Coronavirus' effects on the Stock Market Network
by
Benjamin Liang & Isaac Garcia

Intro

Unlike the 2008 crisis where only one sector was initially affected, the onset of Covid-19 in the United States has quickly affected all sectors of the market. Because of the unprecedented nature of this, we wanted to do a project that helped us understand what was going on in the market. We do this by doing back of the envelope calculations with CAPM and through visualization of a market network cascade with software called Gephi.

CAPM

CAPM is a financial mathematical model that describes the relationship between systematic risk and expected payoffs of a stock.[1]

$$\mathbb{E}[r_i - r_b] = \beta_i \mathbb{E}[r_m - r_b]$$

where

$$\beta_i = \frac{\mathrm{Cov}(r_i, r_m)}{\mathrm{Var}(r_m)}$$ [2]

Where beta is the risk/reward conversion factor of a specific stock.
The L2 linear regression is used on time series returns data to estimate beta.

$$y_t = \alpha + \beta x_t + \epsilon_t$$

$$\widehat{\beta} = \frac{\mathrm{Cov}(x_t, y_t)}{\mathrm{Var}(x_t)} = \frac{\sum_{t=1}^{T}(x_t - \overline{x})(y_t - \overline{y})}{\sum_{t=1}^{T}(x_t - \overline{x})^2}$$

$$\widehat{\alpha} = \overline{y} - \widehat{\beta}\overline{x}$$ [3]

The constant term alpha is regarded to be a stock efficiency shifting factor for it's risk factor.
A goal of this project is to see how market sectors' alpha and beta have changed due to coronavirus.

---

[1] Kenton, Will. "Capital Asset Pricing Model (CAPM)." Investopedia. 2020
[2] Motohiro Yogo, "ECO 362: Financial Investments Capital Asset Pricing Model," Princeton University, Princeton, NJ, Week 6. Pg. 8
[3] Ibid. pg.13-34

Methodology

To calculate sector beta, we've compiled a list of ETFs:

Communication Services:	IYZ
Consumer Discretionary:	XLY
Consumer Staples:	XLP
Energy:	XLE
Financials:	XLF
Health Care:	XLV
Industrials:	XLI
Information Technology:	XLK
Materials:	GDX
Real Estate:	VNQ
Utilities:	XLU

This list of sectors is from Fidelity, and the respective ETFs were determined by market cap i.e. whoever had the largest. The returns of these ETFs are used to represent the overall return of each sector. We believe that the dates representative of the Covid-19 outbreak would be 2020's Jan, Feb, Mar, Apr plus the previous year's December. We compare the betas and alphas of these dates to the betas and alphas of 2019's Jan, Feb, Mar, Apr plus the previous year's December to control for different quarterly market environments.

The Dataset

All the data, along with the RCode, will be uploaded to blackboard as a folder called CAPM. The date range of the daily excel data is 03 Dec, 2018 to 30 Apr, 2019 and 02 Dec, 2019 to 30 Apr, 2020.

Riskless Returns

We downloaded daily yield data from the St. Louis Federal Reserve's website for 1-month Treasuries. The yields are per annum so we divided these returns by 365 so they represent daily returns. We choose the 1-month Treasuries since they are short-term and have the most liquidity.

Market Returns

As explained why in the progress report, we believe that the Wilshire 5000 is the index that best represents the market's return. We did not download this data from WRDS since, as it turns out, WFVK is an inactive ETF. Luckily, the St. Louis Federal Reserve Bank's website has closing price data for this index so we downloaded the market return data from them.

Sector Returns

The closing price data was downloaded from WRDS (TAQ) as planned.

Since the market returns and the sector returns datasets are given as closing prices, we converted these into returns before we ran CAPM.

Results

Last Year:

	alpha	beta	sector
IYZ	-0.003949688	0.8722940	Communication Services
XLY	0.029707872	1.0736189	Consumer Discretionary
XLP	-0.011012574	0.5958201	Consumer Staples
XLE	-0.076438982	1.0549071	Energy
XLF	-0.018635428	0.9316416	Financials
XLV	-0.116010093	0.9058505	Health Care
XLI	0.004040486	1.0730302	Industrials
XLK	0.060405349	1.2599998	Information Technology
GDX	0.092121318	-0.2590363	Materials
VNQ	0.018428293	0.5800051	Real Estate
XLU	0.024362019	0.3206799	Utilities

This year:

	alpha	beta	sector
IYZ	-0.021323626	0.7879643	Communication Services
XLY	0.030451284	0.9344079	Consumer Discretionary
XLP	-0.010888152	0.7325742	Consumer Staples
XLE	-0.272417033	1.3115231	Energy
XLF	-0.158113251	1.2021740	Financials
XLV	0.065647417	0.8311257	Health Care
XLI	-0.136430100	1.0475480	Industrials
XLK	0.140600830	1.1050836	Information Technology
GDX	0.299594074	0.3219928	Materials
VNQ	-0.081924007	1.0966077	Real Estate
XLU	0.001376896	0.9901846	Utilities

The percent change of alpha/beta from last year to this year:

	alpha	beta	sector
IYZ	439.881305	-9.667577	Communication Services
XLY	2.502407	-12.966516	Consumer Discretionary
XLP	-1.129814	22.952261	Consumer Staples
XLE	256.384958	24.325927	Energy
XLF	748.455143	29.038253	Financials
XLV	-156.587677	-8.249137	Health Care
XLI	-3476.576592	-2.374786	Industrials
XLK	132.762219	-12.294938	Information Technology
GDX	225.216876	-224.304113	Materials
VNQ	-544.555594	89.068641	Real Estate
XLU	-94.348184	208.776620	Utilities

Last year mean/variance:

	mean	variance
alpha	0.0002744157	0.003375172
beta	0.7644373493	0.188723157

This year mean/variance:

	mean	variance
alpha	-0.0130387	0.02380027
beta	0.9419260	0.07325456

Percent change of mean/variance from last year to this year:

	mean	variance
alpha	-4851.44005	605.15712
beta	23.21821	-61.18412

Commentary:

The sectors who's betas have increased are: consumer staples, energy, financials, real estate, and utilities, and the sectors who's betas have decreased are: communication services, consumer discretionary, health care, industrials, information tech, materials, as demonstrated by the percent change of alpha/beta table.

From all this, we believe that the percent change of variance is an interesting result. The variance of alpha has increased, and the variance of beta has decreased. You can see the "tightness" of the alpha graphs (figure 1 and 2) decrease and the beta graphs (figure 3 and 4) increases with time. From intuition, we attribute the tightening of beta to a broad market selloff. We hypothesize that people were more or less blindly/uniformly liquidating assets, shoring up cash for downturn. If this is true, the market direction would be more homogenous in the bearish direction, so the slopes measured in the linear regression would also be more homogenous. We hypothesize that the increase in variance of the alpha is in response to the unnatural homogeneity of the risk factor beta. Since stocks now have similar risk but with still differing underlying market fundamentals, there should exist more arbitrage opportunities which would increase the variance of alpha. More research should be done to prove/disprove/investigate this hypothesis. One thing to keep in mind is that this back of the envelope calculation was done with equal weighting for the mean and variance; the market caps of these sectors are clearly different.

Market Cascade Visualization

In order to relate our project more to the course we wanted to visualize the stock market as a network. We found guidance in a paper written by Stanford undergrads for a machine learning course[4]. Their objectives were different from ours as they wanted to identify different sectors via correlation measures and create portfolios based on centrality measures in their network. However, we were interested in their methodology on how to model the stock as a network i.e. the part where the paper models network edges as correlation between stocks. We followed sections 4.2 and 5.2 of this paper to visualize the stock market and model returns over consecutive time periods. We opted to not do computations to identify market sectors, but rather, we just assigned stocks to sectors ad hoc during beforehand.

Data

All the data and code will be uploaded on blackboard in a folder called MarketCascade.

In the Stanford paper, the researchers examined the whole U.S. stock market. We were not interested in visualizing the whole market, just the parts that we thought would hurt/benefit the most from Covid-19, so we made an ad hoc list of all the companies and industries we're interested in. This list of companies, along with the tag of which industry it belongs to, is in the NewSectors excel sheet. We downloaded the closing price data for the list of stocks (common stock only) from the dates December 2nd, 2019 to May 1st, 2020 from WRDS (TAQ).

Upon completing the project, we believe it would've been beneficial to have modeled the entire stock market in our network. However, we were initially hesitant to do so since we initially planned on using NetworkX (python) instead of Gephi to print our graphs. NetworkX is not good at processing thousand node graphs.

Calculations

We followed the methodology of section 4.2 of the Stanford paper to do these calculations.

Two things are needed so Gephi can graph the market network: an adjacency matrix and list of node colors. We calculated the correlation matrix of all stocks from December 2nd, 2019 to May 1st, 2020. Then, the adjacency matrix was created by applying this rule on the correlation matrix:

$$A_{ij} = \begin{cases} 1 & \text{if } c_{ij} >= \theta \text{ and } i \neq j \\ 0 & \text{otherwise} \end{cases}$$ [5]

where theta is the threshold you set.

[4] Sun Wenyue, Chuan Tian and Guang Yang. "Network Analysis of the Stock Market." (2015).

[5] Ibid Section 4.2

The nodes coloring list corresponds to the stock returns list. We calculated returns in seven time periods: two 1-month periods from December 2nd, 2019 then 5 tri-weekly periods till May 1st, 2020. The list of dates is: 2020-01-02, 2020-02-03, 2020-02-24, 2020-03-02, 2020-03-23, 2020-04-13, 2020-05-01. The reason for this non-uniform partitioning is that we wanted to have a two month buffer of market normalcy in our data, but also high visual resolution in the period of market change. We didn't want too many static graphs during the market normalcy period. Returns were calculated by dividing the closing price of the listed date by the closing price of the start date, December 2nd, 2019; by the end, you should have seven lists for each period. We followed the Stanford paper's guidelines by setting all returns less than -20% red, all returns between -20% and 20% (inclusive) green, and all returns larger than 20% blue for our node color vector. [6]

Visualization

Once we have the adjacency matrix and our code color list, we were ready to begin graphing the market network in Gephi. There were many different algorithms available in the software to graph our network. We decided to use the Fruchterman Reingold algorithm as it generated the clearest organization of nodes for a graph of our size. We later noticed that the Stanford paper also used the Fruchterman Reingold algorithm to graph their network[7].

However, we soon realized that we would need to change 2 of our parameters. We noticed that we had far too many sectors, originally, in our data as our network provided no visual distinction between the sectors. Therefore, we consolidated (previously-listed sectors) like cruises and hotels into one called travel. We did this until we had 13 sectors. We consolidated these based on our intuition. For a list of the sectors we created, please reference the legend of our sectors graph.

We also realized that we needed to raise the value of our theta parameter from our initial value of 0.7 (the theta in the Stanford paper[8]) as our initial adjacency matrix was too interconnected to provide an adequate visualization of the different sectors of the economy. After tinkering with the parameter, we found that theta equal to 0.9 generated a clearer separation between the different sectors while still maintaining connections between the sectors. We believe the reason why these stocks were so highly correlated was because of the time frame we used to calculate the correlation. Our data included the massive general downturn the market experienced in March, and we believe that this downturn created the high correlation between stock across different sectors.

Findings

As previously stated, we calculated returns from December 2nd, 2019 to January 2nd, 2020, February 3rd, February 24th, March 2nd, March 23rd, April 13th, and May 1st, 2020. We reduced the time frames from monthly to tri weekly in order to better capture the fall of the market. In our first snapshot we see that the market as a whole was doing well. However, by

---

[6] Sun Wenyue, Chuan Tian and Guang Yang. "Network Analysis of the Stock Market." (2015). Section 5.2
[7] Ibid. Section 5.1
[8] Ibid Section 4.2

February 3rd, our second snapshot shows some stock fall to -20% returns, which is when we color a node and its edges red, before the rest of the market does. The first three stocks to fall into this description were PSX, RLH, and WWE. The fall of PSX is logical due to the decrease in overall movement due to quarantines and fear. We have seen the dramatic fall of gas prices all over the news, so this is no surprise. RLH is a hotel company so it makes sense that this would fall immediately due to fear of travel during this pandmeic. Finally, we see WWE, World Wrestling Entertainment, also fall. This makes sense due to the nature of this industry requiring a venue filled with people and the close proximity of their actors during performances. This kind of industry is unsustainable during this pandemic. We can now see that the pandemic is initially hitting sectors of the market which our intuition would agree with. However, since the list of these stocks is so small, this change could just be noise.

By March 2nd, we can see on our fourth snapshot that the red has spread to other stocks near the original three that fell. The sectors which the red spread further into were Airlines, Airlines/Car manufacturing, Travel, and Gas. This makes sense as the same fundamental reasons why Travel and Gas originally fell are beginning to grip the market more. There is too much fear to travel anymore even if certain parts of the country and world aren't in quarantine at this point.

The dramatic fall the stock market experienced by March 23rd was captured by our fifth snapshot. We can see the interconnected center is completely in the red. The majority of this cluster was composed of Airlines, Air/Car manufacturing, Restaurants, Travel, Gas, Finance and parts of Tech. This is following the trend established by the previous snapshots. The addition of Finance in the red also makes sense as this sector is inherently tied to many other sectors. We also noticed that Tech giants, Apple, Google, and Microsoft, survived the downturn. Other sectors that survived the downturn were Health, Utilities, Consumer Staples, and Real Estate. We believe these sectors survived this fall because during a crisis people are ensuring that their necessities for living are covered first.

Our sixth snapshot and seventh snapshots illustrate the recovery the market has experienced since the beginning of April till April 13th and May 1st respectively. There wasn't a significant difference between the two snapshots as to which sectors were recovering. Both snapshots show that Tech and parts of Restaurant, Communication Services, and Industrial are now recovering. The recovery of Communication Services makes sense given that more people are staying home due to quarantine. The Restaurant sector is also recovering which seems un-intuitive given the circumstances. It's interesting to note that the sectors which are beginning to recover are on the opposite side of where the downturn began on the network. The sectors which survived the initial fall are also continuing to prosper. The way the virus has managed to grip the market in such little time is shocking.

Conclusions

Our CAPM analysis seems to indicate that there is a large uniform market sell off however our networks analysis gives a more nuanced view. From our network, we can now see which sectors are continuing to be impacted by the crisis, which sectors are recovering, and which sectors have been relatively unimpacted by the crisis. The Airlines, Aviation/Car manufacturing, Gas, Travel, Finance sectors are still in the red. Tech and parts of the

Restaurant, Industrial, and Communication Services sectors are beginning to recover. Big Tech companies, Consumer Staples, Utilities, Real Estate, were mostly spared during these time frames. The sectors hurting the most are the ones we would guess via intuition. However, restaurants beginning to recover seems odd. This perceived recovery may be because we sampled too many major restaurants like Mcdonalds, Wingstop, etc which are unlikely to truly represent how the industry as a whole is doing during this crisis.

Sources:

Sun, Wenyue, Chuan Tian and Guang Yang. “Network Analysis of the Stock Market.” (2015).

Mullins, David W. “Does the Capital Asset Pricing Model Work?” *Harvard Business Review*, January 1982

Motohiro Yogo, “ECO 362: Financial Investments Capital Asset Pricing Model,” Princeton University, Princeton, NJ, Week 6.

Kenton, Will. “Capital Asset Pricing Model (CAPM).” Investopedia. Investopedia, January 29, 2020. https://www.investopedia.com/terms/c/capm.asp.

“Wilshire 5000.” Wikipedia. Wikimedia Foundation, March 7, 2020. https://en.wikipedia.org/wiki/Wilshire_5000.

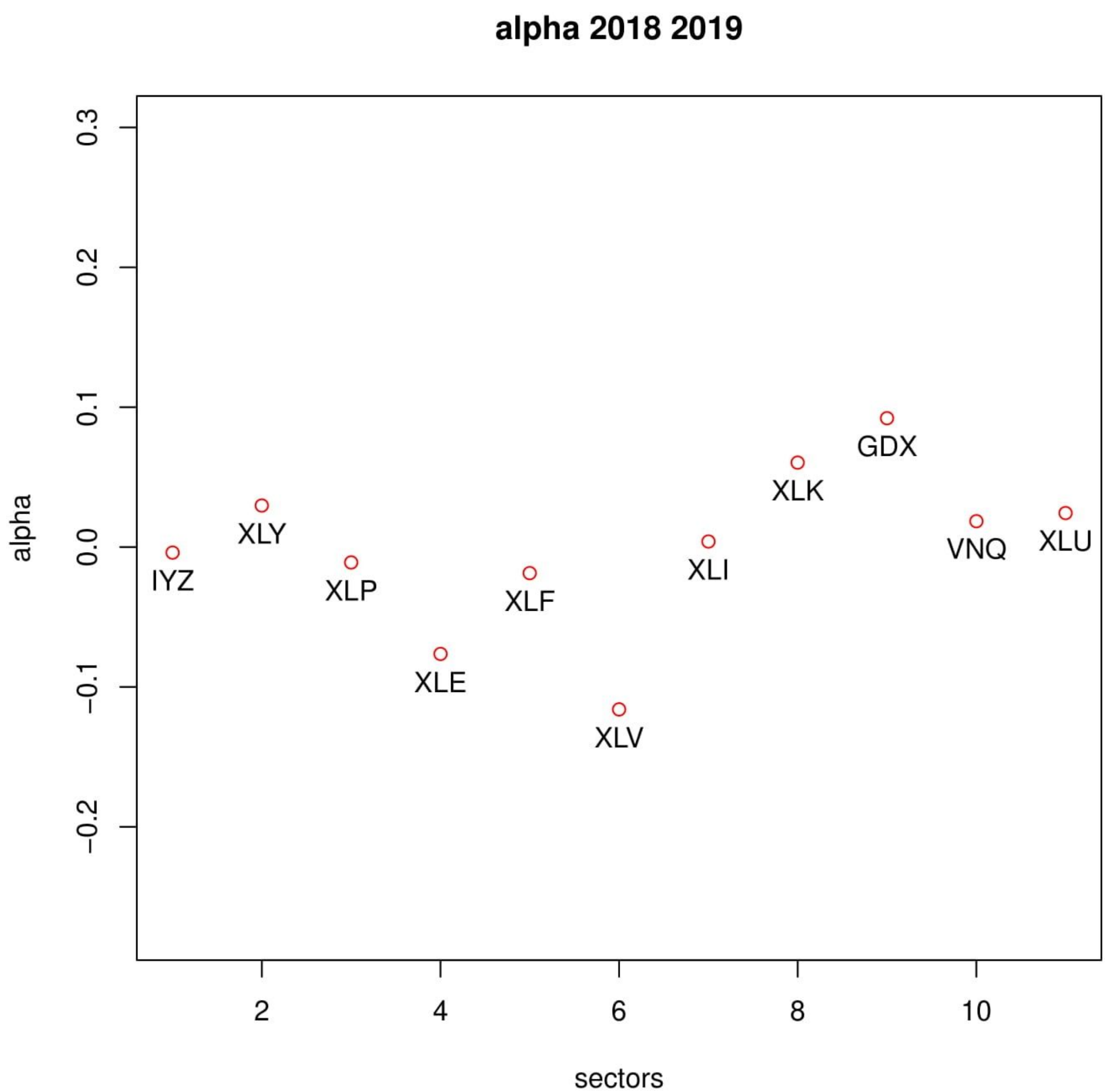

alpha 2018 2019
alpha
sectors
0.3
0.2
0.1
0.0
−0.1
−0.2
2
4
6
8
10
IYZ
XLY
XLP
XLE
XLF
XLV
XLI
XLK
GDX
VNQ
XLU

Figure 1.

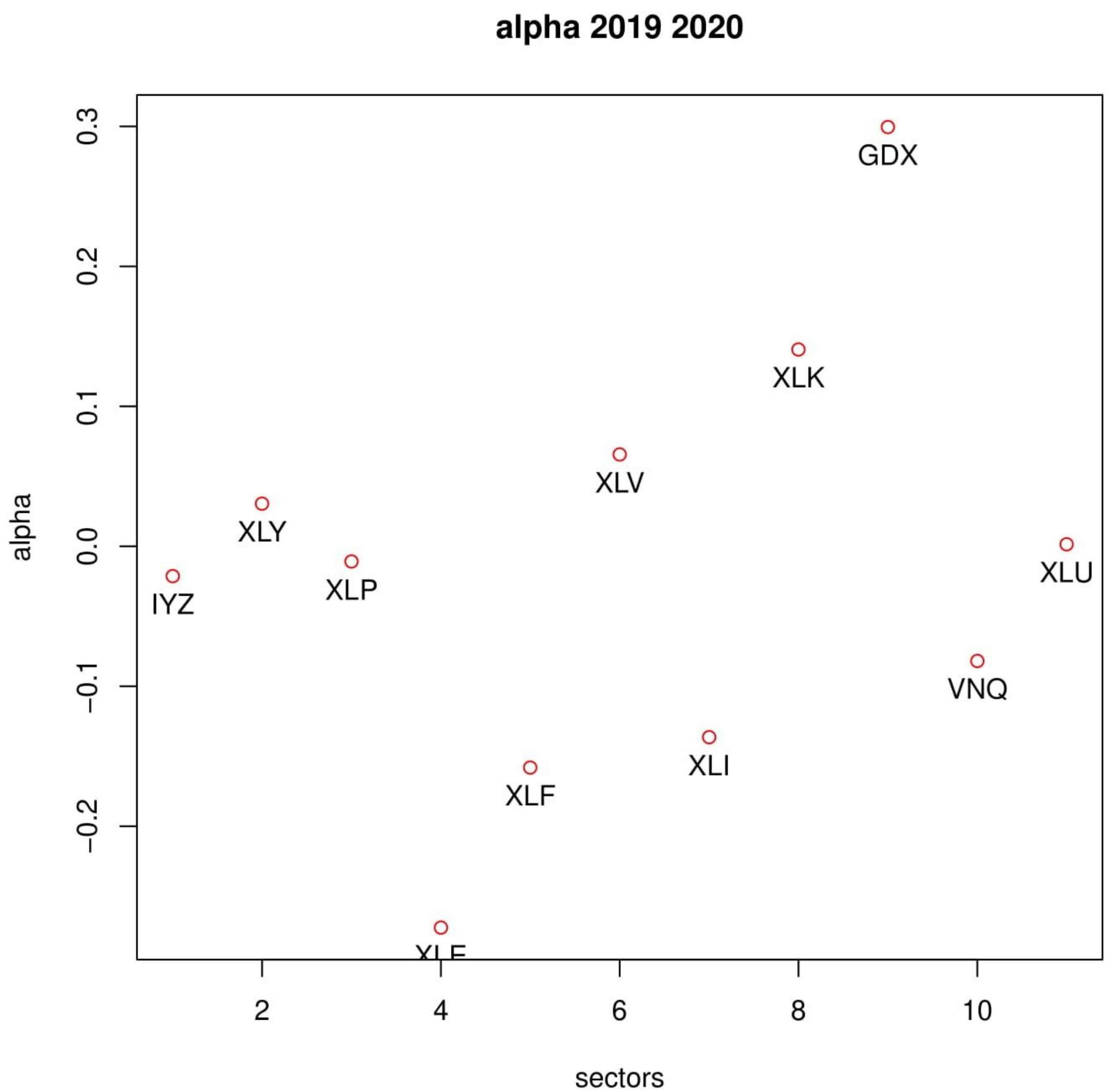

alpha 2019 2020
alpha
sectors
0.3
0.2
0.1
0.0
−0.1
−0.2
2
4
6
8
10
GDX
XLK
XLV
XLY
XLP
IYZ
XLU
VNQ
XLI
XLF

Figure 2.

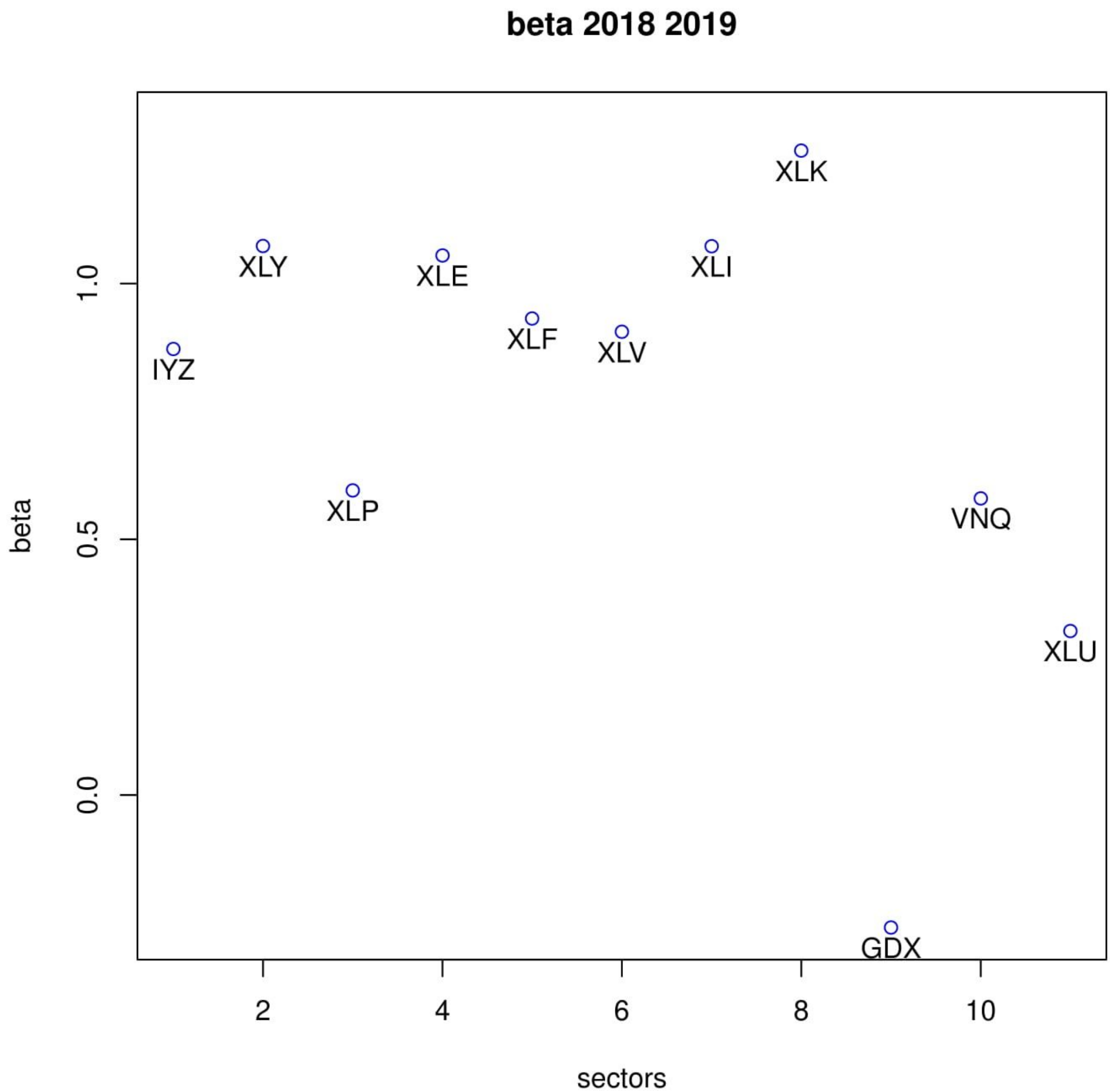

beta 2018 2019
beta
1.0
0.5
0.0
IYZ
XLY
XLP
XLE
XLF
XLV
XLI
XLK
GDX
VNQ
XLU
2
4
6
8
10
sectors

Figure 3.

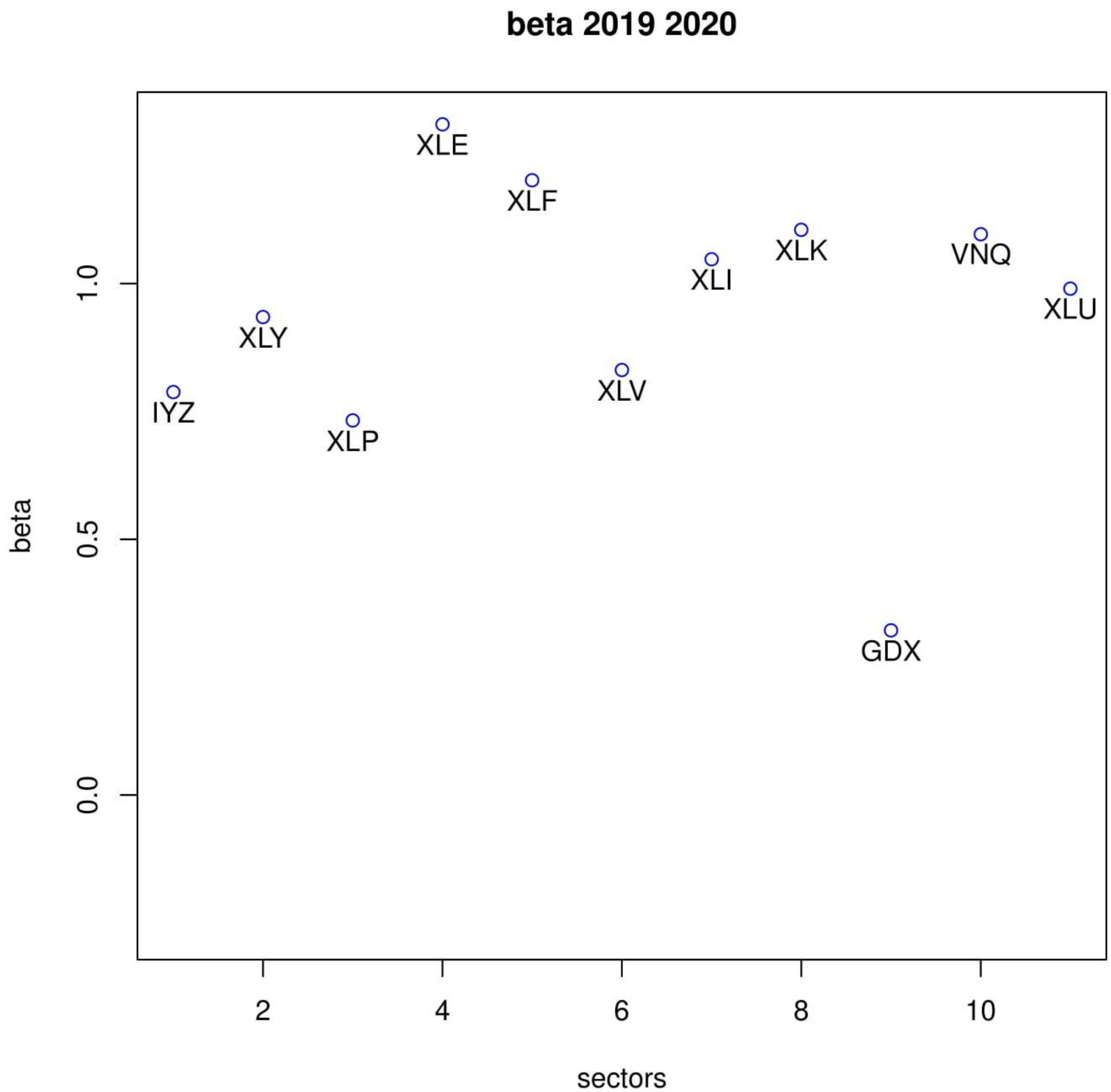
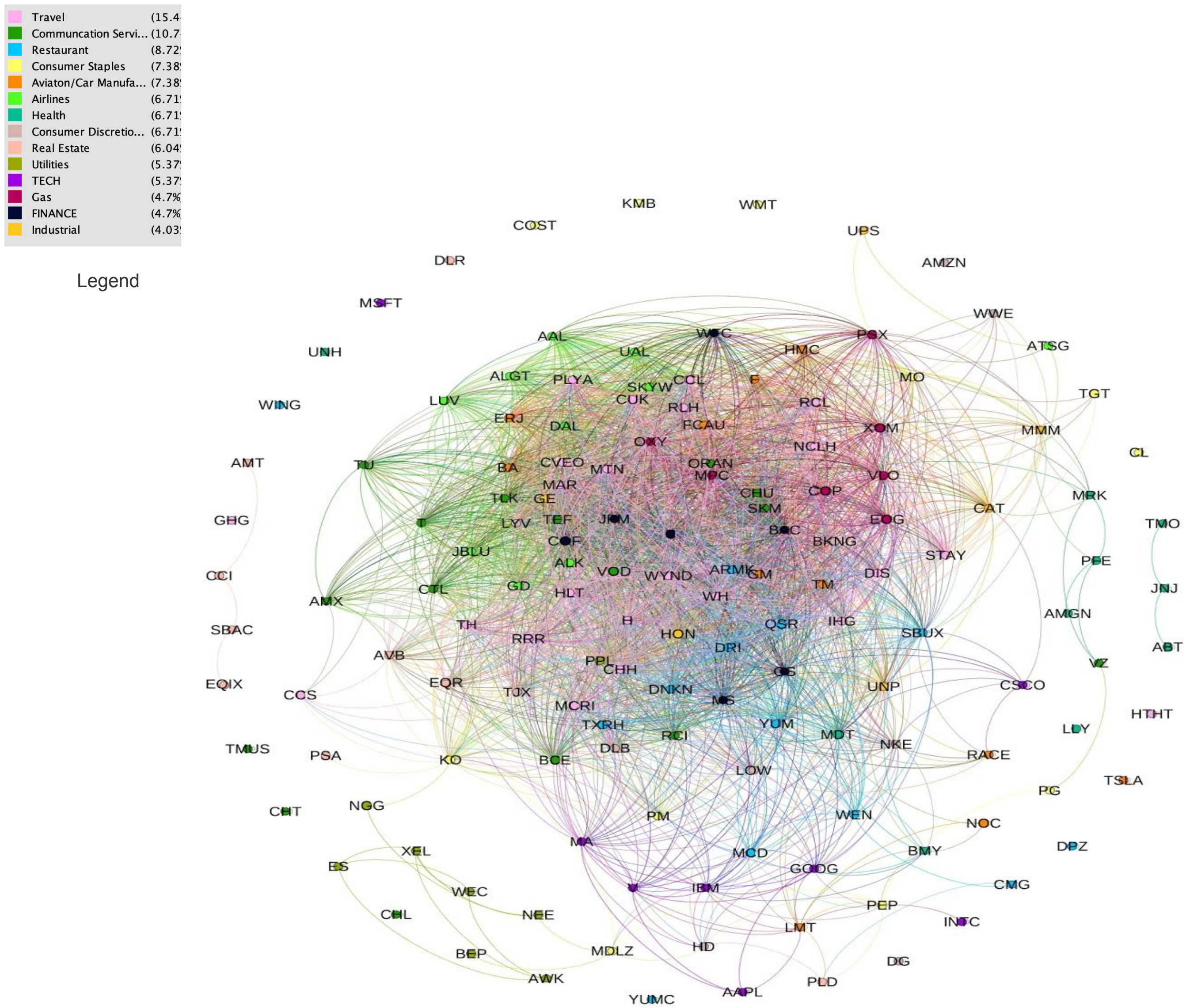
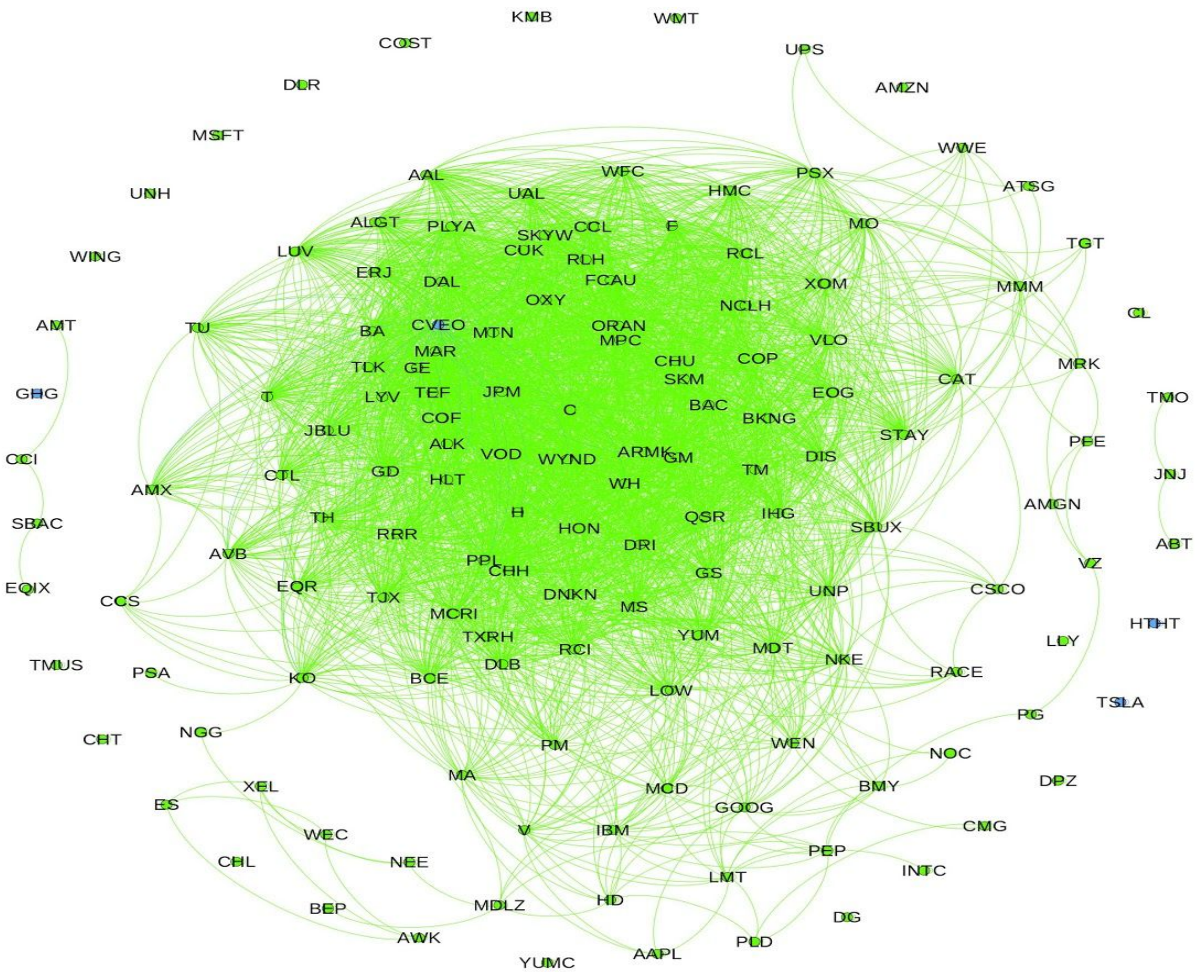

beta 2019 2020
XLE
XLF
XLK
VNQ
XLI
XLU
XLY
IYZ
XLP
XLV
GDX
beta
1.0
0.5
0.0
2
4
6
8
10
sectors

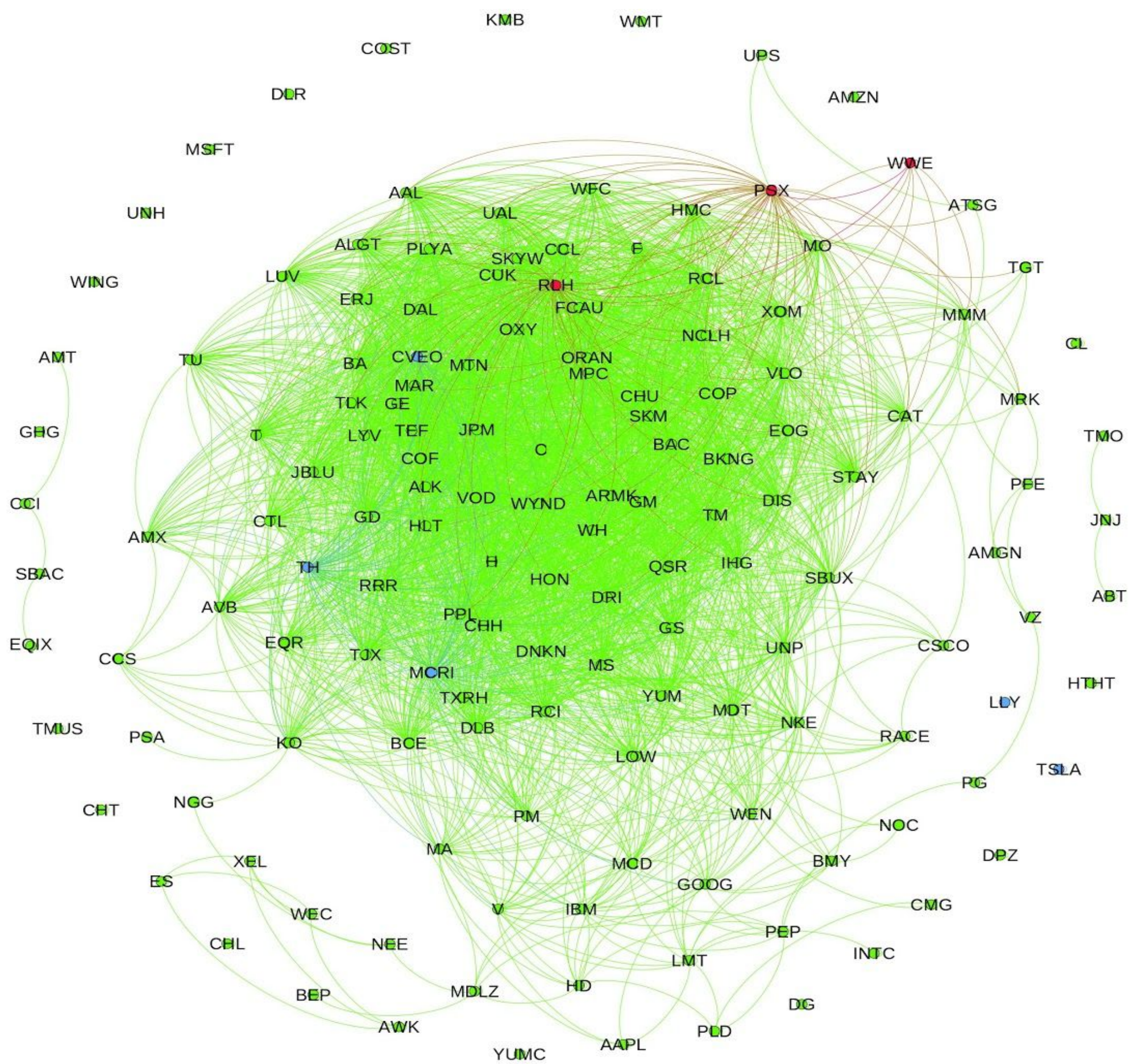
Figure 4.


Travel (15.4
Communcation Servi... (10.7
Restaurant (8.72
Consumer Staples (7.38
Aviaton/Car Manufa... (7.38
Airlines (6.71
Health (6.71
Consumer Discretio... (6.71
Real Estate (6.04
Utilities (5.37
TECH (5.37
Gas (4.7%
FINANCE (4.7%
Industrial (4.03
KMB
WMT
COST
UPS
DLR
AMZN
MSFT
WWE
AAL
WFC
PSX
HMC
ATSG
UNH
UAL
ALGT
PLYA
SKYW
CCL
MO
TGT
WING
LUV
CUK
RLH
RCL
ERJ
DAL
FCAU
XOM
MMM
AMT
OXY
NCLH
CL
TU
BA
CVEO
MTN
ORAN
MPC
VLO
MAR
CHU
COP
MRK
GHG
TLK
GE
SKM
CAT
TMO
TEF
JPM
EOG
LYV
COF
BAC
BKNG
JBLU
STAY
PFE
ALK
VOD
ARMK
CCI
WYND
GM
DIS
JNJ
GD
HLT
TM
AMX
WH
AMGN
TH
IHG
QSR
SBAC
H
HON
SBUX
ABT
AVB
RRR
DRI
VZ
PPL
CHH
GS
EQIX
EQR
UNP
CSCO
CCS
TJX
DNKN
MS
MCRI
YUM
HTHT
TXRH
MDT
LLY
TMUS
PSA
DLB
RCI
NKE
KO
BCE
RACE
LOW
TSLA
PG
NGG
CHT
PM
WEN
NOC
MA
XEL
MCD
DPZ
ES
BMY
GOOG
IBM
CMG
WEC
V
PEP
NEE
CHL
LMT
INTC
HD
MDLZ
BEP
DG
PLD
AWK
AAPL
YUMC
Legend

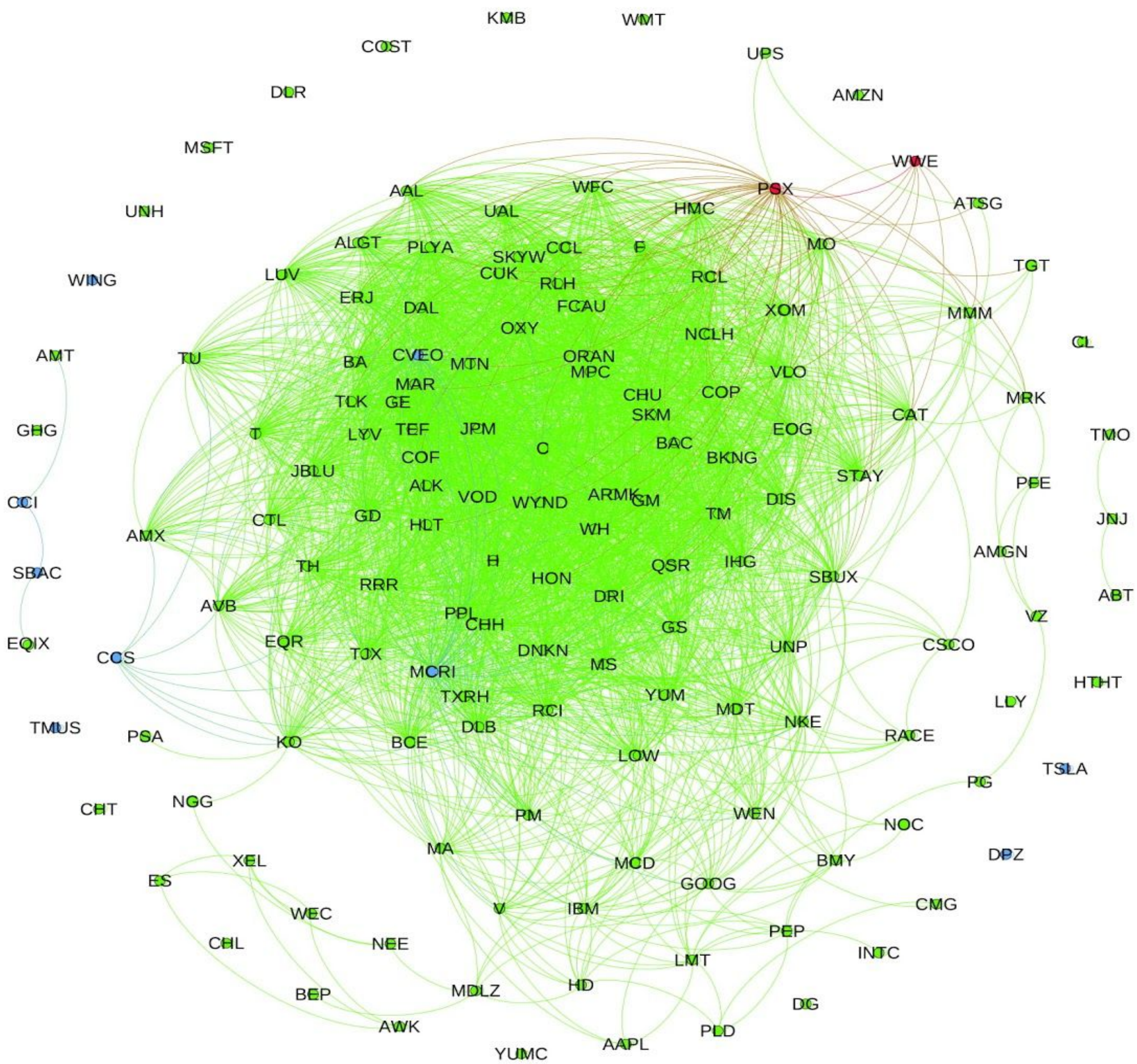
Snapshot 0.



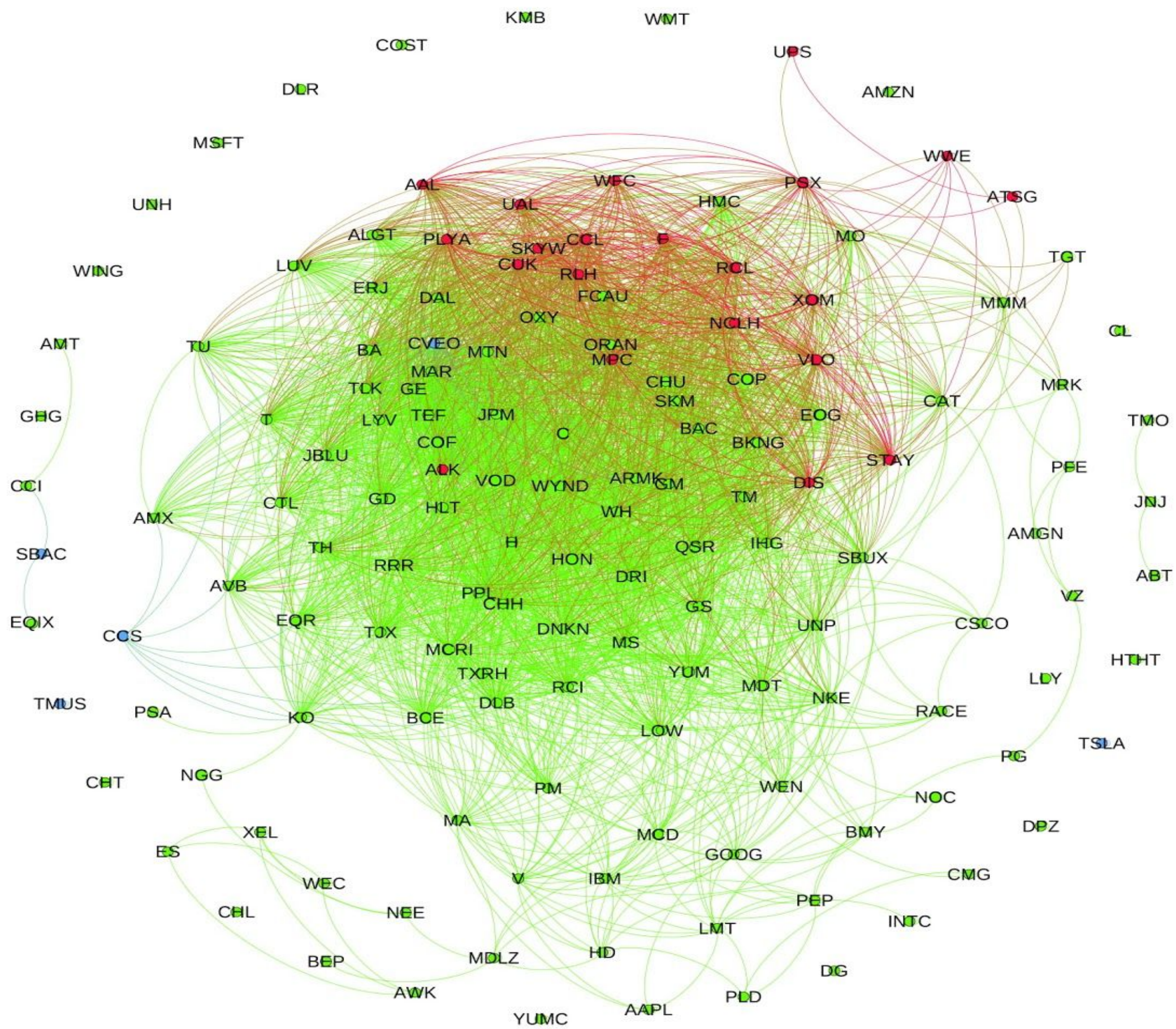
Snapshot 1 (2020-01-02)



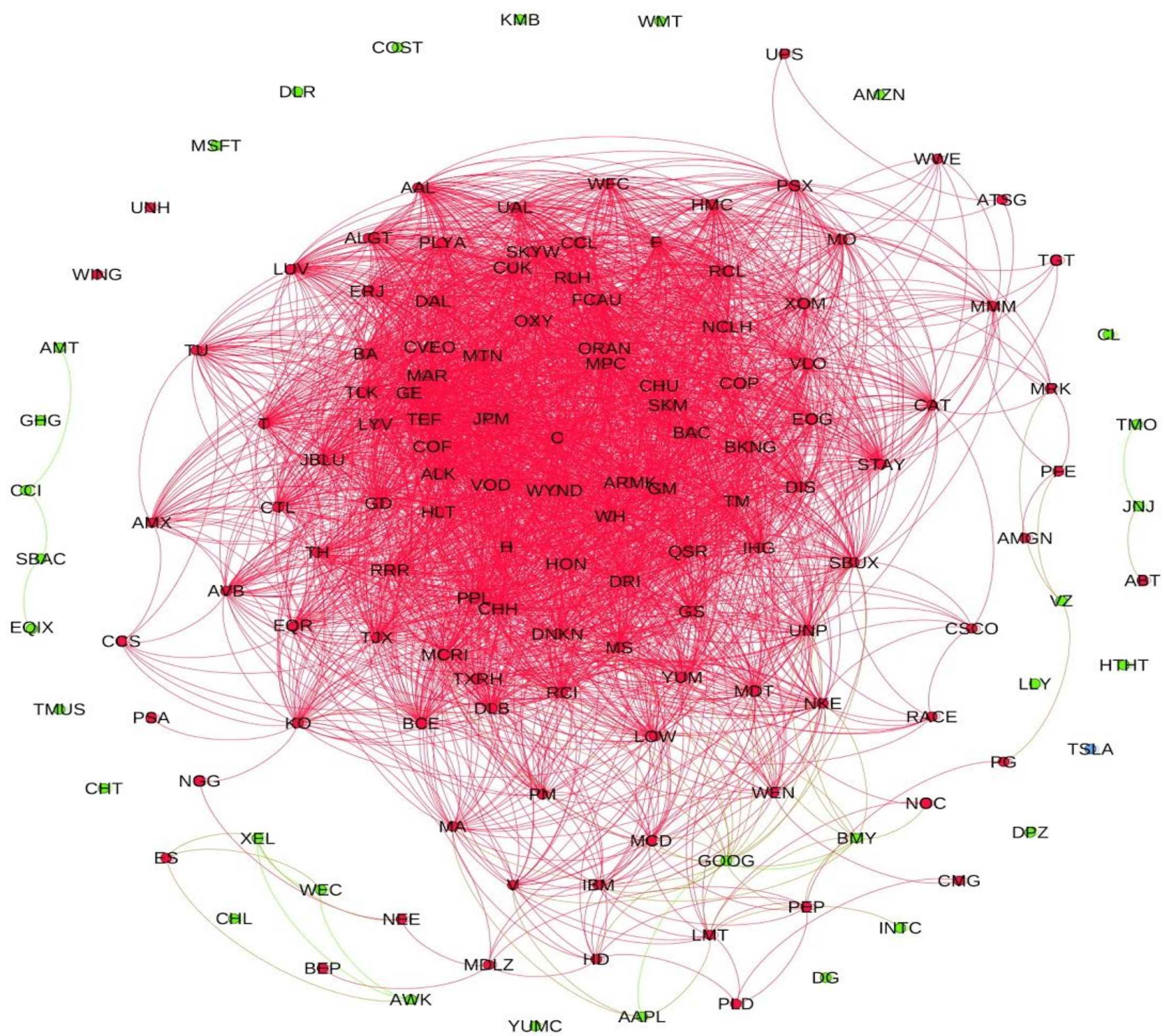
Snapshot 2. (2020-02-03)



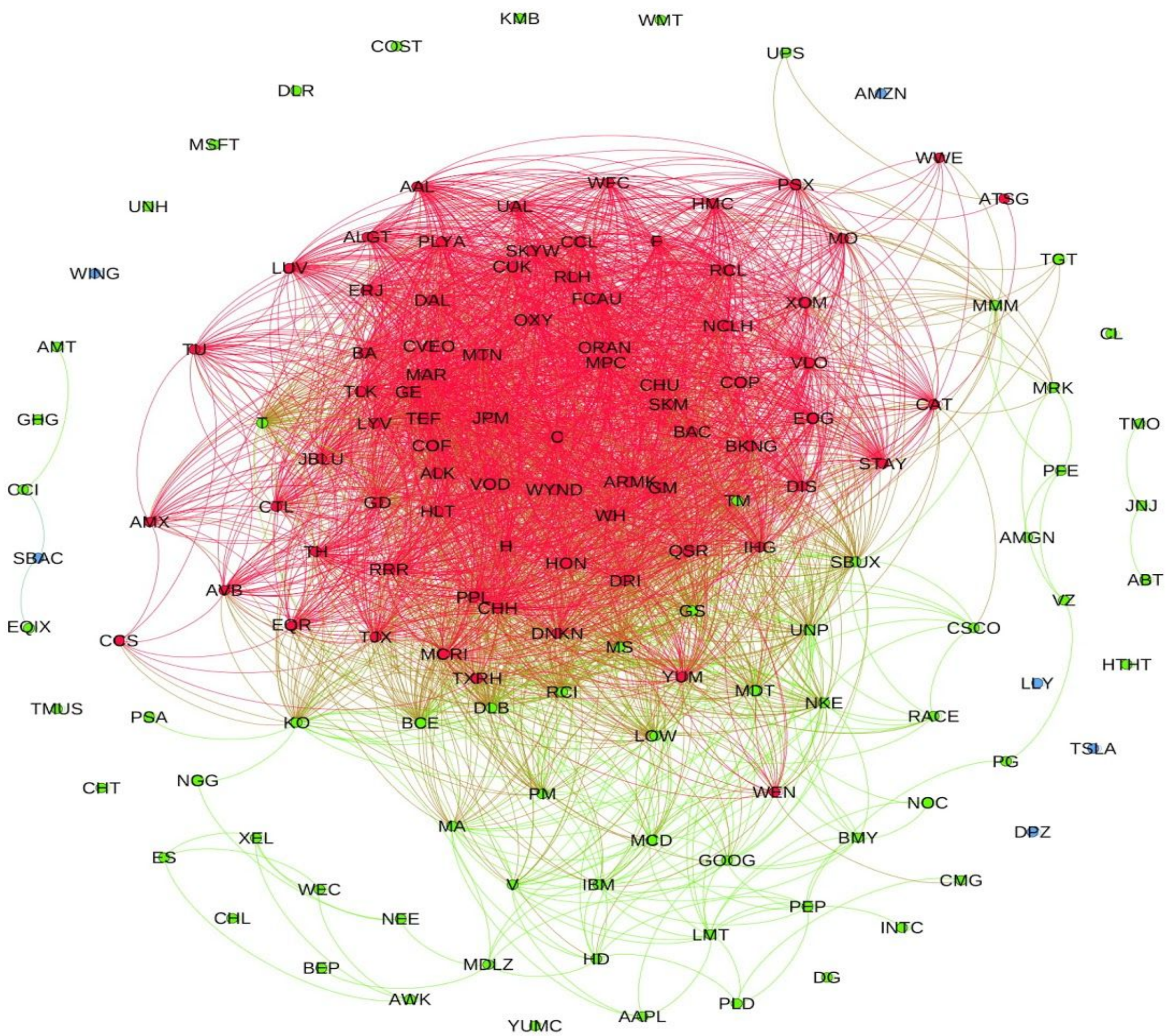
Snapshot 3. (2020-02-24)



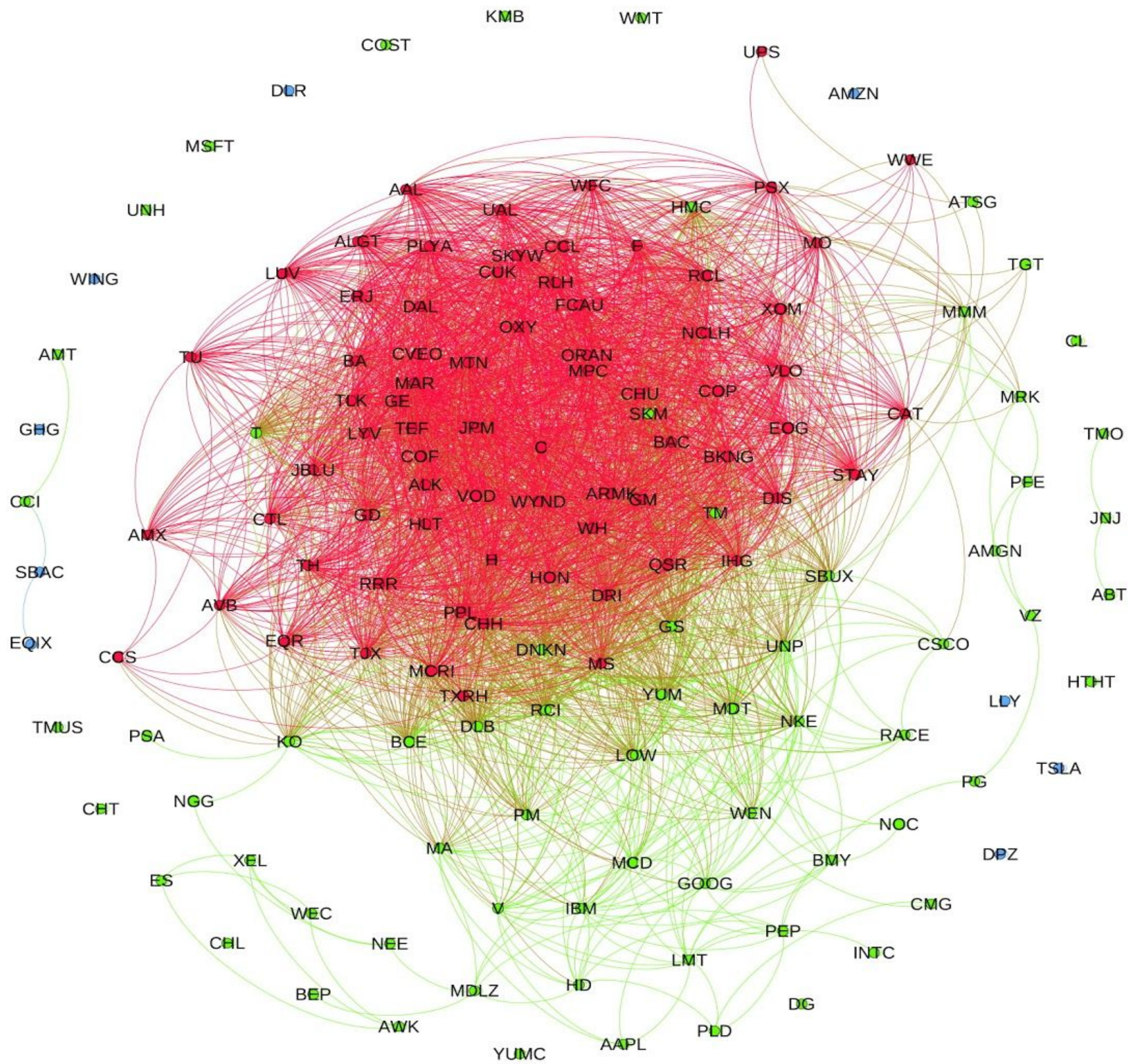
Snapshot 4. (2020-03-02)



Snapshot 5. (2020-03-23)



Snapshot 6. (2020-04-13)



Snapshot 7. (2020-05-01)