International Journal of Financial Management and Economics 2024; 7(1): 171-179

International Journal of Financial Management and Economics

P-ISSN: 2617-9210
E-ISSN: 2617-9229
IJFME 2024; 7(1): 171-179
www.theeconomicsjournal.com
Received: 27-01-2024
Accepted: 03-03-2024

Mehmet Benturk
Faculty of Economics, Administrative and Social Sciences, Istanbul Gelisim University, Istanbul, Turkey

Corresponding Author:
Mehmet Benturk
Faculty of Economics, Administrative and Social Sciences, Istanbul Gelisim University, Istanbul, Turkey

An alternative method to developing a stock market index: Machine learning implementation using higher moments and asset liquidity

Mehmet Benturk

DOI: https://doi.org/10.33545/26179210.2024.v7.i1.283

Abstract
Stock market indexes have a very significant role in asset pricing models, particularly in the Capital Asset Pricing Model (CAPM), as a proxy for the entire wealth in the economy. There is widespread agreement in academia and industry that popular indices are inadequate at reflecting the statistical properties of the market portfolio. This study differs from traditional approaches in two aspects: the index component selection and the components' weight assignment. K-Means machine learning technique is applied through stock moments to exclude outliers' impact on the index and liquidity for index component selection. Principal Component Analysis (PCA) was used to determine index component weights in addition to equal weighting and market-cap weighting to reduce value and growth stocks disproportionate influence on the index. Except for skewness, the PCA-based weighting index results are remarkably similar to the Center for Research in Security Prices (CRSP) market-cap weight index. The PCA-based weighted index has a significantly greater negative skew than other prominent indices.

Keywords: K-Means, PCA-based weighting for index component weighting, stock market index, market portfolio

Introduction
Stock market indices are a significant subject of debate in finance theory due to their critical role in the Capital Asset Pricing Model (CAPM). The CAPM, being an equilibrium market model, weights each security in the market portfolio based on its market capitalization. In practical terms, the market portfolio would contain all feasible assets that are currently available, such as real estate, precious metals, art collections, and any other assets, including the return of human capital. The key issue is that not all risky assets are investable and, in some cases, measurable. Stock market indices are nevertheless employed as a proxy in empirical research, despite their limitations. The pioneering study (known as Roll's critique) regarding CAPM's testability addressed two issues through the market portfolio. The first one is about the market portfolio's mean-variance efficiency, which is a mathematical requirement. The CAPM demonstrates mathematically that the market portfolio is mean-efficient. According to Roll, assuming the proxy is mean-variance efficient makes the CAPM tautological. The most commonly used proxies in theoretical models are the CRSP equal-weighted and CRSP value-weighted indexes. Tested the CRSP equal weighted index, and examined the CRSP value weighted index investigated the mean-variance efficiency of the market portfolio in the CAPM framework and concluded that they are not mean efficient. The second issue is the unobservable market portfolio. Studied measurement error on CAPM estimates (relationship between the asset return and the market return proxy) by using the correlation between the true market portfolio and the proxy. His findings rejected the CAPM's validity due to omitted components from the market analyzed the CAPM's validity by using artificial stock market data for different proxies. Their study confirmed that the abnormal returns were a result of the measurement error. However, in contrast to conclusion, the correlation between the true market portfolio and the proxy is insufficient information to conclude CAPM's estimates are biased.

Many studies have attempted to incorporate alternative proxies in addition to stock market indices to substitute for the omitted asset classes within market portfolio proxies. Tested this argument by adding corporate and government bonds, preferred stocks, real estate, and consumer durables index returns to U.S. stock index returns. In this parallel, tested CAPM at the international level by adding 12 major capital markets from Morgan Stanley Capital International (MSCI) Index to the CRSP and COMPUSTAT data sets to explain the value premium. Examined the CAPM for publicly traded investment funds by using six different proxies. In their model, incorporated human capital by assuming human capital return is linear function of the individual income growth rate. Even though these innovations improve index efficiency, they are still inadequate to reflect the real market portfolio.

This study primarily contributes to the existing literature by focusing on two key aspects to construct a market portfolio proxy. The first aspect involves utilizing clustering, specifically the K-Means algorithm, for the purpose of asset selection. This approach aims to minimize the negative impact of outliers on the proxy and allows for the creation of a suitable subset of stocks that accurately represents the market portfolio. Additionally, PCA is employed in a novel manner to effectively aggregate information from all components, thereby minimizing the loss of information.

Literature Review

The methodology used in this study comprises of two steps. In the initial phase, the K-Means algorithm is used to select assets. The objective of this phase is to reduce the impact of outliers by grouping stocks based on their liquidity and moments. In a similar manner, employed a comparable methodology to mitigate errors within his dataset in order to develop a house price index specific to Saudi Arabia.

The final stage involves assigning weights to the selected stocks. The conventional weighting methodologies have been subject to criticism due to their potential influence on stock categories. While equally weighting leads growth stocks to have an over influence on index performance, capital weighting results same problem for value companies. The 10 largest S&P 500 companies account for 25% of the overall market capitalization. PCA analysis is employed to mitigate these effects. The earliest similar research in the literature was conducted by for alternative stock market indexes. They evaluated their alternative index against the Dow Jones Industrial Average (DJI). Similarly, investigated the market value and volatility premium using a PCA-based replica portfolio. In the context of index component weighting, were the first to employ PCA in order to ascertain the constituent components of the proxy market portfolio. In a previous study conducted by, a similar methodology was implemented to apply PCA for the purpose of medical indexing.

Methodology

Asset Selection

The basic idea behind stock selection involved employing reverse engineering techniques to identify the optimal subset of stocks within a given stock universe. Although the precise market portfolio composition is unclear, the primary systematic risk factors that affect price have been extensively studied in academic literature. K-Means machine learning algorithm was used to cluster the stocks. K-Means is a well-known unsupervised learning algorithm for "clustering." The algorithm, as a vector quantization method, finds the best cluster for each observation that is closest to the mean (cluster centroid). The cluster that is least affected by outlier stocks will be the most suitable cluster for capturing the true market portfolio. K-Means algorithm is executed in two phases with the given features for each year in the data set. The algorithm determines the best clustering number in the first phase, and then labels the data in the second phase based on the cluster number found in the first phase. In the algorithm, the major systematic risk factors in the asset pricing framework are used as features. Hence, prior to running the algorithm, these systematic risk factors were identified.

Determining Features (Risk Factors)

In standard mean-variance analysis, risk is quantified by the standard deviation of asset returns. It is important to note that the standard deviation is an appropriate metric under the assumption that stock returns are normally distributed. Investigated stock return distributions and concluded that they are leptokurtic. His research confirmed the notion that the stock returns distribution belongs to the partisan distribution family, as proposed by. In contrast, empirical analysis revealed that monthly asset return distributions confirm normal distributions. Many pricing studies are conducted using monthly data based on that evidence. Similarly, used daily data and discovered results that supported normal distribution assumptions, but he also noted significant skewness and kurtosis in stock returns and indices. Empirically demonstrated that the daily returns of London Stock Exchange stocks fit Turkey's g- and h-distributions. Investigated time-varying price skewness in the US and global stock markets separately. According to their findings, time-varying price skewness is more important for the global market premium than for the US market premium, explaining some of the negative ex ante market risk premium. Similarly, found empirically that aggregate stock market returns have negative skewness using the CRSP data set from 1973 to 2009. He used a positively skewed firm announcement distribution correlation to explain asymmetric stock returns used moments from S&P 500 options data. Their research found that market skewness and kurtosis explained returns that were above average. Conducted an empirical research of ten developed and ten emerging market economic indices from 1979 to 2016 and found that the generalized lambda distribution is a major characteristic of alternative stock index models. In the CRSP data set, provided a theoretical model for daily returns that indicated systematic skewness pricing in cross-section stock returns. Similarly, there is a vast body of work on portfolio optimization with higher returns, such as, and Moreover, related the skewness and unconditional kurtosis of asset returns to liquidity spirals. In their analysis, the returns of speculators are negatively skewed. This results in a substantial increase in leverage restrictions when market liquidity conditions are poor. Higher moments of asset and market returns (skewness and kurtosis) contribute pricing for extreme events (black swan events) risks. Many studies, including, have shown that if the asset's extreme risk exceeds the market's extreme risk (systemic extreme risk) investors require an additional premium. As a result, in addition to the standard deviation, the skewness and ketosis of asset returns are included as features in the K-Means algorithm.

Asset liquidity is the last risk factor used in clustering. There is a substantial body of literature on the relationship between asset liquidity and returns. Some notable studies include those by Similar to our study, used liquidity as the main factor to construct the index. They used electronic limit order books to measure liquidity in their study.

Consequently, stock returns' standard deviation, skewness, kurtosis, and stock liquidity risk were determined as features in the K-Means machine learning algorithm. The K-Means algorithm generated clusters (stock batches) and labeled the stocks with corresponding batch numbers. Then, each batch's average standard deviation was computed. An index is often considered a neutral investment decision. Hence, the lowest average standard deviation batch, which is least influenced by outliers is selected for the proxy components.

Weighting

As previously stated, equal-weighted and market-cap-weighted indexing are the most commonly used methods in index construction. However, both approaches suffer criticism about their influences, such as equally weighted indices are impacted by low-cap stocks, whereas market-cap weighted indices are influenced by high-cap stocks.

In the literature, PCA has been employed in several studies as an alternative to the equally weighted and capital-weighted indexing against value and growth companies over influence on the index. In fact, PCA is one of the most widely used machine learning techniques for converting data into a lower-dimensional format while retaining as much information as possible. PCA determines principal components by using variables variance and their correlation among them. In more technical terms, each principal component is a linear combination of optimally weighted observed variables. The first principal component preserves the highest amount of variation that was initially present in the original components. The principal components are defined as the eigenvectors of a covariance matrix, resulting in their orthogonality. The proportion of variance accounted for by a principal component is proportional to its eigenvalue. The principal components are essentially the linear combinations of the original variables (stocks in our data) and the weights vector. Previous studies used PCA analysis, using either the first component directly as an index or the positive contribution weights of the first PCA component's elements as index weights. Both approaches result in the loss of information and used it in their studies for index or portfolio construction. This study, on the other hand, differs from them in two ways. To begin, rather than the PCA, another machine learning application, K-Means, is used to handle asset selection. The second distinction aims to solve the loss of information issue. Different than previous studies, the principal component analysis (PCA) components, along with their corresponding contribution weights, are used in an innovative way. First, PCA algorithm was coded to capture 99% of the variance. Hence, if there are n components, the sum of components' variance ratio (PCVR) equals 1.

$$\sum_{j=1}^{n} \mathrm{PCVR_j} = 1 \quad (1)$$

The first component (PC_1) explains the highest proportion of the variance, while the last component (PC_n) explains the least proportion of the variance.

$$PCVR_1 > \mathrm{PCVR}_2 > \cdots \mathrm{PCVR}_{n-1} > PC\mathrm{PCVR}_n \quad (2)$$

If we have p stocks $S_1, S_2, \ldots, S_p$ that results p loading factors ($\emptyset$) for each component. $\emptyset_m = (\emptyset_{1m}, \emptyset_{2m}, \emptyset_{3m}, \ldots, \emptyset_{pm})$ is the loading vector for the m the principal component.

$$\sum_{j=1}^{p} \emptyset_{jm}^2 = 1 \; and \; \emptyset_{jm} \in \mathbb{R} \quad (3)$$

Hence m th principal component can be defined as follows;

$$PC_m = \emptyset_{1m}S_1 + \emptyset_{2m}S_2 + \cdots + \emptyset_{pm}S_p \quad (4)$$

The basic idea is to use the principle component variance ratio (PCVR) and corresponding loading weights to find each feature's (stock) proportional contribution to the component. Hence, n the stocks (S_n) proportional contribution to m the principle component (PC_m) is as follows:

$$\tau_{n,m} = \frac{abs(\emptyset_{n,m})}{\sum_{j=1}^{p} abs(\emptyset_{j,m})} \quad (5)$$

Since there is n principle components, m the feature's (m the stock) weight will be the sum of the product for each component's variance ratio and corresponding proportional contribution to the component. This calculation is very similar to feature importance concept in some machine learning algorithms like decision tree.

$$w_m = \sum_{j=1}^{n} PCVR_j \, \tau_{j,m} \quad (6)$$

w_m is m the stock's weight, which is its proportional contribution to explain total variance.

$$w_1 + w_2 + \cdots + w_p = 1 \quad (7)$$

Data

Except for certificates, American Depositary Receipts (ADRs), Shares of Beneficial Interest (SBIs), Depository Units, exchange-traded funds (ETFs), Real Estate Investment Trusts (REITs), and shares of Berkshire Hathaway Inc., from January 1, 1990, to March 31, 2022, all stocks in the CRSP database were used. The liquidity proxy was measured using trade volume-averaged bid-ask spreads. On days with low trading activity, the bid-ask spreads for low-volume stocks tend to demonstrate a significant width. Hence, averaging bid-ask spreads by using the trading volume of the stocks increased the spreads accuracy. In addition to the returns, moments of returns, and liquidity, The Standard Industrial Classification Codes (SIC) for each stock were added to the data set to track the index's industrial composition over time.

Each year's expected returns and moments of stocks were calculated in a loop by using three years of historical data. The K-Means algorithm was looped for each year from 1995 to 2022. The calculations were coded in the Python environment by using K-Means and PCA algorithms from the Sklearn library. Missing data observations are imputed with the variable averages using the Sklearn library's simple imputer during the PCA process.

Results

Component Selection: As previously stated, the fundamental purpose is to develop a stock market index that meets theoretical requirements. A good index should reflect the sectoral distribution of the overall economy. Stocks in the data set were labeled according to their SIC in order to be analyzed indices sectoral composition. The sectoral distribution of index components over time is displayed in Figure 1. Manufacturers, financiers, and service companies have the largest share of the total. Their trend, however, began to decline gradually after the mid-2000s, whereas the public sector, which primarily represents technology companies, has grown at a rapid pace. The results demonstrate that the component selection well captured the sectoral distribution in the overall economy, as the economy's sectoral tendencies closely resemble the image.

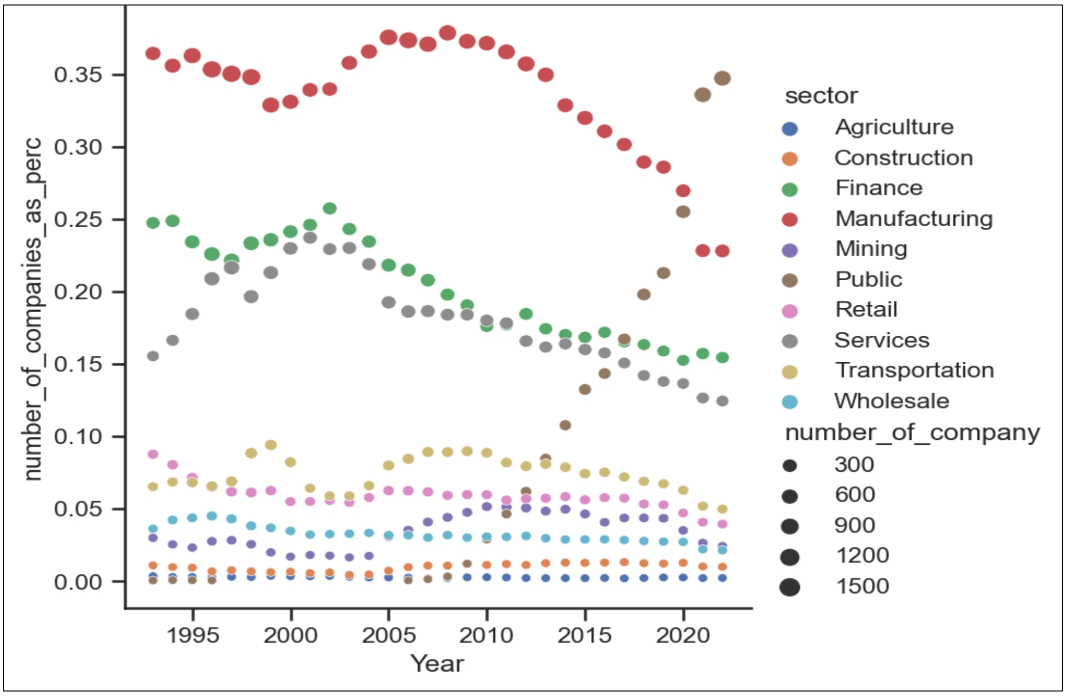


Fig 1: The sectoral compositions of index components

Weights

The weighting methodology directly influences the impact of value and growth companies on the index. Due to their typically longer existence, value companies are less common in both the economy and index compared to growth companies. Conversely, value companies have larger market capitalizations. As a result, value companies have an influence on market-cap weighting indices, whereas growth companies have an influence on equal-weighting indices. Therefore, the weighting methodology has a substantial effect on the performance and statistical properties of the index. PCA-based weighting was used to eliminate these impacts. The equal-weighting, market-cap averaged weighting, and PCA-based weighting indices were compared to the major popular indices, the DJI, S&P 500, Nasdaq Composite Index, and CRSP Value Weighted Index, to analyze performance and moments. Figure 2 shows the estimated kernel density plot for the equal-weighted, market-cap weighted, and PCA-based weighted indexes with the major popular indices returns. At first glance, it is evident that none of the indices exhibit a normal distribution and have similar means and standard deviations. Table 1 shows the Shapiro-Wilk test results, which also fail to support the presence of a Gaussian distribution. For the higher moments, however, the three indices that are generated in this study differ from other major indices.

Table 1: Indices returns normality test statistics

Index	Equal Weighted Index	Market Cap Index	PCA Weighted Index	DJI	S&P 500	NASDAQ	CRSP Market Cap Index
Shapiro-Wilk Statistic	0.912	0.907	0.915	0.896	0.901	0.927	0.907
p Value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 2: Indices' returns moments

Index	Equal Weighted Index	Market Cap Index	PCA Weighted Index	DJI	S&P 500	NASDAQ	CRSP Market Cap Index
Mean	0.0598	0.1019	0.0472	0.0394	0.0411	0.0551	0.0486
Standard Deviation	1.2701	1.4302	1.3698	1.1548	1.1991	1.5435	1.2043
Skew	-0.4970	0.1274	-0.4303	-0.1657	-0.1950	-0.0043	-0.2636
Kurt	9.2749	7.9792	8.4742	12.3280	10.4466	6.4752	9.5893

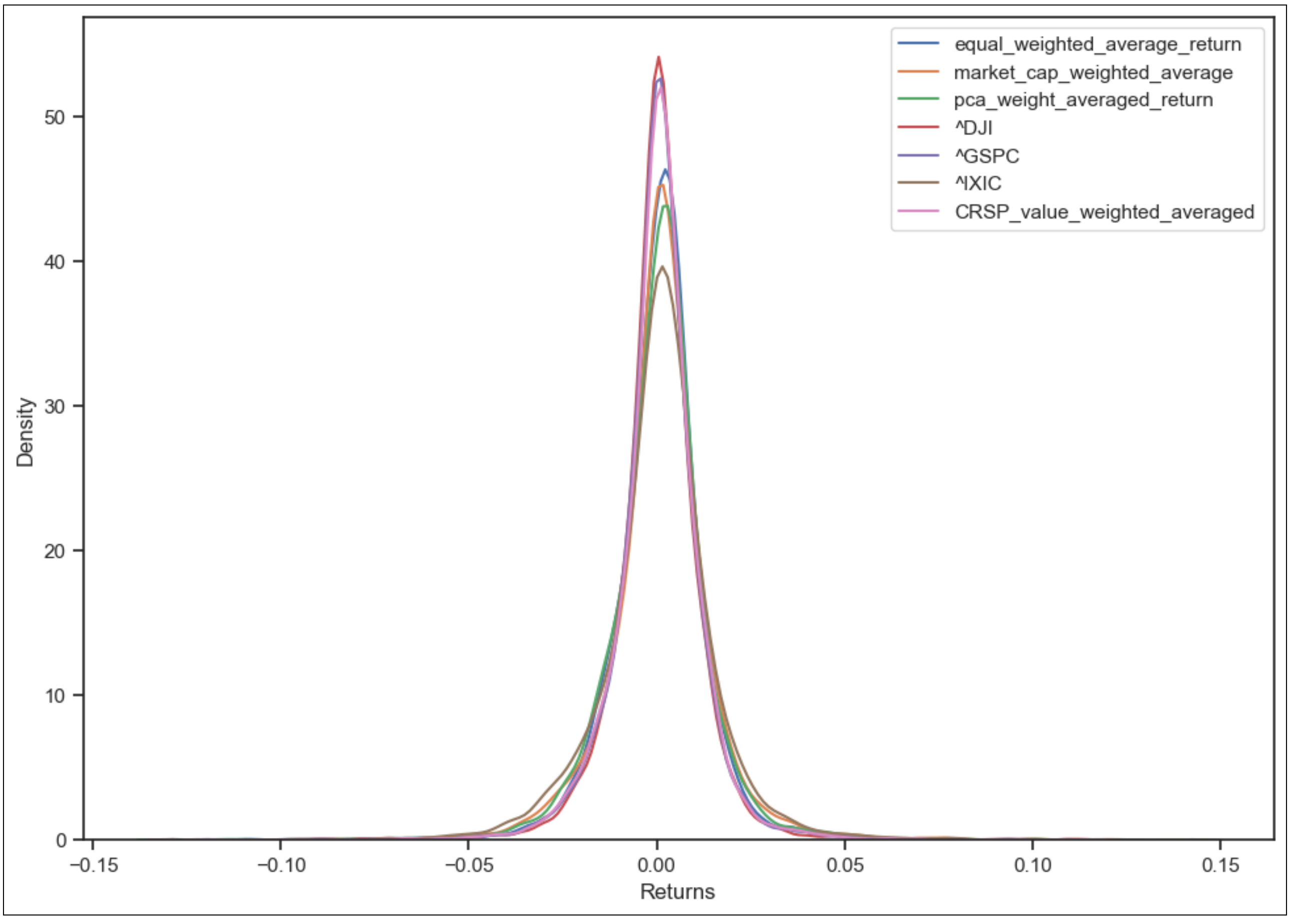
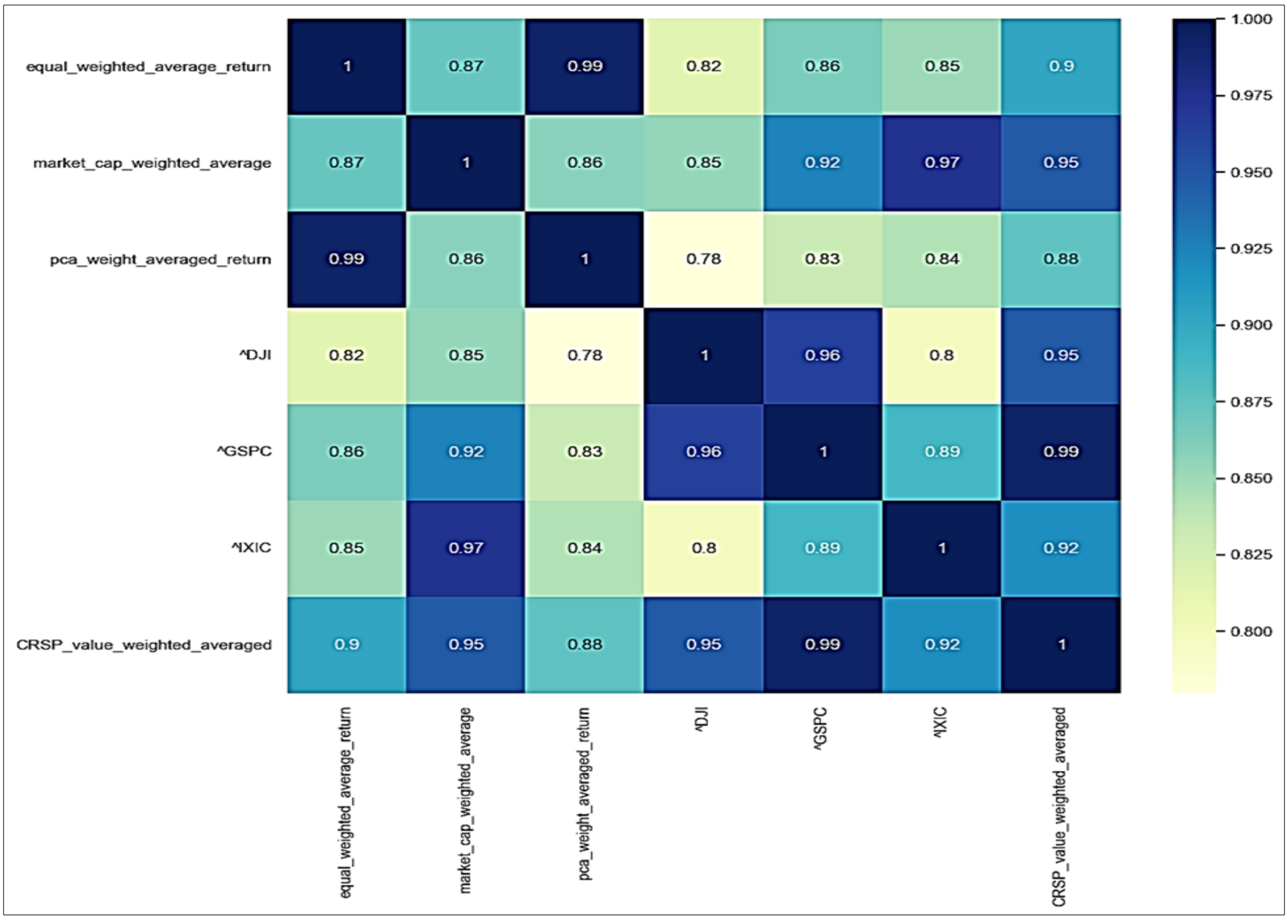


Fig 2: Indices' returns distribution

Table 3: Indices returns correlation matrix



	equal_weighted_average_return	market_cap_weighted_average	pca_weight_averaged_return	^DJI	^GSPC	^IXIC	CRSP_value_weighted_averaged
equal_weighted_average_return	1	0.87	0.99	0.82	0.86	0.85	0.9
market_cap_weighted_average	0.87	1	0.86	0.85	0.92	0.97	0.95
pca_weight_averaged_return	0.99	0.86	1	0.78	0.83	0.84	0.88
^DJI	0.82	0.85	0.78	1	0.96	0.8	0.95
^GSPC	0.86	0.92	0.83	0.96	1	0.89	0.99
^IXIC	0.85	0.97	0.84	0.8	0.89	1	0.92
CRSP_value_weighted_averaged	0.9	0.95	0.88	0.95	0.99	0.92	1

Table 2 shows the indices' return moments. The market-cap-weighted index differs from others in terms of mean and skewness. The mean of the market-cap weighted index is almost twice as high as that of the other indexes, and it is positively skew. As previously discussed, the market-cap weighting results in the over-influence of large-cap companies on the index. Figure 3 shows the cumulative return performance of the indices over time. In comparison to the others, the market-cap weighted index performance is excessively high. Therefore, it can be inferred that the use of market capitalization weighting results in the ineffectiveness of indexing as a proxy measure. In contrast to its return's moments, its correlations with major popular indices (particularly the NASDAQ composite index and the CRSP market-cap index) are extremely high. This brings us back to the misspecification studies mentioned in the introduction. According to ,, and , the theoretical requirement for market portfolio mean-variance efficiency can be rejected or accepted through a proxy as long as its correlation with the real market portfolio is high enough (greater than 0.70). However, as the market-cap weighted index indicates, a proxy that is highly correlated with the real market portfolio can diverge in moments.

The equal and PCA-weighted averaged indices are both highly correlated with the most popular indices. Table 3 shows the correlation matrix for indices returns. While the correlation between the equal-weighted average index and the CRSP Market-Cap weighted index is 0.90, the correlation between the PCA-based weighted index and the CRSP Market-Cap weighted index is 0.88. Aside from the skewness, the moments of the equal-weighted average and PCA-based weighted indices are close to the moments of the major popular indices, as shown in Table 2. The equal-weighted average index return's mean is 0.0598, very close to the NASDAQ, which is 0.0551. Similarly, the means of the PCA-based weighted and CRSP market-cap indices returns are very close to each other, at 0.0472 and 0.0486, respectively. For the returns' second and fourth moments (variance and kurtosis), both the equal-weighted and PCA-based weighted indices are close to the major popular indices. They do, however, differ significantly from the major popular indices in terms of skewness. The equal-weighted and PCA-based weighted average indices exhibit skew nesses of -0.4970 and -0.4303, respectively. Hence, the CRSP Market-Cap weighted index is the closest major popular index to the PCA-based weighted index in terms of skewness. The skewness of the NASDAQ index return is -0.0043, which is the most pronounced deviation compared to both the equal-weighted and PCA-based weighted indexes. As a result, the PCA-based weighted index technique deviates from the major indices due to its skewness. Greater negative skewness indicates greater black swan probabilities in the market, which helps explain stock market crashes such as the 1929 crisis, 1989 Black Monday, and 2000 Dot-com crash.

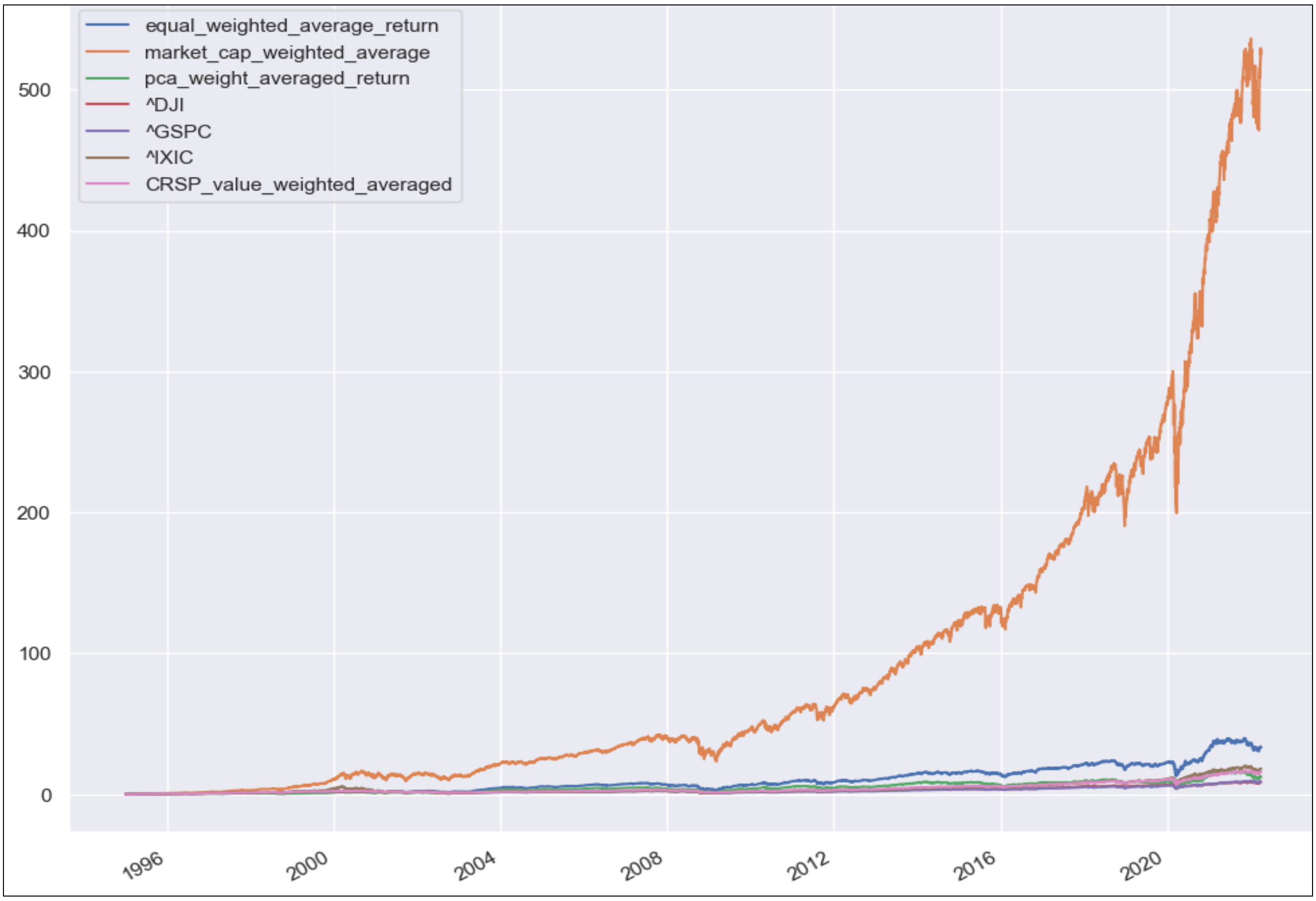


Fig 3: Indices' Compound Returns (03/01/1996=100)

Fig 4: Indices' Compound Returns (Market-Cap Index is excluded), (03/01/1996=100)

Despite having the same components, market-cap-weighted, equal-weighted, and PCA-based weighted indexes perform differently due to weighting. Figure 4 depicts the cumulative return of all indices, while the market capitalization-based index was removed due to its scale difference. Hence, Figure 4 provides a clear comparison of the performance of indices. The equal-weighted index has a relatively more reasonable performance than the market-cap weighted index. There is, however, a significant performance difference when compared to the major popular indices. In contrast to market-cap indices, equal-weighted indices are influenced by small-cap companies, as is the case in this study. Demonstrated in their experimental study that if equal-weighted indices are very inclusive (high component number), as this study's methodology is, they become less representative for the market portfolio. PCA-based weighting, on the other hand, is based on the idea of capturing the greatest amount of variance, which eliminates large or small cap companies' undue influence on the index.

As seen from Figure 4, PCA-based weighting index has similar performance to the CRSP market-cap weighted index. Furthermore, the correlation matrix among the indices that is presented in Table 3 also confirms this inference. PCA-based weighting works perfectly in index weighting to reflect the pure price change relationship among the components by avoiding the capitalization impact. The main difference between these two indices, as shown in Table 2, is skewness. The PCA-based weighted index is noticeably more skewed than the CRSP weighted index (and other major indices). The index's higher negative skewness makes the black swan events we've seen in the markets more understandable. More importantly, it explains why investors accept lower returns on value stocks versus growth stocks. The skewness of the DJI is -0.1657, which is 2.6 times greater than the skewness of the PCA-based weighted index. This conclusion also validates the endogenous implication in the CAPM framework, which is addressed.

Conclusions

The K-Means algorithm classification based on higher moments and asset liquidity provided very consistent stock (component) selection for the market portfolio proxy. The weighting methods made a significant difference in index performance and proxy statistical properties. As expected, the market-cap-based weighting index methodology is overly influenced by the value companies' stocks. The market-capitalization-based weighting index methodology distinguished itself in terms of index performance and statistical properties when compared to other methods and major indices. The performance and statistical features of equal-weighted indexing are more reasonable compared to market-cap weighted indexing. However, due to the nature of equal-weighting, small-cap stocks have a disproportionate influence on the proxy, resulting in a questionable performance difference when compared to the major popular indices.

Alternatively, PCA analysis was used to eliminate market-cap effects on the proxy during the weighting step. As anticipated, the PCA analysis weighting methodology generated satisfactory results. It has a high correlation with the major popular indices. Likewise, except for its third moment (skewness), its other moments (mean, variance, and kurtosis) are very close to major popular indices. The PCA-based weighting index, on the other hand, stands out due to its extremely high negative skewness.

In brief, the PCA-based weighting index will be extremely beneficial in improving pricing model qualities in academic and industrial practices. Considering the size of the index fund industry, this study's methodology will be beneficial to provide objective and transparency-based indices to industry.

References

1. Acharya VV, Pedersen LH. Asset pricing with liquidity risk. Journal of Financial Economics. 2005;75:375-410.
2. Albuquerque R. Skewness in Stock Returns: Reconciling the evidence on firm versus aggregate returns. Review of Financial Studies. 2012;25(6):1630-1673.
3. Alexander C, Dimitriu A. Sources of over-performance in equity markets: Mean reversion, common trends and herding. University of Reading, Discussion Papers in Finance; c2003.
4. Algahtani SN. Constructing a house price index for Saudi Arabia. Journal of Real Estate Portfolio Management. 2022;28(2):166-182.
5. Amenc N, Goltz F, Sourd VL. Assessing the quality of stock market indices: Requirements for asset allocation and performance measurement. EDHEC Risk & Asset Management Research Centre Publication; c2006.
6. Amihud Y, Mendelson H. Asset pricing and the bid-ask spread. Journal of Financial Economics. 1986;17:223-249.
7. Avellaneda M, Lee JH. Statistical arbitrage in the US equities market. Quantitative Finance. 2010;10(7):761-782.
8. Bali TG, Murray S. Does risk-neutral skewness predict the cross-section of equity option portfolio returns? Journal of Financial and Quantitative Analysis. 2013;48(4):1145-1171.
9. Bernstein PL. Points of Inflection: Investment Management Tomorrow. Financial Analysts Journal. 2003;59(2):18-23.
10. Blattberg RC, Gonedes NC. A comparison of the stable and student distributions as statistical models for stock prices. Journal of Business. 1974;47(2):244-280.
11. Brennana MJ, Subrahmanyam A. Market micro-structure and asset pricing: On the compensation for illiquidity in stock returns. Journal of Financial Economics. 1996;41:441-464.
12. Brown KC, Brown GD. Does the composition of the market portfolio really matter? The Journal of Portfolio Management. 1987;13(2):26-32.
13. Brunnermeier MK, Pedersen LH. Market liquidity and funding liquidity. The Review of Financial Studies. 2009;22(6):2201-2238.
14. Chang BY, Jacobs K, Christoffersen P. Market skewness risk and the cross-section of stock returns. Journal of Financial Economics. 2013;110(1):46-68.
15. Chao YS, Wu CJ. Principal component-based weighted indices and a framework to evaluate indices: Results from the Medical Expenditure Panel Survey 1996 to 2011. PLOS One. 2017;12(10):e0186826.
16. Chordia T, Roll R, Subrahmanyam A. Commonality in liquidity. Journal of Financial Economics. 2000;56(1):3-28.
17. Corlu CG, Meterelliyoz M, Tiniç M. Empirical distributions of daily equity index returns: A comparison. Expert Systems with Applications. 2016;55:170-172.
18. Daniel SR. Conditional skewness and asset pricing. Journal of Empirical Finance. 2007;14(1):91-119.
19. Dittmar RF. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. Journal of Finance. 2002;57(2):369-403.
20. Fama EF. The behavior of stock-market prices. The Journal of Business. 1965;38(1):34-105.
21. Fama EF, French KR. Value Versus Growth: The International Evidence. Journal of Finance. 1989;44(5):1975-1999.
22. Feeney GJ, Hester DD. Stock Market Indices: A Principal Components Analysis. Cowles Foundation Discussion Paper No. 175; c1964.
23. Gibbons MR. Multivariate tests of financial models a new approach. Journal of Financial Economics. 1982;10(1):3-27.
24. Harvey CR, Siddique A. Time-varying conditional skewness and market risk premium. Research in Banking and Finance. 2000;3:25-58.
25. Harvey CR, Liechty JC, Liechty MW, Muller P. Portfolio selection with higher moments. Quantitative Finance. 2010;10(4):469-485.
26. Hasbrouck J, Seppi DJ. Common factors in prices, order flows, and liquidity. Journal of Financial Economics. 2001;59(3):383-411.
27. Haugen RA, Baker NL. The efficient market inefficiency of capitalization-weighted stock portfolios. The Journal of Portfolio Management. 1991;17(2):35-40.
28. Huang M. Liquidity shocks and equilibrium liquidity premia. Journal of Economic Theory. 2003;109(1):104-129.
29. Jagannathan R, Wang Z. The Conditional CAPM and the Cross-Section of Expected Returns. Journal of Finance. 1996;51(1):3-53.
30. Jeewon J, Jang J, Kang J. An Intertemporal CAPM with higher-order moments. The North American Journal of Economics and Finance. 2017;39:314-337.
31. Jobson JD, Korkie B. Potential performance and tests of portfolio efficiency. Journal of Financial Economics. 1982;10(3):433-466.
32. Jurczenko EF, Maillet BB, Merlin PM. Efficient frontier for robust higher-order moment portfolio selection. CES Working Paper; c2008.
33. Kandel S, Stambaugh RF. On correlation and inferences about mean-variance efficiency. Journal of Financial Economics. 1987;18(1):61-90.
34. Kleniati MP, Rustem B. Portfolio Decisions with Higher Order Moments. COMISEF working Papers Series; c2009.
35. Kon SJ. Models of Stock Returns - A Comparison. Journal of Finance. 1984;39(1):147-165.
36. Lo AW, Mamaysky H, Wang J. Asset prices and trading volume under fixed transactions costs. Journal of Political Economy. 2004;112(5):1054-1090.
37. Longstaff FA. Optimal Portfolio choice and the valuation of illiquid securities. The Review of Financial Studies. 2001;14(2):407-431.
38. Mandelbrot B. The variation of some speculative prices. The Journal of Business. 1967;40(4):393-413.
39. Mills TC. Modelling skewness and kurtosis in the

London stock exchange FT-SE Index Return Distributions. The Statistician. 1995;44(3):323-332.
40. Ohara M. Presidential Address: Liquidity and Price Discovery. The Journal of Finance. 2003;58(4):1335-1354.
41. Pastor L, Stambaugh RF. Liquidity risk and expected stock returns. Journal of Political Economy. 2003;111(3):642-685.
42. Pedersen LH, Gârleanu NB. Margin-based asset pricing and deviations from the law of one price. The Review of Financial Studies. 2011;24(7):1980-2022.
43. Prono T. Market proxies as factors in linear asset pricing models: Still living with the roll critique. Journal of Empirical Finance. 2015;32:36-53.
44. Renault E, Chabi-Yo F, Leisen DP. Aggregation of preferences for skewed asset returns. Journal of Economic Theory. 2014;152:453-489.
45. Roll R. A critique of the asset pricing theories tests Part I: On past and potential testability of the theory. Journal of Financial Economics. 1977;4(2):129-176.
46. Shah A, Thomas S. Market Microstructure Considerations in Index Construction. ICBOT Research Symposium Proceedings. 1998;27:173-193.
47. Shanken J. Multivariate Proxies and Asset Pricing Relations: Living with the Roll Critique. Journal of Financial Economics. 1987;18(1):91-110.
48. Simmet A, Pohlmeier W. The CAPM with Measurement Error: There's life in the old dog yet!. Journal of Economics and Statistics. 2020;240(3):417-453.
49. Stambaugh RF. On the exclusion of assets from tests of the two-parameter model a sensitivity analysis. Journal of Financial Economics. 1982;10(3):237-268.
50. Strongin S, Petsch M, Sharen G. Beating Benchmarks. Journal of Portfolio Management. 2000;27(2):11-27.
51. Thomas S, Clare A, Motson N. An Evaluation of alternative equity indices, Part 1: Heuristic and Optimised Weighting Schemes. Econometrics: Econometric & Statistical Methods. 2013;2(1).