News and Networks: Using Text Analytics to Assess Bank Networks During COVID-19 Crisis

Sophia Kazinnik, Daniela Scida, John Wu (FRB Richmond)
Cooper Killen (FRB Chicago)

ASSA Annual Meeting
January 7 – 9, 2022

FEDERAL RESERVE BANK OF RICHMOND
Richmond ▪ Baltimore ▪ Charlotte

Motivation

- Studying financial networks is key to understanding:
  - Financial interconnectedness
  - Systemic importance
- Traditionally, bank interdependencies are captured via:
  - Interbank lending data (e.g., Gofman (2011); Afonso, Kovner, and Schoar (2014))
  - Co-movements in market data (e.g., Billio, Getmanzky, Lo, and Pelizzon (2012); Diebold and Yilmaz (2014); Hardle, Wang, Yu (2016))
- Alternatively, one can use text to construct networks: Banks' relationships in the view of public discussion (here, financial news)

This paper

- We study the interconnectedness of large U.S. financial institutions that fall under the Dodd-Frank Act Stress Test (DFAST) umbrella during the events surrounding the stress period related to the COVID-19 pandemic

- Build upon Rönnqvist and Sarlin (2015, Quantitative Finance) "*text-to-network approach*" and construct weekly network matrices based on co-mentioning of banks in news

- Financial connections should be broadly understood as resulting from any financial link (positive or negative) from news that translate into two banks being co-mentioned

Contribution

- We are the first to study the network among stress tested banks
- We study the network dynamics during time of stress and shed light on the impact of COVID-19 events on the network topology
- We construct an index based on our co-occurrence measure to track stress in the financial sector
- We propose using the eigenvector centrality of nodes to rank systemic importance of these financial institutions, and compare it to rankings based on traditional systemic risk measures

Results Preview

- Intuitive patterns of DFAST banks networks based on media narrative
  - Similar types of banks are clustered together (e.g., big 6, trusts, credit cards, IHCs)
  - Core-periphery topology (i.e., largest banks clustered together at the center and IHCs at the periphery)

- During periods of stress, we observe:
  - Denser networks, consistent with the literature
  - More connections across different bank groups (i.e., cross-cluster connectivity increases)
  - Connections across big players are quite stable, while connections at the periphery increase

- More intuitive and stable systemic risk rankings using text-based eigenvector centrality vs traditional systemic risk measures (e.g., SRISK)

Results Preview (cont'd)

- Overall, our methodology allows for:
    - Analysis of both cross-sectional and time dimension elements of systemic risk
    - Frequent and granular updating of both the network topology, the proposed stress index, and the systemic risk rankings
    - Real-time analysis of a financial system's architecture
    - Use of text narrative to better understand the observed connections and changes in patterns

Data: News Articles

- We derive our financial interconnectedness measure from financial news articles:
  - Dow Jones Factiva Analytics database
  - All articles on DFAST banks from top financial news sources from 07/01/2019 - 09/30/2020 DFAST Banks Sources
  - Around 70K articles in total (18K articles with co-mentions)
- We divide our sample into three parts:
  - Pre-pandemic period (July 2019 through February 2020),
  - High stress period (March through April 2020), and
  - Period of a "new normal" (May through September 2020)

Methodology: Network Analysis

- We construct weekly co-occurrence network matrices for our sample period:
  - Non-zero co-occurrences represent the links between every bank-pair
  - Co-occurrence values measure the importance of each connection (i.e., network weights)

Text2Network

- We use *eigenvector centrality* to determine centrally positioned nodes
  - It weighs both the importance of own (i.e., direct) and neighbors (i.e., indirect) connections $\rightarrow$ quality besides quantity of connections matters

Co-occurrence across time

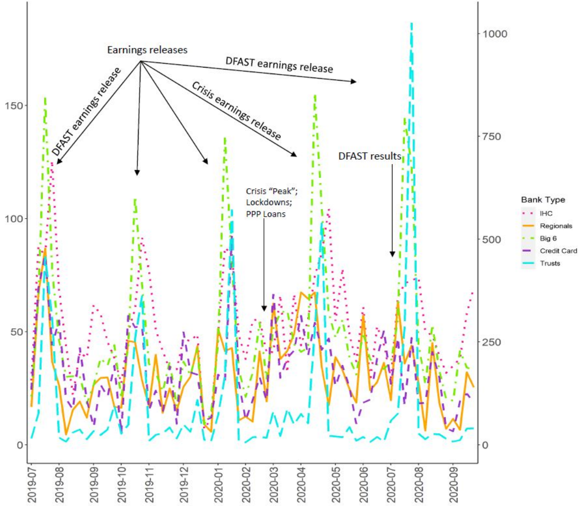
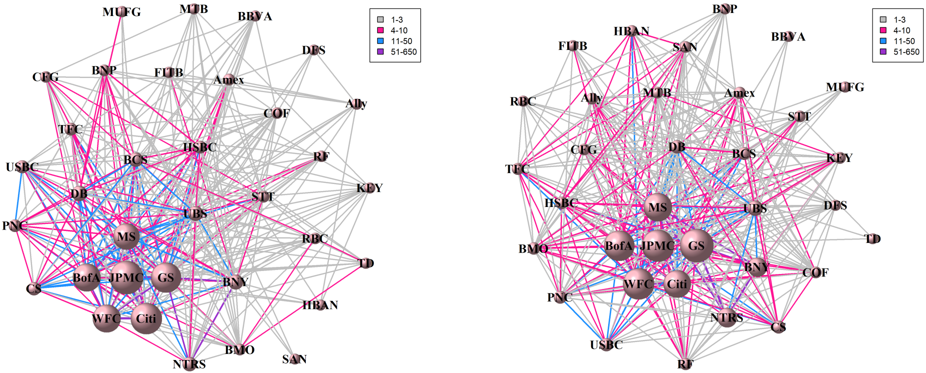


Figure 1: Time series of bank co-occurrences, by bank type
(Big 6 on the right axis)

Network graphs comparison



Panel A. January 2020 Earnings

Panel B. April 2020 Earnings

Figure 2: Network graphs: Pre-crisis vs crisis periods

Network topology comparison

Table 1: Summary statistics of January and April earnings network matrices

	January Earnings	April Earnings	% Change
Number of Connections			
Within *Big 6*	12	12	0%
Within Non-*Big 6*	598	698	16.72%
Between *Big 6* and Non-*Big 6*	131	141	7.63 %
Number of Co-occurrences			
Within *Big 6*	3432	3788	10.37%
Within Non-*Big 6*	1556	1959	25.90%
Between *Big 6* and Non-*Big 6*	1069	1218	26.29 %
Other metrics			
Clustering Coefficient	0.69	0.76	
Average Path Length	1.50	1.41	

Note: January Earnings is 13 - 19, 2020; April Earnings is 13 - 19, 2020. Connections is the number of links and co-occurrences is the number of co-mentions in articles. Clustering coefficient is calculated as the transitivity or connectivity of a network and average path length is the mean shortest path between two nodes.

Measuring Network Connectivity: Stress Index

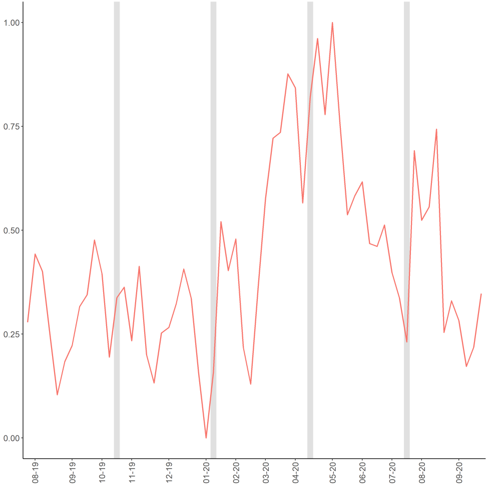


Figure 3: Event-adjusted stress index

Systemic Risk Rankings

Panel A. Eigenvector Centrality Panel B. SRISK

Figure 4: Ranking of *Big 6* Banks: SRISK versus eigenvector centrality

Distance

Robustness Checks

- Monthly vs Weekly Eigenvector Centrality
- Co-occurrence Using Select Publications: Reuters
- Including IHCs in systemic risk analysis

Conclusions

- We investigate the interconnectedness of DFAST bank holding companies by analyzing how they are mentioned together in financial news articles in the context of the COVID-19 induced financial crisis
- Text-based networks provide a real time alternative to traditional network approaches with more traceable connections: Observed patterns seem intuitive in normal times and bank connections become denser cross-type and at the periphery during the peaks of COVID-19 induced financial stress
- Our text-based stress index can be used to help track real-time stress in the financial system
- Our text-based systemic risk measure offers an alternative to traditional measures like SRISK and provides more stable systemic risk rankings in both normal times and during stress periods

Next Steps

- Refine co-occurrence measure by further exploiting the text (e.g., add sentiment)
- Manual classification of articles of our two key weeks (January and April 2020):
  - Assess accuracy of co-occurrence
  - Further investigate narrative behind connections
  - In particular, better understand drivers of new connections (or differences) during stress
- Expand systemic risk comparison to other systemic risk measures (e.g., CoVAR)

Thank You!

Appendix

Table 2: List of DFAST Bank Holding Companies (BHC)

Bank Type	Bank Name	Symbol
Big 6	Bank of America	BofA
	Citigroup	Citi
	Goldman Sachs	GS
	JPMorgan Chase	JPMC
	Morgan Stanley	MS
	Wells Fargo	WFC
Trusts	BNY Mellon	BNY
	Northern Trust	NTRS
	State Street Corp	STT
Credit Card	American Express	Amex
	Capital One	COF
	Discover Financial	DFS

Bank Type	Bank Name	Symbol
Regionals	Ally Financial	Ally
	Fifth Third Bank	FITB
	Huntington Bank	HBAN
	KeyCorp	KEY
	M&T Bank	MTB
	PNC Group	PNC
	Regions Financial	RF
	Truist	TFC
	US Bancorp	USBC
IHC	BBVA Compass	BBVA
	Bank of Montreal	BMO
	BNP Paribas	BNP
	Barclays Bank	BCS
	Credit Suisse	CS
	Deutsche Bank	DB
	HSBC Bank	HSBC
	MUFG Union	MUFG
	Santander Bank	SAN
	TD Bank	TD
	UBS Group	UBS

Back

Table 3: List of news source groups from Factiva Analytics

Code	Name	Notable Examples
TDJW	Dow Jones Newswire	Dow Jones Institutions News
TMNB	Major News and Business Sources	CNN, NY Times, Charlotte Observer
TPRW	Press Release Wires	Business Wires, Nasdaq/Globenewswire
TRTW	Reuters Newswires	Reuters News
SFWSJ	Wall Street Journal Sources	The Wall Street Journal
IBNK	Banking/Credit Sources	American Banker, Financial Times
IFINAL	Financial Services Sources	The Economist, MarketWatch

Back

Methodology: From text to network

- Look at the co-occurrences of entity names in a given news article
- Example: Assume we have the following documents (i.e., news article) in our corpus:
  - Doc 1: Acme Corp banks with both WFC and BoA.
  - Doc 2: The headquarter of WFC is in SF, and BAC's is in Charlotte.
  - Doc 3: In Q3, WFC was fined $1.5B for its dealings with JPMC. WFC plans to appeal.

	WFC	BoA	BAC	JPMC
Doc 1	1	1	0	0
Doc 2	1	0	1	0
Doc 3	2	0	0	1

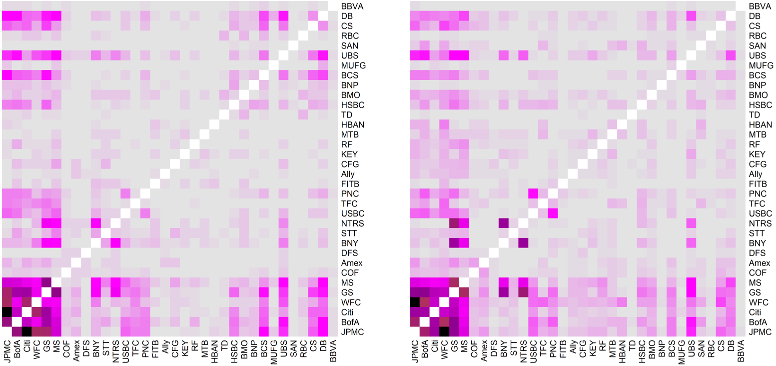
Table 4: Raw term-document matrix: M

	WFC	BAC	JPMC
WFC	3	2	1
BAC	2	2	0
JPMC	1	0	1

Table 5: Co-occurrence matrix: $C = M^T \times M$

Back

Heatmaps



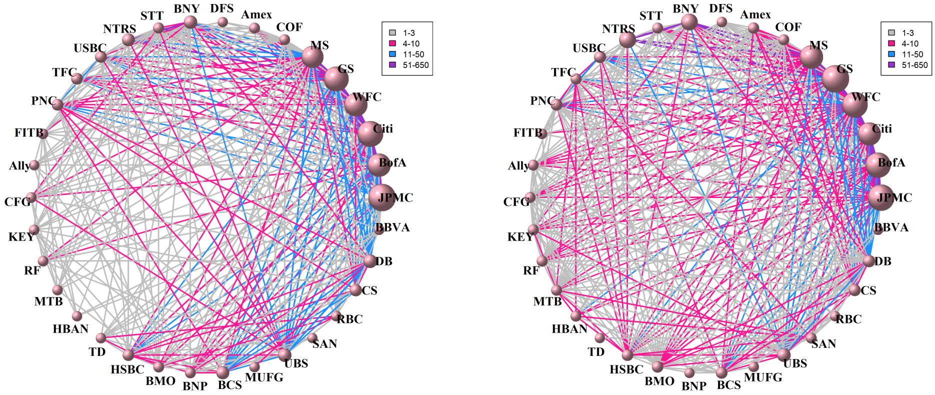
Panel A. January 2020 Earnings

Panel B. April 2020 Earnings

Figure 5: Heatmaps: Pre-crisis vs crisis periods

Back

Circle Graphs



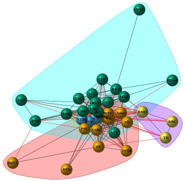
Panel A. January 2020 Earnings

Panel B. April 2020 Earnings

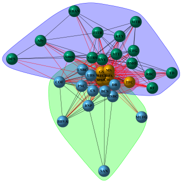
Figure 6: Circle Graphs: Pre-crisis vs crisis periods

Back

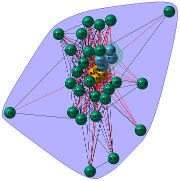
Cluster analysis



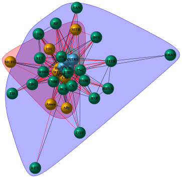
Panel A. Oct 2019 Earnings



Panel B. Jan 2020 Earnings



Panel C. April 2020 Earnings



Panel D. July 2020 Earnings

Figure 7: Cluster analysis across earning release periods

Back

SRISK and Eigenvector Centrality: Travelled Distance

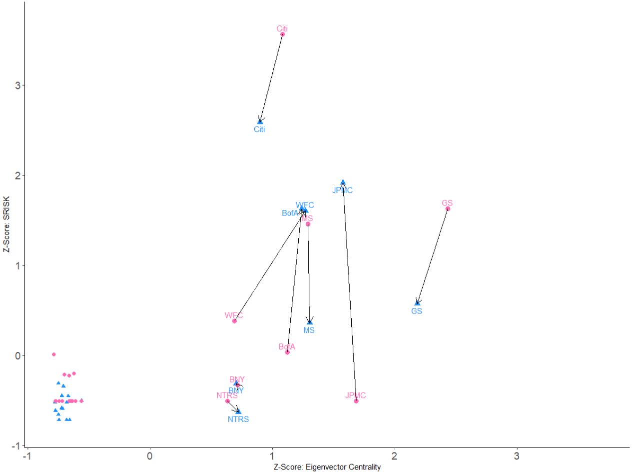


Figure 8: SRISK versus eigenvector centrality z-scores: Pre-crisis versus crisis peak

Back