Research Paper-A Network Approach To Portfolio Selection

Abstract

This research paper treats financial markets as ‘networks’ with securities as nodes, and the correlation between returns as the links between them.It tries to incorporate this model and the individual performance of each security to determine the best portfolio selection, it compares this method to the original Markowitz portfolio theory. It investigates using ML methods that it is preferable to invest in securities with low centrality . Since they have less correlation with others, it allows us to have a diverse portfolio.Next, the study looks at what factors affect a security's importance in this network. It finds that securities that are older, larger in terms of market capitalization, cheaper (possibly undervalued), and financially riskier tend to be more important in this network. Further strategies are developed using this concept of networks combined with individual performance which yield better out of sample and in sample results compared to traditional methods.

Key Words

Network Theory- A theory which considers financial markets as networks with securities as nodes, and the correlation between returns as the links between them

Centrality- Centrality in network theory measures how important a node is within a network. It quantifies a node's influence, connections, or role in information flow. Different centrality measures, like degree or betweenness centrality, help identify key nodes that can impact the network's structure or function, aiding in various applications like social networks and web search algorithms. So here centrality refers to how a security’s returns are related to overall returns.

Portfolio Selection-Portfolio selection is the process of choosing a combination of investments (stocks, bonds, etc.) to achieve a specific financial goal while managing risk. It aims to maximize returns or minimize risk by diversifying assets. Modern portfolio theory, developed by Harry Markowitz, is a fundamental framework for optimizing investment choices.

Mathematical Formulae and Logic-

$$min_w \sigma_p^2 = w^T \Sigma w \quad subject\ to \quad w^T 1 = 1$$

The solution of Eq. (2) is given by:

$$w^*_{minv} = \frac{1}{1^T \Sigma^{-1} 1} \Sigma^{-1} 1$$

This formula is used for the method of minimum variance, given the sum of weights of stocks in portfolio is 1. 1 refers to column vector with all entries 1. wT1=1 indicates sum of weights has to be 1. Then solving this equation using matrices we get the equation using equation 2 as given.

follows.

$$\hat{w}^*_{minv} = \varphi_{minv}\Omega^{-1}\epsilon$$

where $\hat{w}^*_{i,minv} = w^*_{i,minv} * \sigma_i$, $\varphi_{minv} = \frac{1}{\mathbf{1}^{T}\Sigma^{-1}\mathbf{1}}$ and $\epsilon_i = 1/\sigma_i$.

Final Result

$$\hat{w}^*_{mv} = \varphi_{mv}\Omega^{-1}\hat{\mu}^e$$

where $\hat{w}^*_{i,mv} = w^*_{i,mv} * \sigma_i$, $\varphi_{mv} = \frac{R^e}{\mu^{e\,T}\Sigma^{-1}\mu^e}$ *and* $\hat{\mu}^e_i = \mu^e_i/\sigma_i$.

2.2 The relationship between optimal portfolio weights and asset cen

This is the formula for the method of mean variance, here we also include risk free securities. The excess return matrix is given by meu and here the sum of weights need not be 0 since some wealth may be allocated to risk free securities. Putting these we get the above formula.

$$\hat{w}^*_{minv} = \varphi_{minv}\epsilon + \varphi_{minv}\left(\frac{1}{\lambda_1} - 1\right)\epsilon_M v_1 + \Gamma_{minv}$$

$$\hat{w}^*_{mv} = \varphi_{mv}\hat{\mu}^e + \varphi_{mv}\left(\frac{1}{\lambda_1} - 1\right)\hat{\mu}^e_M v_1 + \Gamma_{mv}$$

This formula is used to calculate the weights taking into account the centrality and network theory. Here the link in the network are assigned vectors vi and eigenvalues lamdai .

The proof for this formula iswhose ith main diagonal element is λ_i. Thus the inverse of Ω could be written as

$$\Omega^{-1} = P\Lambda^{-1}P^T = P * diag(1/\lambda_i) * P^T \tag{A.1}$$

$$\Omega^{-1} = \sum_k \left(\frac{1}{\lambda_k} v_k {v_k}^T\right) = \frac{1}{\lambda_1} v_1 {v_1}^T + \frac{1}{\lambda_2} v_2 {v_2}^T + \ldots + \frac{1}{\lambda_n} v_n {v_n}^T \tag{A.2}$$

From Eq. (7) in Section 2.2, we have $\hat{w}^* = \varphi\Omega^{-1}\hat{\mu}^e$. By adding and subtracting $\varphi\hat{\mu}^e$ from this expression we get

$$\hat{w}^* = \varphi\hat{\mu}^e + \varphi\left[\Omega^{-1} - I\right]\hat{\mu}^e \tag{A.3}$$

Using the preliminary results above-mentioned in this appendix, we know that the matrix $\Omega^{-1} - I$ has the same eigenvectors with eigenvalues equal to $\frac{1}{\lambda_k} - 1$ for $k = 1 \ldots n$. Therefore, A.3 is stated as follows:

$$\hat{w}^* = \varphi\hat{\mu}^e + \varphi\left[\sum_k \left(\frac{1}{\lambda_k} - 1\right) v_k {v_k}^T\right]\hat{\mu}^e \tag{A.4}$$

Given that eigenvector centralities refer to the elements of the eigenvector corresponding to the largest eigenvalue, we define $\Gamma = \varphi[\sum_{k=2}^{n}(\frac{1}{\lambda_k} - 1)v_k {v_k}^T]\hat{\mu}^e$. Then, A.4 is stated as

$$\hat{w}^* = \varphi\hat{\mu}^e + \varphi\left(\frac{1}{\lambda_1} - 1\right) v_1 {v_1}^T \hat{\mu}^e + \Gamma \tag{A.5}$$

$$\hat{w}^* = \varphi\hat{\mu}^e + \varphi\left(\frac{1}{\lambda_1} - 1\right)\left({v_1}^T\hat{\mu}^e\right) v_1 + \Gamma \tag{A.6}$$

Interpretation of Data tables

Table 1
Market capitalization, traded volume, and centrality by economic sectors.
This table reports the total and mean firms' centrality by economic sectors in the d-S&P500 dataset. Market capitalization and traded volume are in millions of dollars and correspond to the end of year 2012. The column Firms gives the number of companies in each economic sector. The columns denoted by % present the percentages with respect to the total of each of the preceding variables.

Economic sector	Firms	%	Market cap.	%	Traded vol.	%	Centrality	
							Total	Mean
Finance, insurance, and r. estate	37	19%	2,254,158	20%	70,788	19%	2.74	0.0741
Mining	17	9%	656,688	6%	26,178	7%	1.23	0.0722
Transp., comm., elect, gas, sanit. s.	27	14%	1,145,092	10%	36,988	10%	1.88	0.0696
Manufacturing	87	44%	5,338,584	48%	156,881	43%	6.01	0.0691
Retail trade	12	6%	745,486	7%	27,130	7%	0.82	0.0682
Services	13	7%	829,671	7%	37,314	10%	0.89	0.0682
Wholesale trade	5	3%	142,423	1%	6011	2%	0.33	0.0656
Construction	2	1%	44,905	0%	3073	1%	0.13	0.0646
Total	200		11,157,008		364,364			

Table 1- It contains data regarding companies of all sectors,their no. of firms,market cap,traded volume and centrality. This table shows that finance sector despite not having the highest market cap,traded volume still has the highest centrality.This leads us to the conclusion that securities that are older, larger in terms of market capitalization, cheaper, and financially riskier have more centrality.

Table 2
Optimal portfolio weights as a function of the individual and systemic stock's dimensions considering a cross-sectional approach v_i is the centrality of stock i.. σ_i is the standard deviation of stock i. SR_i is the Sharpe ratio of stock i. N is the number of observations (stocks) in the cross-sectional regressions. The regression R^2 is adjusted for degrees of freedom. Each row of the table reports OLS estimations of Eqs. (10) and (11), respectively, where *t-statistics* are in parentheses.*, **

	v_i	σ_i	SR_i	N	R^2
$w^*_{i,minv}$	-0.740 (-4.05)***	-1.755 (-6.38)***		200	0.231
$w^*_{i,mv}$	-0.778 (-3.64)***		0.761 (4.77)***	200	0.179

* At 5% level

Table 2- After performing ML analysis on the dataset we get the linear equation with weights of various parameters shown in the table above. The values of vi(vector) indicating centrality is negative for both indicating optimal portfolio weights have a negative correlation with centrality (investing in stocks with less centrality=better portfolio),also w has a negative correlation with standard deviation indicating a better portfolio is one with stocks having less variance.W also has a positive correlation with sharpe ratio indicating a better portfolio is formed by investing in stocks having higher sharpe ratio.

Table 3
Optimal portfolio weights as a function of the individual and systemic stock dimensions taking a time series approach We consider a 60-day rolling window estimation procedure for each variable. Each t denotes a 60-day rolling window. $\bar{v}_t$ is the mean centralities at each t. $(\frac{\sigma_{v,t}}{\bar{v}_t})$ is the coefficient of variation of the centrality distribution at t. π_t is the correlation between centralities, v_{it} and standard deviations, σ_{it} at t. ρ_t is the cross-sectional correlation between centralities, v_{it}, and Sharpe ratios, SR_{it}, at t. N is the number of observations (stocks) in the cross-sectional regressions. The regression R^2 is adjusted for degrees of freedom. Each row of the table reports OLS estimation of Eqs. (14) and (15). *t-statistics* are in parentheses.*, **

	$\bar{v}_t$	$\frac{\sigma_{v,t}}{\bar{v}_t}$	π_t	ρ_t	N	R^2
$\bar{\bar{v}}_{minv,t}$	-3.324	-0.0707	-0.00613		2522	0.573
	(-19.04)***	(-33.80)***	(-26.34)***			
$\bar{\bar{v}}_{mv,t}$	-2.661	-0.0720		0.00627	2522	0.457
	(-10.83)***	(-24.50)***		(17.79)***		

* At 5% level.

Table 3- This table adopts a time series approach as compared to the cross-sectional approach adopted earlier. Here the optimal weights are calculated as a linear equation of rolling mean centrality,COV of centrality,rhoi(t) and pi(t).The negative correlation with mean centrality indicates investing in securities with low centrality is better choice, pi(t) refers to the correlation between centrality,standard deviation so negative value increases a trade off occurs between choosing low variance and low centrality. Similarly pho(t) refers to correlation between sharpe ratio and centrality so a trade off occurs to choice between high sharpe and high centrality.(in general sign refers to phase + indicates both should be decreasing or increasing,-ve indicates one decreasing-one increasing). Negative correlation with COV indicates variance in invested stocks should be as little as possible.

Table 4
Out-of-sample performance of portfolio strategies.
We report the out-of-sample Sharpe ratio, variance and turnover for portfolio strategies. The benchmark strategy is denoted as "*All stocks*", that is naïve strategies applied to all of the stocks in the dataset. Following the *ρ-dependent* strategy, when ρ is higher than 0.2, we diversify among highest central stocks and otherwise, diversify among the lowest central stocks. The *Reverse ρ-dependent* approach takes an opposite investment decision to the ρ-dependent strategy. The *Highest Sharpe Ratio* strategy refers to diversifying naively among stocks with the highest level of Sharpe ratios. In *Lowest Central* and *Highest Central* strategies, we diversifying among lowest and highest central stocks, respectively. The *minv-cc* and *mv-cc* strategies refer to minimum-variance and mean-variance strategies with short-selling constraints. We considered 20 stocks for our portfolios. The p-values are computed following the procedure in Ledoit and Wolf (2008) and Ledoit and Wolf (2011) based on a studentized circular block bootstrap with block size equal to 5 and number of bootstrap samples equal to 5000.

Panel A	d-S&P (Avg ρ: -0.0556)			d-FTSE (Avg ρ: -0.1853)		
	Sharpe Ratio	Variance	Turnover	Sharpe Ratio	Variance	Turnover
All Stocks	0.471	0.066	0.141	1.153	0.025	0.138
ρ-dependent	0.724	0.051	0.149	2.523	0.014	0.164
	(0.0125)	(0.0010)		(0.0033)	(0.0009)	
Reverse ρ-dependent	0.315	0.110	0.126	0.549	0.026	0.061
	(0.1523)	(0.0009)		(0.0100)	(0.014)	
Highest Sharpe Ratio	0.438	0.038	0.141	1.201	0.018	0.138
	(0.8239)	(0.0009)		(0.8007)	(0.0009)	
Highest Central	0.315	0.110	0.126	0.549	0.026	0.061
	(0.1523)	(0.0009)		(0.0100)	(0.014)	
Lowest Central	0.724	0.051	0.149	2.523	0.014	0.164
	(0.0125)	(0.0010)		(0.0033)	(0.0009)	
minv_cc	0.448	0.040	0.115	1.589	0.015	0.127
	(0.9568)	(0.0020)		(0.5482)	(0.0010)	
mv-cc	0.573	0.067	0.148	1.089	0.025	0.137
	(0.8339)	(0.8951)		(0.9535)	(0.6563)	
Panel B	m-NYSE (Avg ρ: 0.1157)			d-NYSE150 (Avg ρ: 0.2323)		
	Sharpe Ratio	Variance	Turnover	Sharpe Ratio	Variance	Turnover
All Stocks	0.754	0.019	0.057	1.241	0.021	0.127
ρ-dependent	0.591	0.025	0.065	2.083	0.019	0.140
	(0.3997)	(0.2553)		(0.0664)	(0.3367)	
Reverse ρ-dependent	0.552	0.034	0.049	0.968	0.028	0.113
	(0.0233)	(0.0009)		(0.5216)	(0.0020)	
Highest Sharpe Ratio	0.531	0.017	0.039	1.0571	0.018	0.103
	(0.1694)	(0.2248)		(0.5781)	(0.0110)	
Highest Central	0.555	0.032	0.049	1.201	0.033	0.106
	(0.0299)	(0.0009)		(0.8671)	(0.0009)	
Lowest Central	0.583	0.028	0.065	1.942	0.014	0.147
	(0.0831)	(0.0009)		(0.1927)	(0.0009)	
minv_cc	0.777	0.015	0.046	1.310	0.013	0.109
	(0.9302)	(0.1558)		(0.9734)	(0.0110)	
mv-cc	0.456	0.054	0.053	1.353	0.039	0.122
	(0.3422)	(0.0010)		(0.9502)	(0.0010)	

NYSE150 dataset. Nevertheless, in the cases where ρ is relatively low, as in the d-S&P and d-FTSE datasets, our results are in line

Table 4- This table shows the values of sharpe ratio,variance,turnover when different strategies are applied on the different datasets. In all these datasets we can see that the values of sharpe ratio are better and variance is less for our strategy(p-dependent) compared to all other strategies. This shows that the factors we are taking are correct in predicting portfolios. The turnover(cost incurred in portfolio readjustment) is a bit on the higher side but when we trade on longer durations(ie.20 days) the transaction costs are negligible compared to the returns

Table 5
Annualized risk-adjusted returns for ρ-dependent strategy with 5% winsorisation.
We report annualised risk-adjusted returns for different settings of ρ-dependent strategy on the four Carhart (1997) factors, MKT, HML, SMB, MOM with 5% winsorisation. The estimation window is considered to be 1000 days (192 months) for daily (monthly) datasets. The t-statistics are reported in parentheses.

Portfolio setting		Alpha (5% winsorizing)		
Stocks	Holding period	d-S&P	d-FTSE	m-NYSE
10	1d/1m	14.75	35.97	5.1
		(3.37)**	(6.37)**	(2.26)*
	20d/12m	17.19	36.01	4.49
		(3.97)**	(6.21)**	(1.98)*
20	1d/1m	10.57	33.65	4.42
		(3.02)**	(6.65)**	(2.31)*
	20d/12m	11.97	32.84	4.37
		(3.45)**	(6.52)**	(2.36)*
50	1d/1m	10.52	28.07	5.45
		(3.63)**	(5.17)**	(3.23)**
	20d/12m	10.08	28.08	5.82
		(3.50)**	(5.18)**	(3.40)**

** Indicate significance at the 1% level.

Table 5- The given table takes into account the transaction cost and has calculated risk adjusted returns taking that into account. The data in the table shows that the returns increase if we increase the holding period since transaction costs are offset while trading for longer durations.

Section-Wise Understanding

Section 1- Here the author delves into the intersection of traditional portfolio theory and financial network analysis. It uncovers that the centrality of stocks in financial networks impacts portfolio selection, with highly central stocks increasing portfolio risk. The author proposes a "ρ-dependent strategy" that considers the correlation between systemic and individual stock characteristics in the network.Systemic characteristics refers to a security's impact on the financial network whereas individual characteristic refers to the securities own's performance considering sharpe ratio,variance etc. This strategy outperforms traditional portfolio methods in out-of-sample testing. The study not only connects portfolio theory with network analysis but also provides a practical approach to enhance portfolio selection by factoring in the network structure of financial markets, improving wealth allocation and risk management.

Section 2-This section goes further and helps develop a relation between the weights ,returns, eigenvalues etc.It also shows how these factors are derived from Markowitz portfolio theory.It helps us to give complete detail on how network theory works and the dependent parameters.

Section 3-It indicates the datasets which are used for analysis ie. d S&P-the daily returns for 200 highly capitalized constituents of S&P index.m-NYSE- monthly returns of the 200 firms listed in NYSE.d-FTSE-daily stock returns of 200 most capitalized constituents of FTSE-200 index.

Section 4-The study examines centrality within the financial network, focusing on the d-S&P dataset. It finds that financial firms are the most central within the stock market network, despite manufacturing being the largest sector by market capitalization and traded volume. Moreover, central stocks tend to exhibit lower risk-adjusted returns and higher systematic risk. The study employs an econometric approach, identifying key drivers of centrality, such as firm capitalization, lower stock prices, older companies, and reduced cash holdings. Additionally, stocks' stability in centrality status over time is explored, with a strong tendency for consistent centrality levels observed among listed stocks.

Section 5-This section employs a cross-sectional and time-series approach to analyze stock market dynamics, focusing on centrality and portfolio optimization. Using the Minimum Spanning Trees (MST) technique, the study constructs financial market networks that filter excessive information from stock co-movements. It reveals that centrality plays a significant role in shaping optimal portfolio allocations, with a preference for low-central securities. Furthermore, a time-series perspective shows that centrality is influenced by a dynamic interplay between the individual and systemic dimensions of stocks. A key finding is the trade-off between stock centrality, standard deviation, and Sharpe ratio, with portfolio weights adapting to balance these factors. Regression analyses support these observations, emphasizing the importance of centrality distribution, systemic risk, and trade-off conditions in portfolio optimization. Overall, the study sheds light on the complex relationship between centrality and portfolio composition in stock markets.

Section 6-The study by DeMiguel et al. (2009) challenges conventional wisdom in portfolio management. They investigate the performance of a straightforward investment strategy called 1/N, where you distribute your wealth equally among all available assets. This seemingly unsophisticated strategy consistently holds its ground against, and sometimes even surpasses, more complex methods based on the work of Harry Markowitz, who introduced Modern Portfolio Theory. Markowitz's theories attempt to manage the estimation errors involved in building an investment portfolio.DeMiguel's research findings have been corroborated by other academics, including Jobson and Korkie (1980), Michaud (2008), and Duchin and Levy (2009). Collectively, these studies suggest that the 1/N strategy is a surprisingly effective way to manage investments, especially for those who prefer a simple and transparent approach. Furthermore, DeMiguel and his colleagues propose an innovative investment strategy known as the ρ-dependent strategy. This method utilizes network centrality rankings of assets. Centrality rankings are determined based on how interconnected or important an asset is within the financial market. The ρ-dependent strategy identifies groups of assets by their centrality rankings and then allocates wealth according to a parameter called ρ. For low ρ values, it invests more in low-centrality stocks, while for high ρ values, it favors high-centrality stocks.What makes this strategy intriguing is that it shows potential for outperforming the 1/N rule and other Markowitz-based strategies in various datasets. It delivers higher Sharpe ratios, which measure risk-adjusted returns, and reduces portfolio variance, which is an indicator of risk.

Even after accounting for transaction costs associated with buying and selling assets, the ρ-dependent strategy offers advantages, particularly when investments are held for longer periods.The only challenge with this strategy is determining the ideal value for the threshold parameter ρ, which requires further research. Nevertheless, the ρ-dependent strategy represents a compelling alternative to traditional portfolio diversification approaches, offering a fresh perspective on how to construct an investment portfolio that can deliver better risk-adjusted returns.

Conclusion

Network theory is a much better way of making an investment portfolio. Strategies based on combining network theory and individual stock performance are extremely efficient and give better returns as compared to traditional methods. The research paper also provides a ‘p’ based investment strategy which can be used to make portfolios and has been shown to have better performance. Further scopes of research include investigating networks which are directional(ie. correlation is different in different directions)