0.1393*** (0.0183)	-0.3982*** (0.1091)
Author Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	14,960,000	3,938,220	3,603,116	3,682,404	3,736,260
Number of authors	20,000	5,265	4,817	4,923	4,995

non-r/wallstreetbets averages

Pre-shock average	0.9938	0.0238	0.1278	0.4698	3.3677
Shock average	1.0720	0.1844	0.3116	0.6330	3.1734
Post-shock average	1.0134	0.2166	0.3124	0.6255	2.9116
Change from pre-shock to shock	8%	676%	144%	35%	-6%
Change from pre-shock to post-shock	2%	811%	144%	33%	-14%

Complemented engagement on r/wallstreetbets with engagement on other subreddits

r/wallstreetbets averages

Pre-shock average	0.0526	0.0022	0.0092	0.0233	0.1763
Shock average	0.5402	0.2292	0.3448	0.4699	1.1256
Post-shock average	0.0543	0.0284	0.0313	0.0397	0.1184
Change from pre-shock to shock	928%	10,451%	3,638%	1,914%	539%
Change from pre-shock to post-shock	3%	1,208%	239%	70%	-33%

However, the most engaged users *substituted* their engagement during the shock

Regression	Full Sample	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Other subreddits flag (not r/wallstreetbets)	0.9412*** (0.0191)	0.0216*** (0.0003)	0.1186*** (0.0010)	0.4465*** (0.0034)	3.191*** (0.0667)
Shocked period (relative to pre-period)	0.4876*** (0.0160)	0.2270*** (0.0112)	0.3356*** (0.0195)	0.4465*** (0.0275)	0.9493*** (0.0530)
Post-shocked period (relative to pre-period)	0.0018 (0.0035)	0.0262*** (0.0023)	0.0221*** (0.0033)	0.0163*** (0.0026)	-0.0579*** (0.0133)
Other subreddits flag * shocked period	-0.4094*** (0.0268)	-0.0664*** (0.0154)	-0.1518*** (0.0230)	-0.2833*** (0.0306)	-1.144*** (0.0987)
Other subreddits flag * post period	0.0178 (0.0280)	0.1666*** (0.0126)	0.1625*** (0.0118)	0.1393*** (0.0183)	-0.3982*** (0.1091)
Author Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	14,960,000	3,938,220	3,603,116	3,682,404	3,736,260
Number of authors	20,000	5,265	4,817	4,923	4,995

non-r/wallstreetbets averages

Pre-shock average	0.9938	0.0238	0.1278	0.4698	3.3677
Shock average	1.0720	0.1844	0.3116	0.6330	3.1734
Post-shock average	1.0134	0.2166	0.3124	0.6255	2.9116
Change from pre-shock to shock	8%	676%	144%	35%	-6%
Change from pre-shock to post-shock	2%	811%	144%	33%	-14%

Substitution of engagement for the most engaged users pre-shock

r/wallstreetbets averages

Pre-shock average	0.0526	0.0022	0.0092	0.0233	0.1763
Shock average	0.5402	0.2292	0.3448	0.4699	1.1256
Post-shock average	0.0543	0.0284	0.0313	0.0397	0.1184
Change from pre-shock to shock	928%	10,451%	3,638%	1,914%	539%
Change from pre-shock to post-shock	3%	1,208%	239%	70%	-33%

Additionally, post-shock engagement is highly dependent on pre-shock levels of engagement...

Regression	Full Sample	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Other subreddits flag (not r/wallstreetbets)	0.9412*** (0.0191)	0.0216*** (0.0003)	0.1186*** (0.0010)	0.4465*** (0.0034)	3.191*** (0.0667)
Shocked period (relative to pre-period)	0.4876*** (0.0160)	0.2270*** (0.0112)	0.3356*** (0.0195)	0.4465*** (0.0275)	0.9493*** (0.0530)
Post-shocked period (relative to pre-period)	0.0018 (0.0035)	0.0262*** (0.0023)	0.0221*** (0.0033)	0.0163*** (0.0026)	-0.0579*** (0.0133)
Other subreddits flag * shocked period	-0.4094*** (0.0268)	-0.0664*** (0.0154)	-0.1518*** (0.0230)	-0.2833*** (0.0306)	-1.144*** (0.0987)
Other subreddits flag * post period	0.0178 (0.0280)	0.1666*** (0.0126)	0.1625*** (0.0118)	0.1393*** (0.0183)	-0.3982*** (0.1091)
Author Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	14,960,000	3,938,220	3,603,116	3,682,404	3,736,260
Number of authors	20,000	5,265	4,817	4,923	4,995

non-r/wallstreetbets averages

Pre-shock average	0.9938	0.0238	0.1278	0.4698	3.3677
Shock average	1.0720	0.1844	0.3116	0.6330	3.1734
Post-shock average	1.0134	0.2166	0.3124	0.6255	2.9116
Change from pre-shock to shock	8%	676%	144%	35%	-6%
Change from pre-shock to post-shock	2%	811%	144%	33%	-14%

r/wallstreetbets averages

Pre-shock average	0.0526	0.0022	0.0092	0.0233	0.1763
Shock average	0.5402	0.2292	0.3448	0.4699	1.1256
Post-shock average	0.0543	0.0284	0.0313	0.0397	0.1184
Change from pre-shock to shock	928%	10,451%	3,638%	1,914%	539%
Change from pre-shock to post-shock	3%	1,208%	239%	70%	-33%

Less engaged users increased their engagement significantly post-shock...

...with the most engaged users substituting their engagement

Regression	Full Sample	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Other subreddits flag (not <i>r/wallstreetbets</i>)	0.9412*** (0.0191)	0.0216*** (0.0003)	0.1186*** (0.0010)	0.4465*** (0.0034)	3.191*** (0.0667)
Shocked period (relative to pre-period)	0.4876*** (0.0160)	0.2270*** (0.0112)	0.3356*** (0.0195)	0.4465*** (0.0275)	0.9493*** (0.0530)
Post-shocked period (relative to pre-period)	0.0018 (0.0035)	0.0262*** (0.0023)	0.0221*** (0.0033)	0.0163*** (0.0026)	-0.0579*** (0.0133)
Other subreddits flag * shocked period	-0.4094*** (0.0268)	-0.0664*** (0.0154)	-0.1518*** (0.0230)	-0.2833*** (0.0306)	-1.144*** (0.0987)
Other subreddits flag * post period	0.0178 (0.0280)	0.1666*** (0.0126)	0.1625*** (0.0118)	0.1393*** (0.0183)	-0.3982*** (0.1091)
Author Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	14,960,000	3,938,220	3,603,116	3,682,404	3,736,260
Number of authors	20,000	5,265	4,817	4,923	4,995

non-r/wallstreetbets averages

Pre-shock average	0.9938	0.0238	0.1278	0.4698	3.3677
Shock average	1.0720	0.1844	0.3116	0.6330	3.1734
Post-shock average	1.0134	0.2166	0.3124	0.6255	2.9116
Change from pre-shock to shock	8%	676%	144%	35%	-6%
Change from pre-shock to post-shock	2%	811%	144%	33%	-14%

r/wallstreetbets averages

Pre-shock average	0.0526	0.0022	0.0092	0.0233	0.1763
Shock average	0.5402	0.2292	0.3448	0.4699	1.1256
Post-shock average	0.0543	0.0284	0.0313	0.0397	0.1184
Change from pre-shock to shock	928%	10,451%	3,638%	1,914%	539%
Change from pre-shock to post-shock	3%	1,208%	239%	70%	-33%

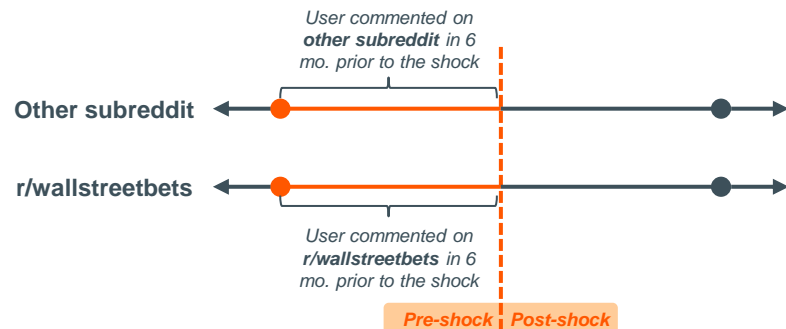
...while the most engaged users actually decreased their engagement post shock

So far we've looked at a broad group of users – what happens if we look at pre-existing r/wallstreetbets users v. new r/wallstreetbets users?

Group A: Pre-shock r/wallstreetbets commenters

Two conditions:

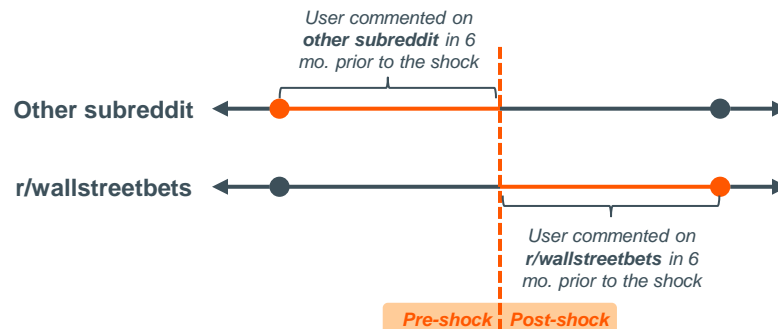
1. Pre-shock other subreddit comments: the user must have commented on any subreddit other than r/wallstreetbets in the 6 months **prior to the shock**
2. Pre-shock r/wallstreetbets comments: the user must have commented on r/wallstreetbets in the 6 months **prior to the shock**



Group B: Post-shock r/wallstreetbets commenters

Two conditions:

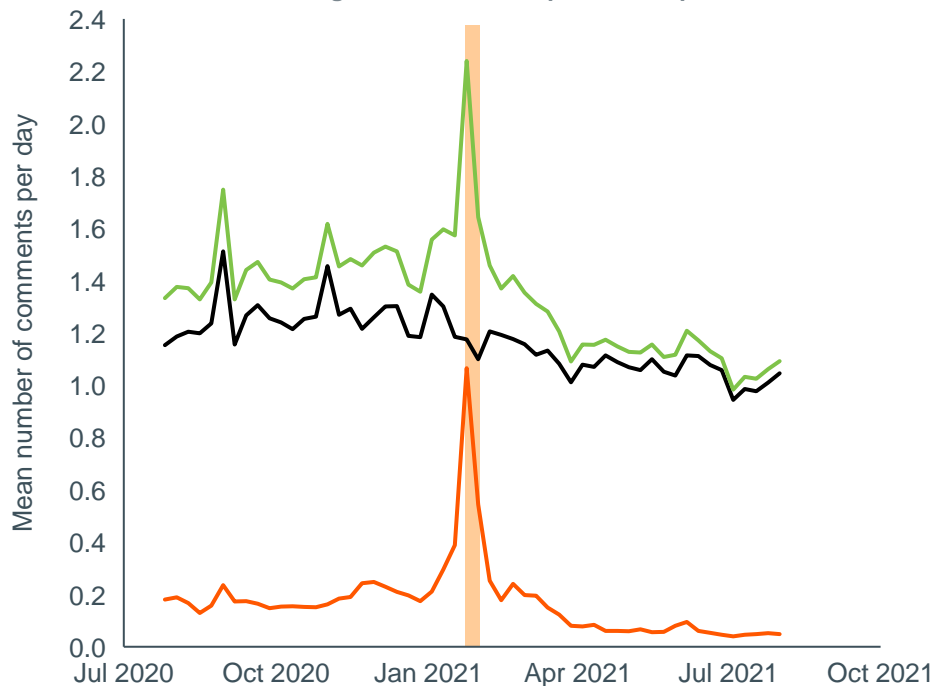
1. Pre-shock other subreddit comments: the user must have commented on any Subreddit other than r/wallstreetbets in the 6 months **prior to the shock**
2. Post-shock r/wallstreetbets comments: the user must have commented on r/wallstreetbets in the 6 months **after to the shock**



Comparison of Groups A and B (pre- and post-shock r/WSB commenters) shows only a few noticeable differences...

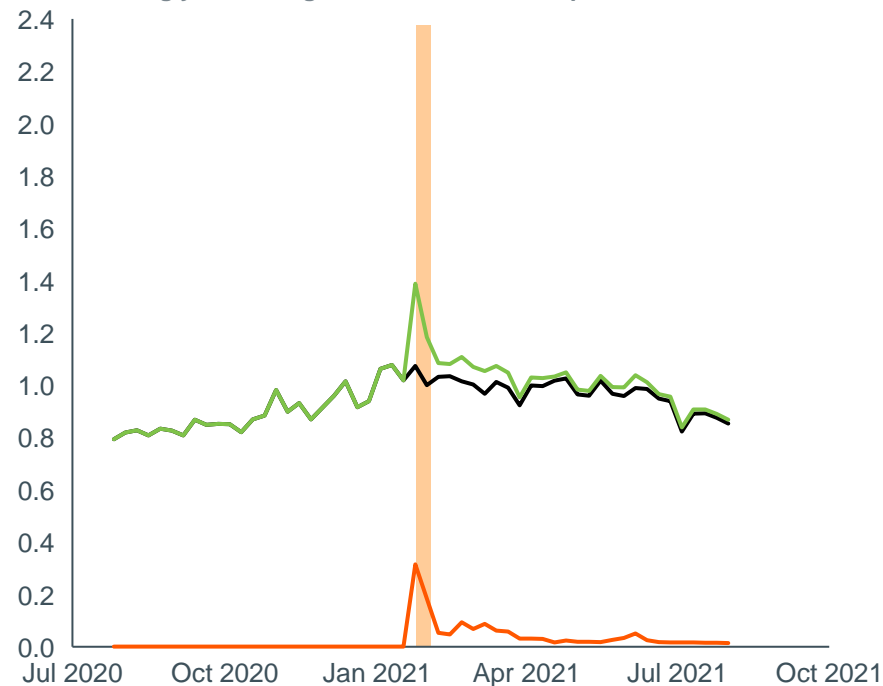
Group A: Pre-shock r/WSB commenters

Substitution during the shock with rapid decline post-shock



Group B: New r/WSB commenters post-shock

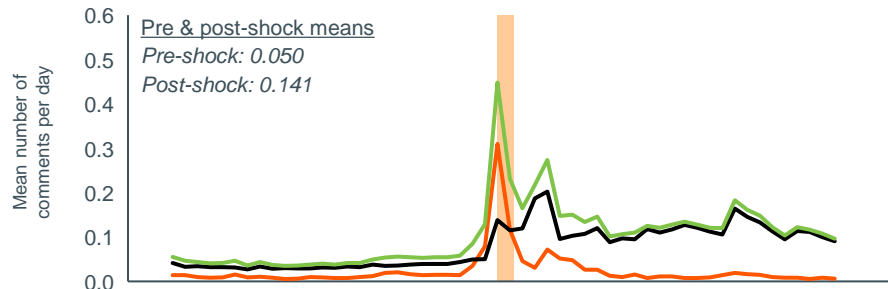
Seemingly flat during the shock with flatter post-shock



...however, when splitting by engagement metrics, it's clear there was a more pronounced effect on low engagement Group B commenters

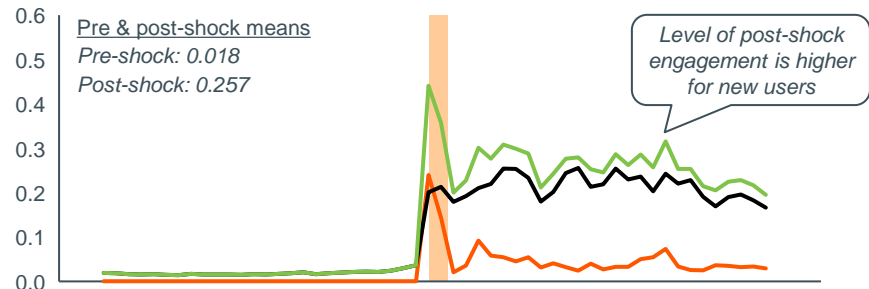
Group A, Quartile #1: Pre-shock r/WSB commenters

Relatively lower increase in post-shock engagement



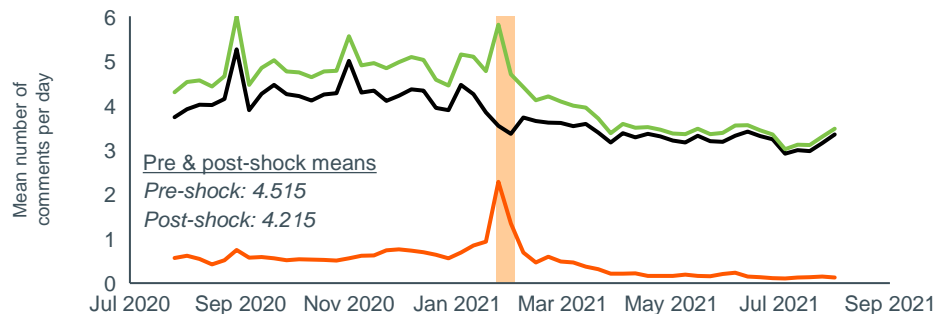
Group B, Quartile #1: Post-shock r/WSB commenters

Substantial increase in post-shock engagement



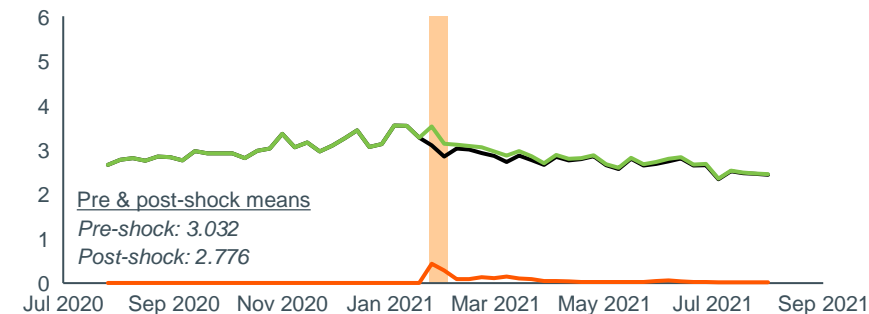
Group A, Quartile #4: Pre-shock r/WSB commenters

Pronounced decreased in total comments



Group B, Quartile #4: Post-shock r/WSB commenters

More muted decrease of most engaged users

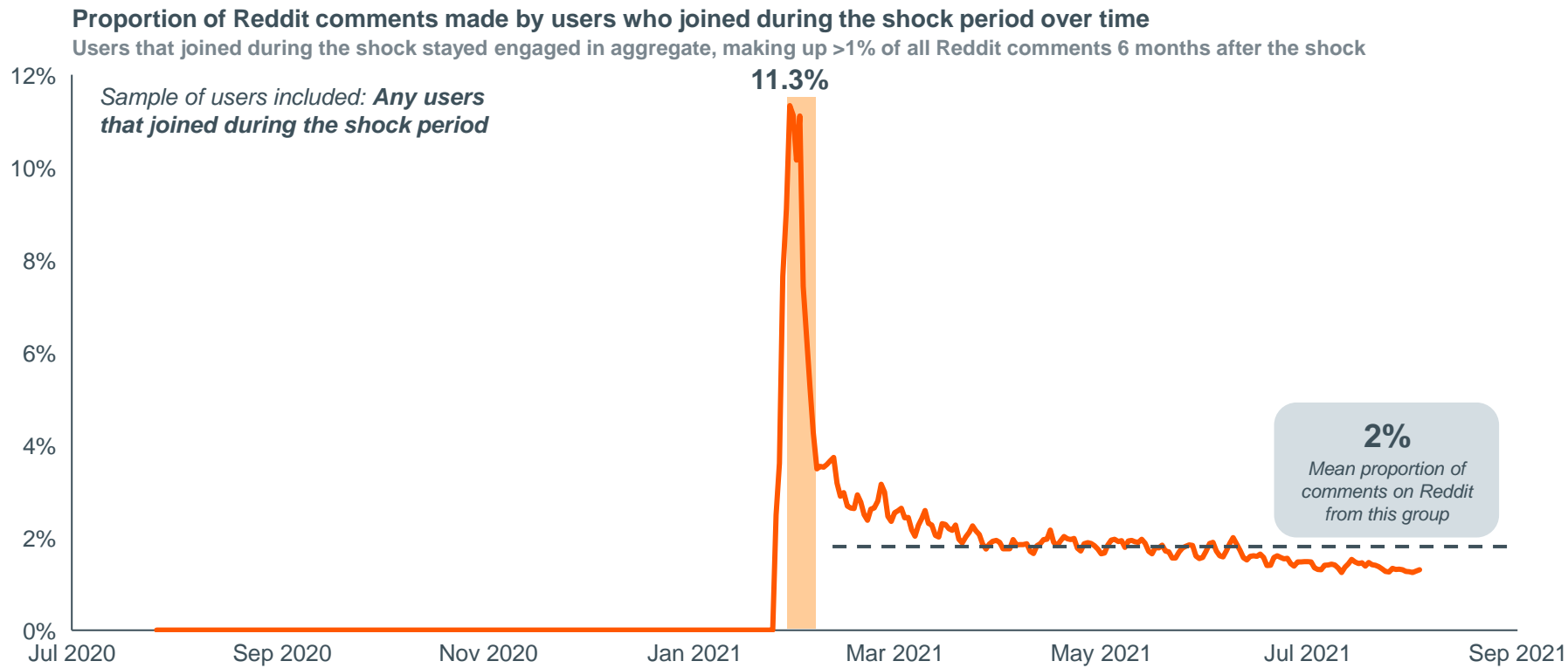


CONFIDENTIAL

— Mean comments on r/wallstreetbets
— Mean comments on all other subreddits
— Mean total comments

KEYSTONE

Additionally, when new users joined during the shock period, *they continued to engage following the shock, at significant levels*



Shout out to the Econ Research Team and Shared Services Team

THANK YOU!

**Michael Gates
Evangelos Constantinou
Nimo Suleyman
Carolina de Iruarrizaga**

**Vitoria Rabello de Castro
Jill Furzer
Catherine Moon**

**Data Shared Services
Aaron White and Bohan Li**