Analyzing the Dynamic Evolution of Stock Market Networks Through Temporal Graph Modeling

Omer Idgar, Idan Bibi

August 21, 2023

Abstract

Financial markets' intricate nature demands advanced analytics. This study analyzes S&P 500 stocks (2015-2022) using dynamic correlation network techniques. Employing visualizations, community detection, and centrality metrics, it uncovers relationships among influential market entities. Findings reveal external events' impact on market-wide synchronization and changing correlations across sectors. Centrality metrics track key stocks' network importance shifts. Results depict complex interconnectivity, market maturation, and growing significance of intrinsic stock factors and sectoral drivers. This study underscores network science's role in understanding financial intricacies.

1 Introduction

Financial markets exhibit complex dynamics driven by the intricate web of relationships between influential entities such as stocks, sectors, and external events. Understanding the structure and evolution of these relationships provides valuable insights into market behavior, facilitating more informed investment decisions and policymaking. This study utilizes correlation network analysis, an innovative application of network science, to investigate the interconnectedness and changing dynamics within stock markets.

In network science, financial markets can be modeled as complex adaptive systems consisting of numerous interacting components. Correlation network analysis specifically involves constructing networks where nodes represent stocks, and edges connect strongly correlated stocks based on similarity in price movement. This framework allows the application of graph theory techniques to analyze properties of the system as a whole and relationships between individual nodes. Prior research has demonstrated the utility of correlation networks in elucidating market structure, risks, critical transitions, and the impact of events [1–3].

This study focuses on constructing correlation networks from the stocks comprising the S&P 500 index, a market-capitalization-weighted index of leading U.S. companies. The S&P 500 represents over 80% of the U.S. equity market capitalization, making it a prime benchmark for the overall market. Moreover, the diversity of sectors and industries encompassed within the S&P 500 offers rich potential for analyzing cross-sector and intra-sector relationships. This study constructs correlation networks for the S&P 500 over an 8-year period (2015–2022), capturing both steady market trends and periods marked by volatility stemming from major events like the COVID-19 pandemic.

The analysis employs an array of techniques, including dynamic community detection, clique analysis, centrality metrics, and interactive network visualizations, to derive nuanced, multidimensional insights. The key focus areas are: (1) evaluating the evolution of network structure and properties over time; (2) identifying critical transition periods marked by changing correlation patterns; (3) elucidating relationships and recurring trends between sectors; and (4) tracking shifts in individual nodes' centrality and community affiliations.

This study aims to demonstrate correlation networks' capabilities in monitoring financial systems, detecting early warning signals, and providing context for market movements. The insights gained contribute to strategic investment and policy decisions, keeping pace with an increasingly complex environment. Moreover, the research expands the growing repertoire of correlation network techniques, further establishing their role in investigating economic systems and complex adaptive networks more broadly.

2 Related Work

The exploration of stock correlation networks has become increasingly pivotal in financial research. Initial groundwork by Mantegna [6] led to the construction of networks based on stock price correlations, offering a fresh perspective on market dynamics. Building on this, Onnela et al. [7] focused on the topological aspects of asset graphs, providing deeper insights into market structures and systemic risks. Boginski et al. [2] took a different approach by detecting cliques in the networks, revealing groups of stocks with similar behaviors.

Research has also spanned across diverse markets. For instance, Brida and Risso [3] used minimum spanning trees to uncover hierarchical structures in global stocks, while Huang et al. [5] delved into Chinese stock markets, emphasizing the role of thresholding in network analyses. Song et al. [8] took a longitudinal approach, showcasing the potential of correlation networks in understanding market shifts, especially during significant financial events. Notably, Huang and Wang [4] incorporated network science metrics to identify key banks in the Chinese market, underscoring the link between network centrality and systemic risk.

Collectively, these endeavors demonstrate the growing importance and utility of stock correlation network analysis in the financial domain.

3 Methodology

In this section, we will details the data collection, preprocessing, network construction, and analysis techniques employed to study the dynamic stock correlation networks.

3.1 Data Collection

Daily closing price data for S&P 500 constituent stocks spanning 2015-2022 was extracted using the yfinance[1] Python package. Stock symbols and sector information were gathered from the S&P 500 metadata dataset obtained from Kaggle[2]. In total, pricing data was retrieved for 500 stocks across the 8-year period.

3.2 Data pre-processing

Several preprocessing steps were applied to clean and transform the raw data:

Step One: Stocks with incomplete data across the full time-frame were filtered out to ensure stocks with comprehensive historical data were analyzed.

Step Two: Daily log returns were calculated for each stock by taking the logarithmic difference between consecutive closing prices, as shown in Equation 1:

$$r(t) = ln(P(t)) - ln(P(t-1)) \tag{1}$$

Where $P(t)$ is the stock's adjusted closing price on day t.

Step Three: Missing values in the returns data were imputed using backward fill and forward fill techniques.

This yielded a returns dataframe with daily log returns for every stock and trading day in the studied period. To visualize the impact of preprocessing, Figure 1 compares the raw closing prices and preprocessed log returns for Apple.

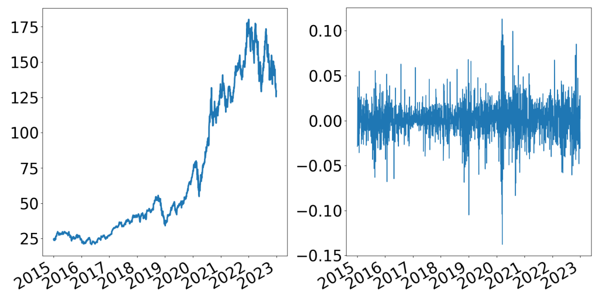

Figure 1: Raw closing prices vs Daily log returns for Apple

3.3 Network Construction

Following established techniques in prior literature [6, 7], periodic correlation matrices were generated by calculating Pearson correlation coefficients between the daily returns of each pair of stocks over rolling annual periods.

These matrices were used to construct dynamic correlation networks, where nodes represent stocks, and edges connect stocks with correlations exceeding a defined threshold.

In Figure 2, we provide a visual example of the correlation network as it appeared in 2017, with nodes colored by their respective sectors.

[1]https://pypi.org/project/yfinance/
[2]https://www.kaggle.com/datasets/tomasmantero/top-tech-companies-stock-price

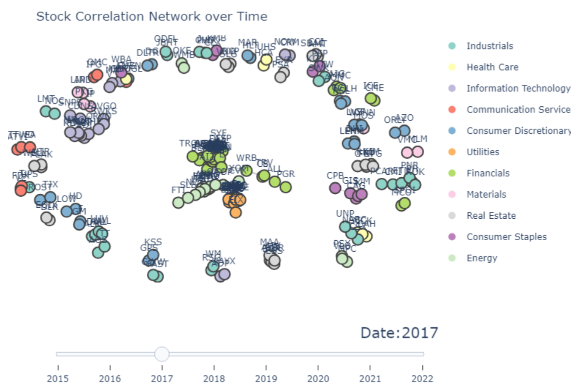

Figure 2: Correlation network in 2017

3.4 Network Analysis

A comprehensive analysis was carried out to understand the complex interconnections and relationships among stocks.

3.4.1 Topology & Community Shifts

The topology and evolution of the overall network structure was examined through tracking foundational properties over time. Density provided insights into market interconnectivity, components indicated cohesion, and centralization identified shifts in dominance.

Interactive visualizations enabled exploration of network transformations across different time periods. The Louvain algorithm was utilized for community detection, uncovering densely connected clusters [1]. Analysis elucidated the longitudinal progression of detected communities in terms of composition, sector alignment, and single stock affiliations.

3.4.2 Sector Relationships

We used average annual correlations and scatterplots to analyze sector relationship trends. Correlation heatmaps identified strongly or weakly connected sector pairs, while time-based subgraphs revealed evolving interconnectivity.

3.4.3 Clique Analysis

Maximal cliques of fully interconnected stocks were identified [9]. Composition and sectoral makeup of detected cliques were analyzed across years. Changes around major events provided perspective on external shocks' impact on local subgraph connectivity. This multifaceted network analysis methodology facilitated a nuanced exploration of complex stock market dynamics from the broader market scale down to individual stock behaviors.

3.4.4 Individual Stock Analysis

Centrality metrics were calculated to quantify individual stocks' positional significance within the network. Interactive visualizations traced stocks' dynamic community affiliations over time. Betweenness centrality indicated bridging ability, closeness reflected reach, and eigenvector captured embeddeness within influential clusters.

4 Results

Our study offers an incisive examination of stock dynamics and sector relationships between the years 2015 and 2022. Through detailed visualization techniques, we present intricate relationships that have evolved over these years.

4.1 Topology & Community Shifts

The network's evolution provides an in-depth look into the intricate dynamics of stock market communities over time.

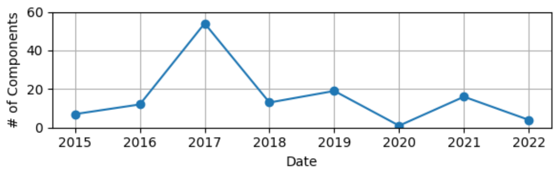

Figure 3: Number of components over time

Figure 3 visualizes the changes in the number of stock market communities. A primary observation from this data is the emergence of a single, dominant component in 2020. This pattern is synchronous with the widespread influence of the COVID-19 pandemic on stock market behaviors, leading to a shift in the market's community structure.

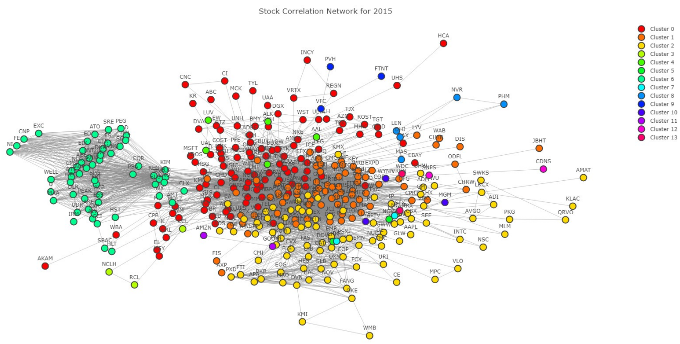

Figure 4: Community detection snapshot for 2015

Expanding further on the dynamics, Figure 4 offers a static representation, focusing on the community landscape in 2015. From this particular snapshot, it is clear that while several communities aligned closely with specific sectors. These deviations, not strictly following sectoral lines, provide intriguing insights into the multifaceted relationships and interactions between stocks, often governed by overarching market dynamics.

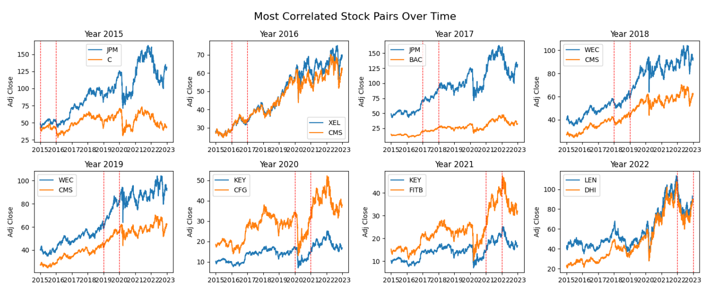

Figure 5: Correlated stock pairs over time

Figure 5 encapsulates the consistent correlations between stocks across multiple years. It's not just about identifying stock pairs with persistent behaviors but also understanding the wider implications for market participants. Recognizing these enduring stock interactions can play a crucial role in strategizing investments, underlining the necessity for a nuanced understanding of long-term stock dynamics.

4.1.1 Sector Relationships

Grasping the intricate interplay among financial market sectors is vital for astute portfolio management and strategy development.

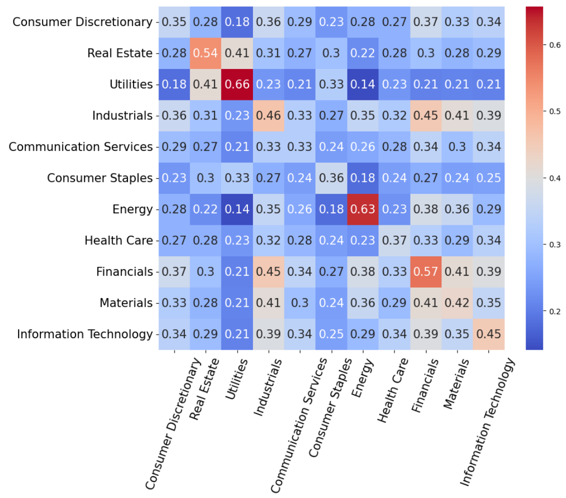

Figure 6: Average sector correlations heatmap over years

Our analysis, visualized in Figure 6, reveals average yearly correlations between sectors, filtering out short-term noise to highlight enduring relationships.

Anomalies in self-correlations, such as those in Consumer Discretionary and Communication Services, hint at their inherent diversity. On the other hand, the robust self-correlations in Utilities, Energy, and Real Estate reflect their relative stability and consistent demand.

Delving into inter-sector dynamics, the 0.45 correlation between Financials and Industrials suggests shared sensitivities to macroeconomic factors. Likewise, the 0.41 correlation between Utilities and Real Estate might be influenced by common reactions to interest rate changes. Notably, many sector pairings exhibit low correlations, indicating distinct behaviors over extended periods.

In summary, this heatmap underscores the importance of understanding sector interdependencies for informed financial decision-making.

4.1.2 Clique Analysis

Clique analysis unveils tightly-knit groups of stocks with high mutual correlations, shedding light on persistent stock interdependencies over time.

The visualization in Figure 7 captures 2015's largest clique, representing a segment of an interactive Plotly-based community analysis. This tool allows users to track the evolution of major cliques yearly, with node colors denoting their sectors. From 2015 to 2019, the Financials sector stood out, suggesting a robust interdependence within its stocks. In contrast, 2020, shaped by the pandemic, saw a diversification of stocks within the main clique. However, 2021 witnessed a resurgence of Financials, highlighting the sector's

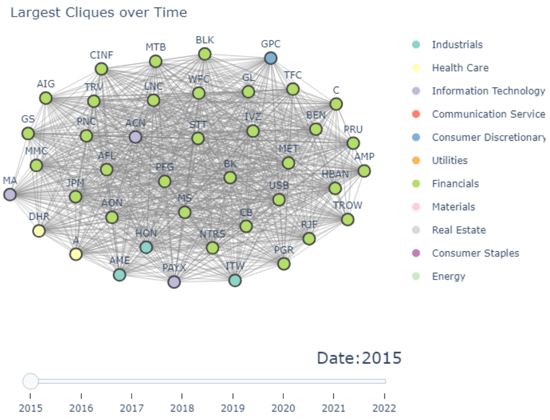

Figure 7: Largest clique in 2015

resilience. By 2022, the mix evolved further, integrating Financials with sectors like Consumer Discretionary and Information Technology, reflecting broader market adaptability.

Overall, clique analysis underscores the dynamic interplay among stock sectors, revealing shifts and patterns over time.

4.1.3 Individual Stock Analysis

This analysis delves into the behavior of individual stocks, emphasizing their centrality measures and sectoral interactions.

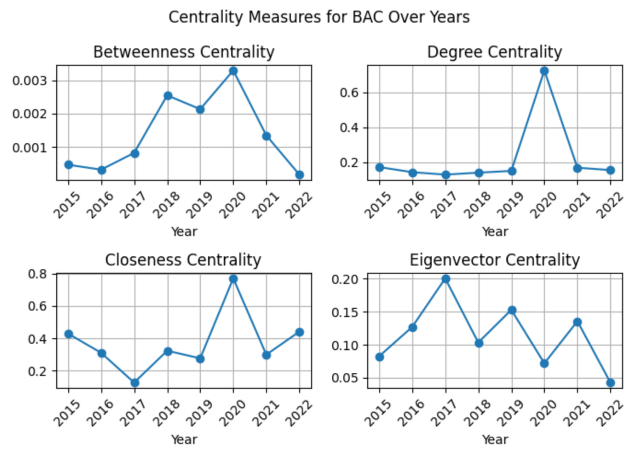

Figure 8: Centrality measures for Bank of America stock

Figure 8 charts centrality measures for the Bank of America Corp. (BAC) from 2015 to 2022. Key observations include BAC's evolving role in bridging stocks as indicated by Betweenness Centrality, its consistent market interactions via Degree Centrality, its unique position during the COVID-19 pandemic revealed by the spike in Closeness Centrality in 2020, and its association with influential stocks peaking in 2017 as suggested by Eigenvector Centrality.

Figure 9 showcases cross-sector correlations for stocks like AAPL (Apple), GOOGL (Alphabet), and AMZN (Amazon). For instance, Apple's ties to both the Information Technology and Consumer Discretionary sectors underscore its software and hardware impacts. Google's reach extends from search to cloud services, and Amazon's influence is felt in both e-commerce and entertainment sectors.

Conclusively, this analysis offers a detailed look into how key stocks maneuver and influence the broader market.

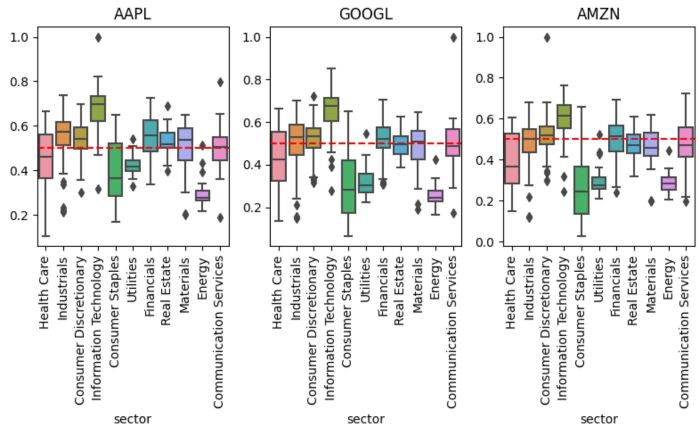

Figure 9: Cross-sector correlations grid for sample stocks

5 Discussion

Our network analysis of the S&P 500 stocks provides an in-depth view of market dynamics, using both broad and specific approaches to explore stock market complexity.

We identify periods of high interconnectedness, often shaped by major events, and chart the changing nature of stock behavior and sector relationships. Our methods also reveal the market's resilience to uncertainties and the enduring influence of key sectors.

The insights gained are valuable for policymakers and investors, helping them anticipate market volatility and manage systemic risks effectively.

We also offer sector-specific findings that indicate economic and interest rate influences, aiding investors in risk management. The study delves into individual stock influences through centrality and cross-sector metrics.

While insightful, the study has limitations. Exploring other global indices and incorporating external economic variables could provide more context. Nonetheless, our research bridges network science and finance, offering clarity on market complexities.

6 Conclusions & Future Work

In this work, we conducted an extensive correlation network analysis of S&P 500 stocks. Multidimensional techniques provided insights into market structure, interconnectedness, dependencies, and dynamics over an 8-year period. The findings reveal fundamental patterns and turning points in stock relationships.

Key contributions include characterization of volatility regimes, quantification of systemic risks, demonstration of intricate sectoral interplay, and insights into enduring stock correlations. The work highlights network analysis' capabilities in investigating financial systems.

Future research can build on these findings by expanding scope, improving methods, and connecting network insights with predictive models. This synthesis of data science, network science, and finance holds promise for comprehensively understanding market complexities.

Acknowledgments

The authors would like to acknowledge the valuable assistance provided by ChatGPT in generating and refining the content presented in this paper.

References

[1] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

[2] Vladimir Boginski, Sergiy Butenko, and Panos M Pardalos. Statistical analysis of financial networks. *Computational statistics & data analysis*, 48(2):431–443, 2005.

[3] J Gabriel Brida and W Adrián Risso. Hierarchical structure of the german stock market. *Expert Systems with Applications*, 37(5):3846–3852, 2010.

[4] Wei-Qiang Huang and Dan Wang. Systemic importance analysis of chinese financial institutions based on volatility spillover network. *Chaos, Solitons & Fractals*, 114: 19–30, 2018.

[5] Wei-Qiang Huang, Xin-Tian Zhuang, and Shuang Yao. A network analysis of the chinese stock market. *Physica A: Statistical Mechanics and its Applications*, 388(14): 2956–2964, 2009.

[6] Rosario N Mantegna. Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, 11:193–197, 1999.

[7] J-P Onnela, Anirban Chakraborti, Kimmo Kaski, and Janos Kertesz. Dynamic asset trees and black monday. *Physica A: Statistical Mechanics and its Applications*, 324(1-2):247–252, 2003.

[8] Dong-Ming Song, Michele Tumminello, Wei-Xing Zhou, and Rosario N Mantegna. Evolution of worldwide stock markets, correlation structure, and correlation-based graphs. *Physical Review E*, 84(2):026108, 2011.

[9] Etsuji Tomita, Akira Tanaka, and Haruhisa Takahashi. The worst-case time complexity for generating all maximal cliques and computational experiments. *Theoretical computer science*, 363(1):28–42, 2006.