

News and Markets in the Time of COVID-19

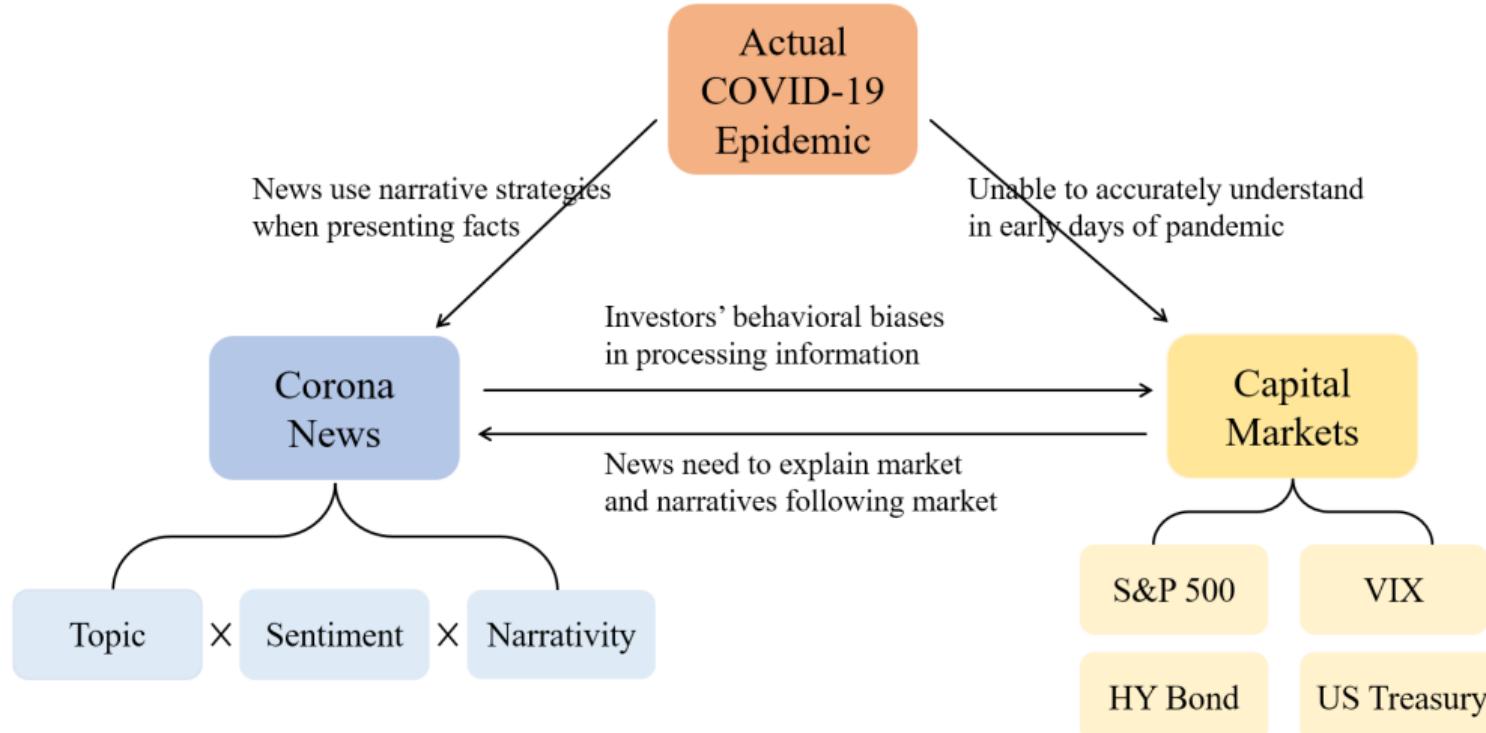
Harry Mamaysky (2024, *JFQA*)

解读人：李子明

武汉大学金融系

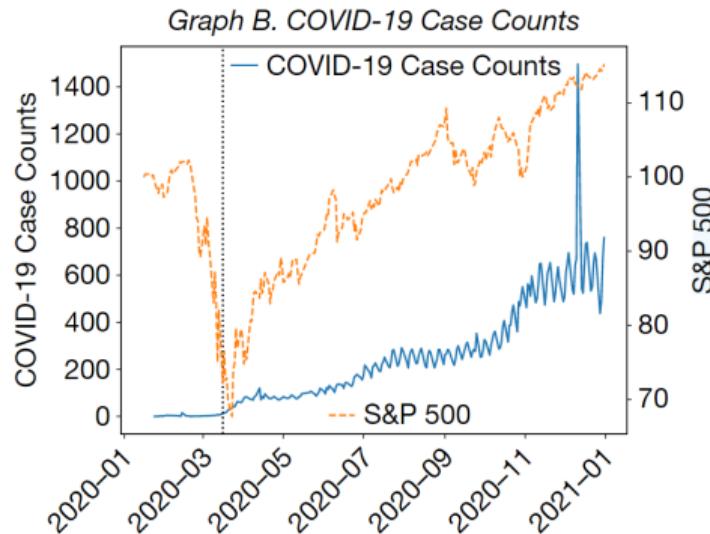
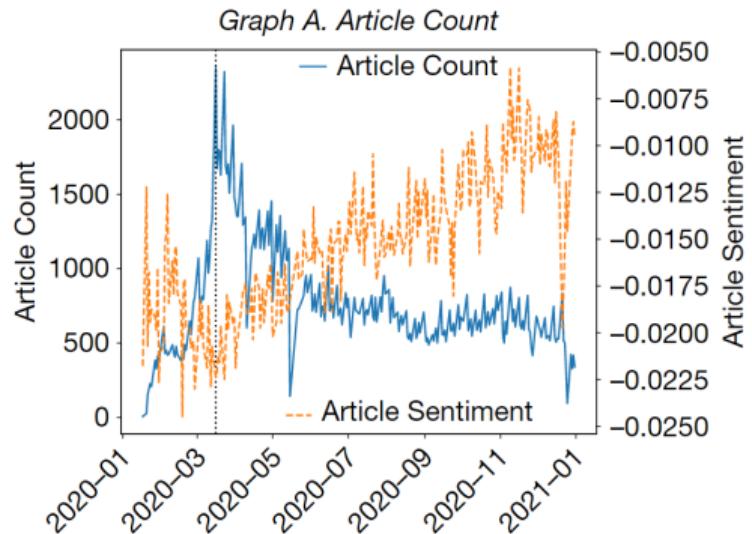
2025.11.30

Overview



Background

- Disease epidemics v.s. Narrative epidemics
 - Disease epidemics: spread of biological COVID-19 viruses.
 - Narrative epidemics: spread of beliefs that COVID-19 lead to economic disaster.



Motivation

- Which of epidemics had more influence on markets?
 - COVID-19 disease incidence rise steadily throughout all of 2020.
 - Volume and sentiment of COVID-19 news peak in March and then slowly fall.
 - S&P500 index bottom after peak of narrative epidemics and set new alltime highs.
- Understanding market behavior requires understanding evolution of **narrativity**.
 - Early days of pandemic lack of first-hand knowledge, most information came from news.
 - Narrativity capture how quickly idea spread through interpersonal communication and degree to which it influence collective beliefs.
- High narrativity news attract investor attention via **availability cascade mechanism** (Kuran and Sunstein, 1999).
 - More narrativity news ⇒ More investor positioning ⇒ High narrativity news more plausible ⇒ Further media coverage and market response

Research Question

- How economic disaster narrative of news shape financial and economic outcomes during COVID-19 pandemic?
- Whether market and news topic exhibit contemporaneous or lead-lag relationship, hypersensitivity or overreaction, depending on narrativity?

Research Hypothesis

- **H1(Contemporaneous Relationship)**: High narrativity topics systematically associated with frequent contemporaneous market responses than factual topics.
- **H2(Hypersensitivity)**: VIX impact market responses to high narrativity topics but not to low narrativity topics.
- **H3(Reversals)**: Hypersensitivity associated with contemporaneous overreaction and next day reversals.
- **H4(Causal Connectivity)**: High narrativity news series are highly connected in news-markets Granger causality network.
- **H5(Rationality)**: H1-H4 do not hold under rational information hypothesis.

Contribution

- Contribute to literature on impact of COVID-19 on firm-level outcomes.
 - Prior: study correlation between text and contemporaneous or future outcome, do not explore how text impacted by past economic variables.
 - This paper focus on feedback linkages between outcomes and text.
- Contribute to literature that study COVID-19 social media text.
 - Prior: COVID-19 news explain large portion of Feb-Mar stock return volatility (Baker et al., 2020); COVID-19 contagion risk is priced (Arteaga-Garavito et al., 2021).
 - Prior assume subjective end date of COVID-19, this paper emphasize formal break for analysis crisis and stark difference in news-markets regimes pre- and post-break.
- Contribute to literature that study narrativity of news.
 - Prior: higher narrativity forecast higher crash attention and volatility (Goetzman, 2022).
 - This paper show impact of narrativity greater during peak crisis time and subside as markets and news normalize.

Data

- News corpus
 - News corpus: 189,548 Thomson Reuters news mentioning **coronavirus** or **COVID-19** from 2019 to Dec. 2020
 - 1987 corpus: 164 Wall Street Journal news in 3 days after Oct. 19, 1987 stock collapse
 - Fed corpus: 155 FOMC statements from Feb. 4, 1994 to May 4, 2022
- Market data
 - 5 asset classes: S&P 500 index, VIX index, FTSE High-Yield Market index, US 2- and 10-year Treasury yields
- Corona case data
 - Global confirmed COVID-19 cases from Johns Hopkins Coronavirus Resource Center
 - Update frequency: Once a day around 23:59 (UTC)
 - Day t increase in global COVID-19 case count as day $t + 1$ value of corona series

Sentiment of News

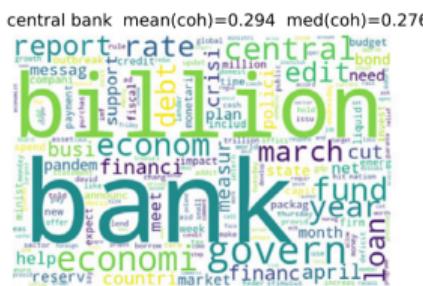
- Article j sentiment: $\text{SENT}_j = \frac{\text{POS}_j - \text{NEG}_j}{\text{TOTAL}_j}$
- POS_j and NEG_j : positive and negative words (Loughran and McDonald, 2011)
- TOTAL_j : total words excluding stop words and negated words (Das and Chen, 2007)
- Daily aggregate sentiment: $\text{SENT}_t = \frac{1}{N_t} \sum_{j \in \{\text{day } t \text{ articles}\}} \text{SENT}_j$

Topic of News

- Stem words in corpus: drop stop words and common occur words, retain all words occurring more than 10 times in any month
- LDA estimated using 72,263 corona news through end of Apr. 2020
- End date choose to ensure clean break tests (break not driven by changing topic)
- 12 topics optimally balance desire for more topics (better characterize news flow) with preference of coherent (topics contain words sensibly belong together)
- Calculate topic loadings for all 189,548 corona news, article j loading on k th topic $f_{j,k}$
- Daily aggregate topic frequency: $f_{t,k} = \frac{1}{N_t} \sum_{j \in \{\text{day } t \text{ articles}\}} f_{j,k}$

Word Clouds for Markets and Central Bank Topics

Graph A. Wordclouds for Select Topics and Sentiment Words



markets mean(coh)=0.345 med(coh)=0.342



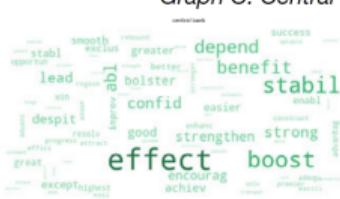
Graph B. Markets: Sentiment Words



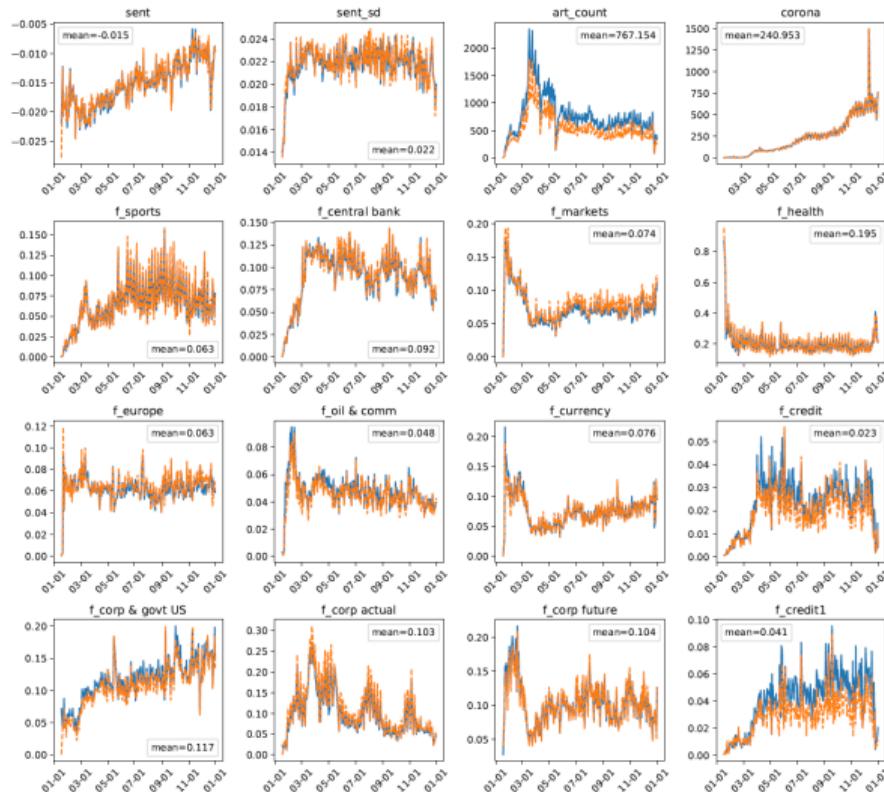
close

The word cloud is centered around the word "close". Other words visible include "fear", "loss", "contract", "eas", "volatil", "worry", "limit", "fallout", "damag", "break", "worst", "weaken", "concern", "drop", "worst", "recessi", "specul", "decid", "crisi", "distract", "slow", "shock", "shock", "draw", "liquid", and "markets". The words are rendered in different sizes and colors, with larger and bolder ones being "close", "fear", "loss", and "contract".

Graph C: Central Bank: Sentiment Words



Topic Frequencies Over Time

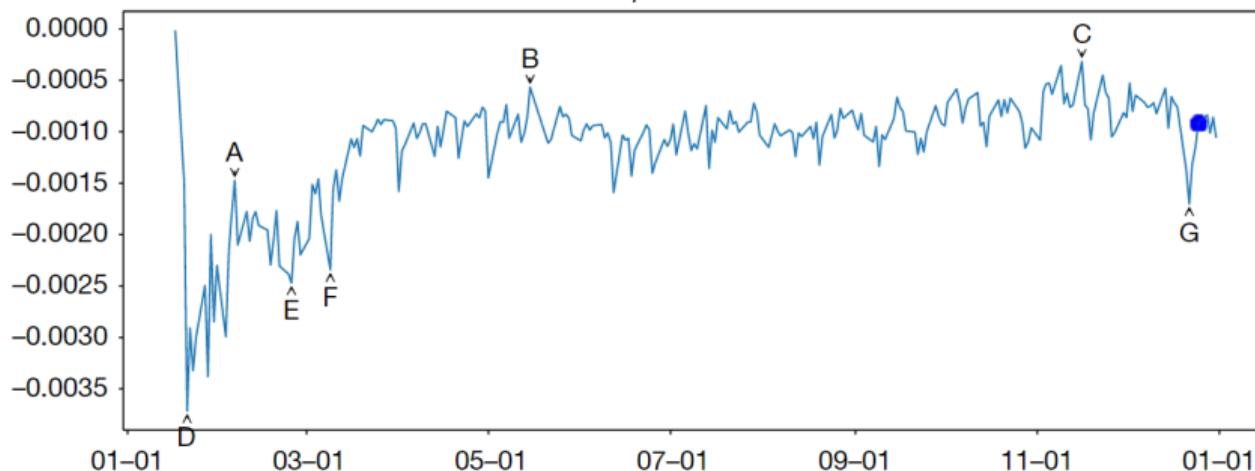


Topical Sentiment

- Topical sentiment: product of topic frequency and daily sentiment
 $\text{SENT}_{t,k} = \text{SENT}_t \times f_{t,k}$ (Calomiris and Mamaysky, 2019)
- Topical sentiment measures extent to which daily negative or positive news flow concentrates in specific topics.
- 12 topics capture rich heterogeneity of news flow about crisis: "health impact of corona" \Rightarrow "central bank intervention" "credit impacts"
- This substantial time series and cross-sectional variation will prove useful when relating to market activity across multiple asset classes

Evolution of Topical Sentiment

Graph A. Markets



A: BUZZ-Aussie energy index sees best day in over a month as oil prices jump (0.06)

B: No articles

C: MIDEAST STOCKS-Most major Gulf markets gain; Saudi index flat (0.03)

D: EMERGING MARKETS-Latam assets keel as China virus sours risk appetite (-0.04)

E: EMERGING MARKETS-Latam currencies trade lower as coronavirus concerns rise (-0.07)

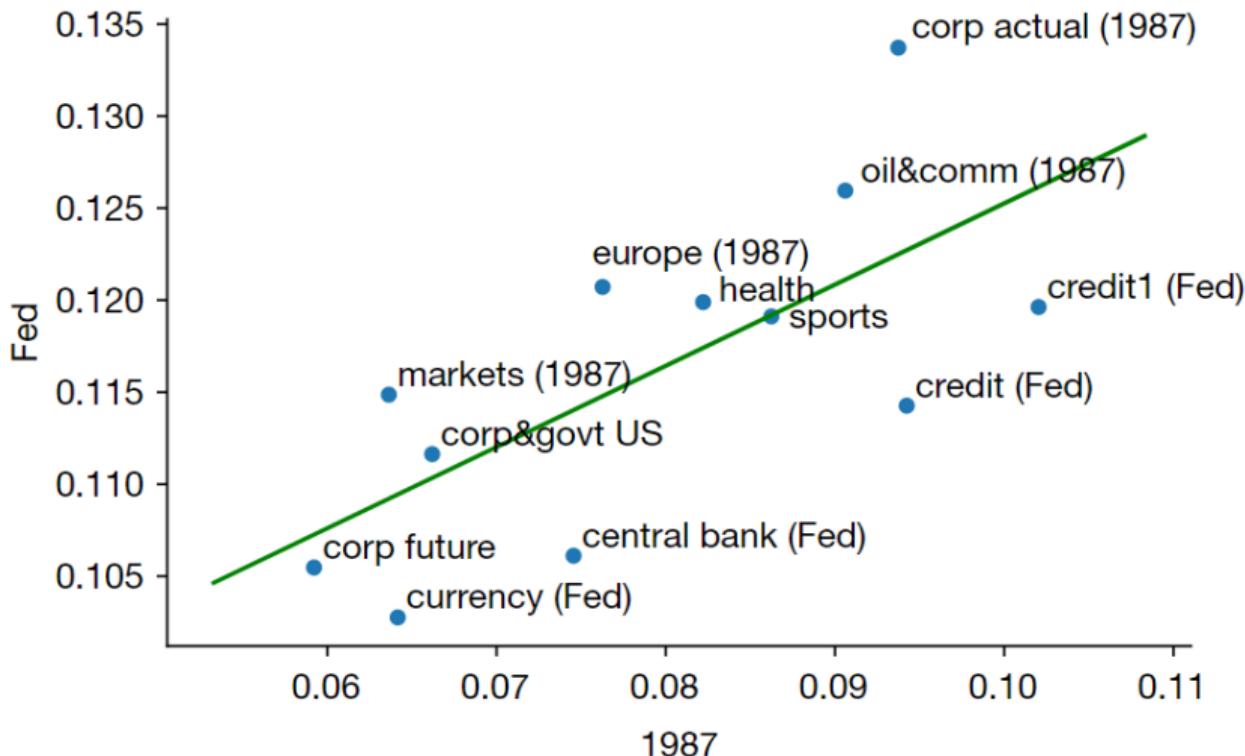
F: CANADA STOCKS-Toronto stocks eye worst day since 2008 as energy sector reels from oil crash (-0.05)

G: RUBBER-Japan futures tumble on Shanghai plunge, virus fears (-0.06)

Topic Narrativity

- Narrativity: similarity of news to corpus with high narrative content (GKS, 2022)
- High narrativity corpus: Oct. 19, 1987 stock collapse dominate media crash narratives
- Low narrativity corpus: FOMC statements are factual communications describing underlying economic conditions opposed to narratives
- Proximity of topics to corpora: $C_k \equiv \left(\sum_{w=1}^W (\tau_k(w) - \tau_C(w))^2 \right)^{1/2}, C \in \{1987, FED\}$
- W : words common to both topic and corpus
- $\tau_k(w)$: loading on word w of topic k
- $\tau_C(w)$: fraction of corpus C represented by word w
- High narrativity topic: lower textual distance to 1987 corpus
- Low narrativity topic: lower textual distance to Fed corpus

Fed Distance vs 1987 Distance

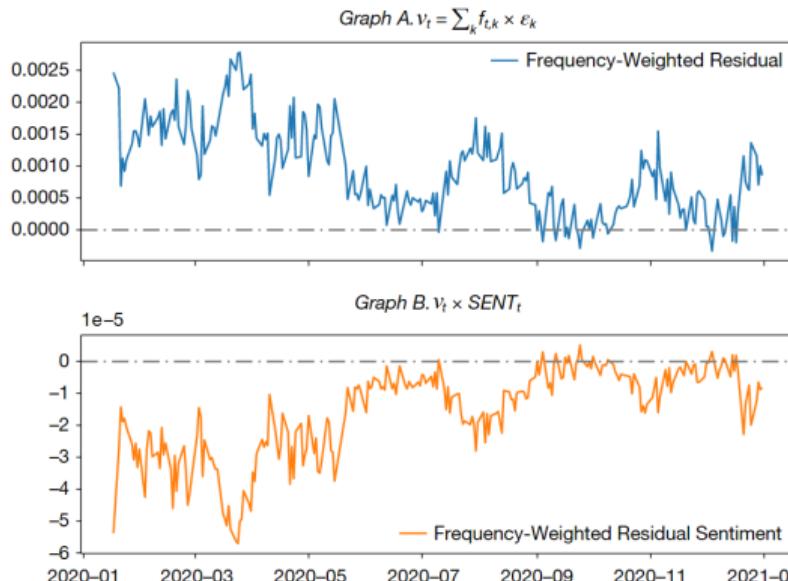


Headlines for Representative Articles by Topic

Label	1987	Fed	Headline	Sent	Date
CREDIT	0.094	0.114	Fitch Takes Action on 14 Italian Banking Groups on Coronavirus Disruption Fitch Rates Texas Instruments' \$750 Million of Five-Year Senior Notes 'A+' India extends suspension of bankruptcy filings S.Korea fin min says to boost loans to developing countries fighting coronavirus	-0.042 0.011 -0.111 0.042	03-24 03-03 09-24 04-26
CENTRAL_BANK	0.075	0.106	RPT-BUZZ-Replay-EUR/USD doubts, sterling exposed, yen setback BUZZ-EUR/USD-Sold on virus led broad risk 'off' USD strength	-0.126 0.052	06-24 12-20
CURRENCY	0.064	0.103	Fitch Ratings: USPF Housing Defines Coronavirus Scenarios for Loan Program Models Fitch to Rate BANK 2020-BNK30; Presale Issued BUZZ-Hershey: Falls on Q1 profit miss, massive sales decline in China	-0.040 0.019 -0.092	04-30 12-08 04-23
CREDIT1	0.102	0.120	Australia's Fortescue sees strong steel demand in 2021 Rugby-Champions Cup, Challenge Cup quarter-finals postponed due to coronavirus	0.060 -0.119	12-08 03-16
SPORTS	0.086	0.119	Soccer-Five positive in latest Premier League COVID-19 tests	0.091	10-12
CORP_&_GOVT_US	0.066	0.112	McDonald's accused of firing worker who sued over COVID-19 claims – Bloomberg Pelosi says bipartisan talks on COVID-19 relief making 'great progress'	-0.123 0.062	06-19 12-10
HEALTH	0.082	0.120	Zimbabwe police arrest critics ahead of anti-government protests Britain making good progress with antibody tests – junior minister	-0.144 0.069	07-20 04-27
OIL_&_COMM	0.091	0.126	UPDATE 1-LNG tanker diverted from China in sign of weaker demand CBOT wheat closes firm on strong demand	-0.063	02-04
MARKETS	0.064	0.115	Indian stocks suffer worst day in history as coronavirus shuts businesses, cities BUZZ-Australian financials extend gains to fifth day on hopes of economic rebound	-0.078 0.081	03-23 11-10
EUROPE	0.076	0.121	Refinitiv Newscasts – Confusion at Heathrow as UK cut off from Europe Refinitiv Newscasts – BioNTech confident vaccine can beat new mutation	-0.116 0.045	12-21 12-22
CORP_ACTUAL	0.094	0.134	BRIEF-Tritech Group Updates on Business Disruptions due to COVID-19 BRIEF-H2O Innovation Presents Update on COVID-19 and Ensures Continuity of its Operations for its Customers	-0.119 0.082	04-07 03-17

Indicator of Narrativity

- Topic narrativity: relative proximity to 1987 vs Fed corpora $FED_k = a + b \cdot 1987_k + \varepsilon_k$
- Daily narrativity: $v_t = \sum_{k=1}^{12} \varepsilon_k \times f_{t,k}$
- Daily narrative sentiment: $s_t = v_t \times SENT_{t,k}$



Contemporaneous Markets–News Specification

- Assume investor expectations of future log return are linear in observable state variables (Campbell and Shiller, 1988)
- Contemporaneous relationship between news and returns: $h_{t+1} = c + b^\top w_{t+1} + e_{t+1}$
- h_{t+1} : asset return from day t to $t + 1$
- w_{t+1} : vector of news flow and other information available at $t + 1$
- e_{t+1} : unobservable information at $t + 1$

Specialized Markets–News Regression

- $h_{t+1} = c + b_1 h_t + b_2 h_{t-1} + b_3 N_{t+1} + b_4 N_{t+1} (\text{VIX}_t^{10} - \overline{\text{VIX}}^{10}) + b_5 \text{VIX}_t^{10} + e_{t+1}$
- Two lagged returns h_t and h_{t-1} : control for dependence of current news on past returns and auto-correlation of returns
- N_{t+1} : news variable measured at day $t + 1$
 - Aggregate daily sentiment SENT_{t+1}
 - Daily standard deviation of article-level sentiment SENT_SD_{t+1}
 - Daily article count ART_COUNT_{t+1}
 - One of the 12 topical sentiment series $\text{SENT}_{t+1,k}$
 - COVID-19 case count series CORONA_{t+1}
- Crisis shock make economy high-informational and high-volatility (GMS, 2024)
 - VIX_t^{10} : 10-day moving average of VIX index at day t
 - $N_{t+1} (\text{VIX}_t^{10} - \overline{\text{VIX}}^{10})$: capture effect of N_{t+1} depend on volatility currently high or low

Regressions of Daily Market Returns on News flow (Pre-Break)

	SP500			VIX			HY			GT2			GT10		
	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1
	SENT	0.761***	0.085***	-0.701***	-0.081***	0.603***	0.104***	-0.347**	0.616***	0.093**	0.906***	0.574***	0.092*	-0.218*	
SENT_SD	-0.742**	-0.106*	1.137***	0.857***	0.117***	-1.098***	-1.100**	-0.180***	1.181***				0.574***	0.092*	1.231***
FACTOR	0.773***	0.078***		-0.705***	-0.066***		0.607***	0.101***	-0.493***	0.826***	0.081***		0.725***	0.100***	-0.368**
FACTOR_ALL	0.522*			-0.471*				-0.360*	0.943***	0.087***		0.703***			
ART_COUNT	-1.240**			1.485***	0.034*	-0.578*	-0.792**	-0.045**	1.506***	-0.957*					
CORONA		-0.046**	0.868***		0.046**	-0.823***		-0.041**	0.222*			0.401**	0.426**	0.067**	0.505***
SPORTS	0.940**		-0.613*	-1.041***			0.645**	0.035*	-1.333***	1.129**			0.891*		
CENTRAL_BANK									0.615*						
MARKETS	0.857***	0.075***		-0.798***	-0.069***		1.079***	0.110***	-0.442*	0.426*	0.078***	-0.892***	0.960***	0.139***	-1.077**
HEALTH	1.053***	0.104**		-1.099***	-0.113***		1.245***	0.171***	-0.892***				0.940*		
EUROPE	0.985***	0.078*		-0.829***			0.676***	0.101***							-0.617**
OIL_&_COMM	0.663***	0.215***		-0.460***	-0.187***		1.073***	0.223***			0.817***	0.169***	-0.808***	1.110***	0.232***
CURRENCY	0.771**			-0.681**						1.262***	0.148***		1.127***	0.139***	-0.863***
CREDIT										0.576*	0.061*				
CORP_&GOVT_US							0.338**		0.379*						-0.077**
CORP_ACTUAL		0.297**			-0.403**										
CORP_FUTURE										0.531**	0.089***		0.457***	0.096***	
CREDIT1															

Mean $|b_3| = 0.815$, SIG = 53, JNT_SIG = 38, HYPER = 38, LEAD = 29, UNDER = 2, OVER = 17, OVER + HYPER = 16.

SIG: b_3 significant; **JNT_SIG** (joint significance): b_3 and b_4 both significant; **HYPER** (hypersensitivity): b_3 and b_4 both significant and same sign; **LEAD** (lead-lag relationship): c_3 significant; **UNDER** (underreaction): b_3 and c_3 both significant and same sign; **OVER** (overreaction or reversal): b_3 and c_3 both significant and opposite signs; **OVER + HYPER** (overreaction and hypersensitivity): b_3 and c_3 opposite sign while b_3 and b_4 same sign

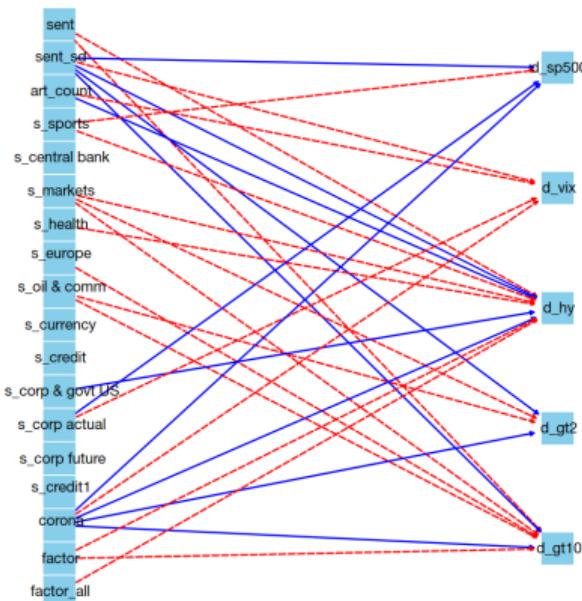
Lead-Lag Relationships

- Granger causality test framework
 - Model 1 (without X): $Y_t = a_0 + \sum_{i=1}^p a_i Y_{t-i} + \varepsilon_t$
 - Model 2 (with X): $Y_t = b_0 + \sum_{i=1}^p b_i Y_{t-i} + \sum_{i=1}^p c_i X_{t-i} + u_t$
 - c_i jointly significant in F-test \Rightarrow X Granger-cause Y
 - X precedes Y in time series and has predictive value (not strict causal relationship)
- $\rho_{t+1} = c_0 + c_1 \rho_t + c_2 \rho_{t-1} + c_3 \tau_t + c_4 \tau_t (\text{VIX}_t^{10} - \overline{\text{VIX}}^{10}) + c_5 \text{VIX}_t^{10} + e_{t+1}$
- τ Granger-cause ρ if c_3 significant at the 10% level or better
- News to Markets: markets series as response variables and news series as test variable
- Markets to News: news series as response variables and markets series as test variable

Granger Causality Network: News to Markets

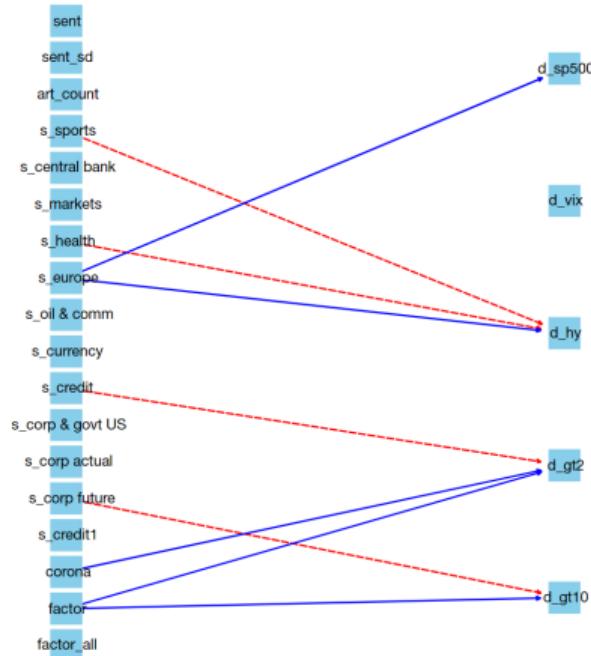
Graph A. Early Subperiod

Granger causality network 2020-01-17 to 2020-03-15
Mean $|c_3|=0.719$ Num signif=29 Num joint signif=23 Num same sign=23



Graph B. Late Subperiod

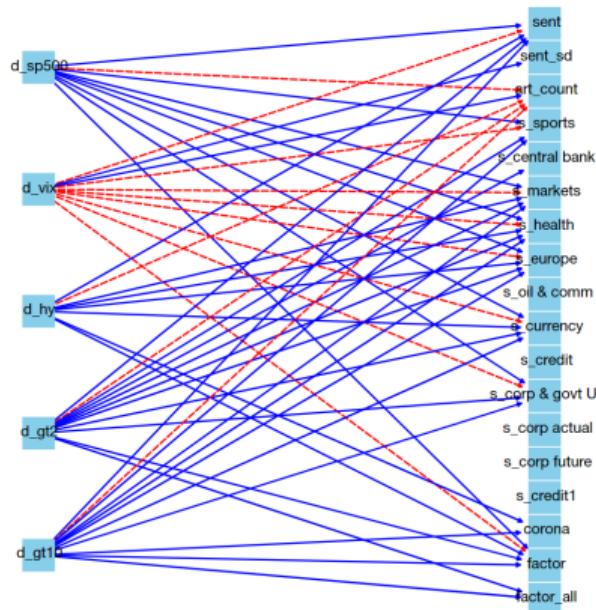
Granger causality network 2020-03-15 to 2020-12-31
Mean $|c_3|=0.126$ Num signif=9 Num joint signif=4 Num same sign=4



Granger Causality Network: Markets to News

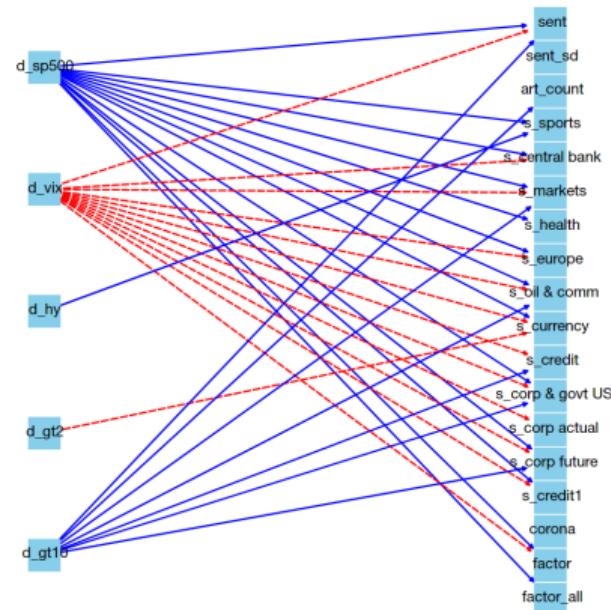
Graph A. Early Subperiod

Granger causality network 2020-01-17 to 2020-03-15
Mean $|c_3|=0.403$ Num signif=49 Num joint signif=43 Num same sign=0



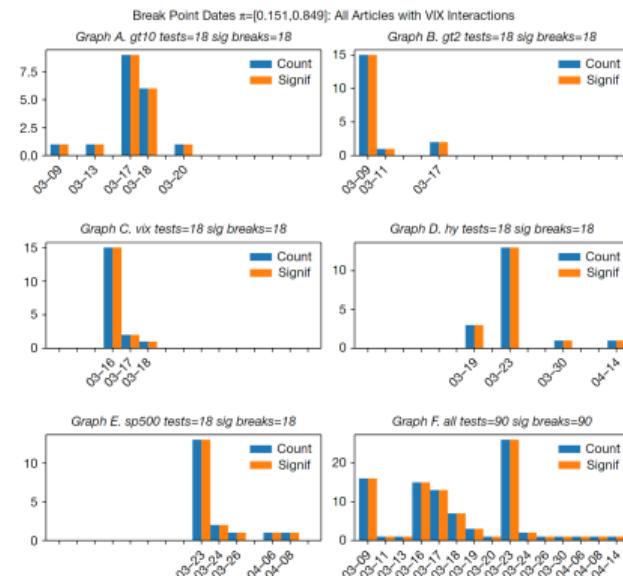
Graph B. Late Subperiod

Granger causality network 2020-03-15 to 2020-12-31
Mean $|c_3|=0.240$ Num signif=34 Num joint signif=23 Num same sign=0



Structural Break

- Assume a structural break exist at known point $t = \pi T$, $\pi \in [0, 1]$
- Structural break detected by Chow test statistic: $\phi(t) = \frac{(SSR - SSR_e - SSR_l)/k}{(SSR_e + SSR_l)/(N_e + N_l - 2k)}$
- All 5 asset classes a regime break in March \Rightarrow Sunday, Mar. 15, 2020 as cutoff date



Regressions of Daily Market Returns on News flow (Post-Break)

	SP500			VIX			HY			GT2			GT10		
	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1
SENT	0.375***	0.027**					0.472***	0.035***							
SENT_SD				0.169**											
FACTOR	0.154***	0.018***		-0.221**			0.209***	0.023***		0.147**			0.073*	0.119*	
FACTOR_ALL	0.166**			-0.242**			0.241***								0.076*
ART_COUNT							-0.354*						-0.015***		
CORONA				-0.230*									0.028***	0.149*	
SPORTS							0.288***	0.045***	-0.204**						
CENTRAL_BANK	0.337***	0.023**					0.444***	0.028**					0.013**		
MARKETS	0.281***	0.031***		-0.140**	-0.012*		0.249***	0.033***							
HEALTH	0.232**	0.022***					0.288***	0.023***	-0.108*						
EUROPE	0.434***	0.033***	0.194**	-0.330**			0.456***	0.037***	0.176*						
OIL_&_COMM							0.203**								
CURRENCY	0.230***			-0.203**	-0.023*		0.244***	0.030***					-0.083*		
CREDIT															
CORP_&_GOVT_US							0.016*	0.201***	0.026***				-0.026**		
CORP_ACTUAL								0.162*					0.014**		
CORP_FUTURE	0.139**						0.124**	0.018**							-0.073*
CREDIT1															

Hypothesis Tests: Difference Caused by Narrativity

- Regression on asset cross-sections with news topics:
$$O_{m,k} = a_m + b_{\text{Fed}} \times \text{FED}_k + b_{1987} \times 1987_k + \varepsilon_{m,k}$$
- $O_{m,k}$: outcome variable for market m and news topic k relationship
- Higher narrativity news \Rightarrow More powerful and varied impact on market \Rightarrow Stronger feedback and causal relationship with market

	CONT-E	HYP-E	REV-E	GRNG-E	M2T-E	CONT-L	HYP-L	REV-L	GRNG-L	M2T-L
1987	0.815*** (0.000)	0.834*** (0.000)	0.115 (0.270)	1.493*** (0.000)	1.063*** (0.000)	0.747** (0.012)	0.589** (0.020)	-0.079 (0.394)	0.268 (0.367)	0.198 (0.195)
FED	-0.589 (0.115)	-0.751** (0.044)	-0.198 (0.285)	-1.133*** (0.000)	-0.476*** (0.000)	-0.294* (0.052)	-0.136 (0.381)	-0.031 (0.396)	0.041 (0.892)	0.152* (0.098)
HYP			0.260** (0.038)					0.129 (0.287)		
No. of obs.	60	60	60	60	60	60	60	60	60	60
R ²	0.135	0.174	0.192	0.278	0.245	0.264	0.188	0.064	0.027	0.062

Future Research Directions

- Are feedback loops a hallmark of market crises in general?
 - Past crises: Dot-com crash of 2000–2001, Global Financial Crisis of 2008–2009, European sovereign debt crisis of 2010–2012
- Narrative strategies of Company disclosure
 - Can firms strategically use narratives to stabilize investor beliefs and firm value?
 - How do narratives interact with fundamental information in disclosure?
 - How do narratives evolve through disclosure over time?