Covariance Structure and Systematic Risk of Market Index Portfolio

Lukáš Pichl

International Christian University
Osawa 3-10-2, Mitaka, Tokyo 181-8585, Japan
lukas@icu.ac.jp

Abstract. A set of 18 stocks, selected as the current components of the Dow Jones Index, for which the historical daily closing data quoted at the US market are available for over four decades, is studied. Within this portfolio, we construct a market index with static weights, defined as the relative aggregate trading amounts for each stock. This market portfolio is studied by means of correlation and covariance analysis for the times series of logarithmic returns. Although no measure defined at the correlation/covariance matrices could be found as a definite precursor of market crashes and bubbles, which thus appear as a rather sudden phenomenon, there is an increase in the covariance measures for large absolute values of the logarithmic return of the index. This effect is stronger for the negative values of the log return, corresponding to the market crash case, during which the first principal component of the covariance matrix tends to describe larger proportion of the total market volatility. Periods of low volatility in the market can be characterized by rather significant spread of the relative importance of the first principal component. This finding is common also for the case of dynamically constructed market index, for which the weights are computed as the coordinates of the first principal component eigenvector using short-term covariance matrices.

1 Introduction

Diversification of assets over sectors with negative correlations represents a way of portfolio risk reduction. Taking a financial market as an example, there is, however, a limit that can not be overcome, the systematic risk of the market itself. This is important at the times of turbulence, when the structure of the temporal correlation matrix changes, and all assets exhibit a uniform, upward or downward trend. The latter case is more common, in which the panic drives the entire market to the crash, and the mere fact of exposure to the market, regardless of the particular portfolio composition, ultimately results in large losses. It is an interesting question whether these situations can be predicted based on the market indicators alone, such as the set of all stock prices and the trading volumes. For that purpose, we have selected a set of 18 stocks from the Dow Jones Index and followed them back to the past for the period of

W. Chu et al. (Eds.): DNIS 2015, LNCS 8999, pp. 172–179, 2015.

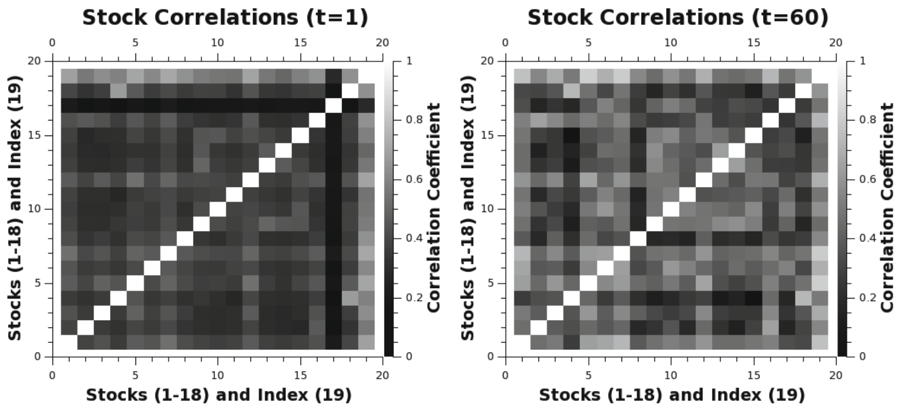
four decades. Within this limited portfolio universe, we study the dynamics of correlations among the stocks, focusing especially on the covariance matrix. These data are certainly different from the Dow Jones Industrial Average Index (DJIA), since the particular stocks were selected so that no index reconstitution takes place. This procedure can be understood as a sort of specification of static restrains for investment portfolio. In effect, we obtain four decades long time series starting from 1974, which include not only the daily closing prices, but also the daily trading volumes in the units of shares, all available at the site of Yahoo Finance. The first question we followed was whether there exists any precursor to the times of turbulence based on the values of the short-term correlation and covariance matrices. Using scalar products of vectors of logarithmic returns for individual stocks, cosine similarities among returns of shares and indices, Frobenius norm of the covariance matrix, sum of the off-diagonal correlation matrix coefficients, etc., resulted in no significant precursor that could be used to predict large systematic risk, characterized by large absolute values of the logarithmic return of the market index. In other words, the periods of market turbulence and the correlation collapse appear practically simultaneously, as far as the time scale of daily price changes is concerned for the present dataset. We therefore focus on the statistics of covariance matrix features, in relation to the distribution of the logarithmic returns of the market index.

Financial risk and turbulence effects in complex systems [1, 2] have been thoroughly studied both in the fields of econometrics and econophysics. Cross-correlations among assets are subject to time dynamics and exhibit a variety of stylized facts [3]. Among the number of econometric approaches developed for the study of multivariate time series [4], cross-correlations and their dynamics constitute the first line of research [5–8]. For instance, Preiss et al. found an intriguing scaling law for the relation of the sum of the off-diagonal correlation matrix elements and the normalized returns of DJIA on the time scales between 10 and 60 trading days. Persistence of the collective trends in stock market is studied in [9], and the general framework of cross-correlations is complemented by a number of recent studies applied to the particular markets [10]. Although we present the matrix of correlation coefficients for the present dataset in Fig. 1, most of the present study is based on the covariance matrix for the logarithmic returns of individual stock components. Because of the use of the logarithmic returns, this matrix is also dimensionless, albeit not normalized within the range of -1 and 1. The Principal Component Analysis (PCA [11]) is applied to the covariance matrix. This is because the variable with the highest variance in the logarithmic return will dominate the first principal component, which we prefer to the standardization of pairwise correlations. In particular, we study the share of the total variance explained by the first eigenvalue to the total variance within the dataset, thus standardizing our analysis within the range of 0 and 1.

The rest of the paper is organized as follows. In Section 2, the basic variables are defined and the correlation properties of the present dataset pictured. In Section 3, we investigate the properties of the covariance matrix. This is done using first the sum of the off-diagonal covariance elements of the logarithmic returns for individual stock pairs, and then by using the share of the total covariance matrix explained by the first eigenvalue (PCA analysis). The implications of the results on the inner market portfolio dimension are discussed in Sec. 4, and we conclude with summarizing remarks in Sec. 5.

2 Construction of Market Index

The 18 components of the portfolio universe are listed in Table 1 by their commonly used trading name abbreviation. Figure 1 shows the correlation coefficients for the 18 stocks listed in Table 1. This is the standard Pearson's correlation coefficient for the time series of logarithmic returns,



(a) Correlation coefficient on daily scale (b) Correlation matrix for 60 trading days

Fig. 1. Correlation matrix image for selected portfolio of 18 stocks listed in DJIA. Portfolio index is constructed and labeled as the component No. 19. The correlation coefficients may vary with the time scale step size and the length of the time series (10,290 days between 1974-01-02 and 2014-10-13).

Table 1. Stock specification by trading name abbreviation

AXP (1), BA, CAT, CVX, DD, DIS, GE, IBM, JNJ, KO, MCD, MMM, MRK, PFE, PG, UTX, WMT, XOM (18)

$$R_i(t) = \log \frac{P_i(t+1)}{P_i(t)}, \quad i = 1, \ldots, 18, \quad t = 1, \ldots T. \tag{1}$$

The label 19 in Fig. 1 is reserved for the market portfolio index, weights of which are defined by the individual share trading amounts $M_i(t)$ (in USD),

$$w_i = \frac{\sum_{t=1}^{T} M_i(t)}{\sum_{i=1}^{18} \sum_{t=1}^{T} M_i(t)}, \quad T = 10,290. \tag{2}$$

The weights are normalized and all positive. Since the matrices of the correlation coefficients include only positive values, an implicit market index derived from the first principal component of the correlation matrix could be used instead. It is known that for the positive correlation matrix, all coordinates of the first principal component vector can be selected with the positive sign. In practice, the role of either index is sufficiently representative, and we thus work with the one given in Eq. (3). The weights are summarized in Table 2. Figure 2

Table 2. Index weights for 18 stocks in Table 1

w[01]=0.034914,	w[02]=0.026008,	w[03]=0.032335,	w[04]=0.041831,
w[05]=0.022976,	w[06]=0.046245,	w[07]=0.129413,	w[08]=0.096037,
w[09]=0.059690,	w[10]=0.076222,	w[11]=0.035123,	w[12]=0.027158,
w[13]=0.061539,	w[14]=0.085853,	w[15]=0.056945,	w[16]=0.028327,
w[17]=0.048806,	w[18]=0.090579		

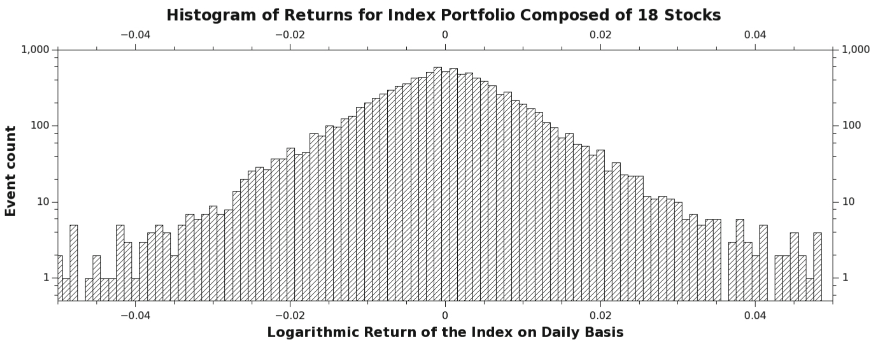


Fig. 2. Statistical distribution of logarithmic returns for the portfolio index constructed from 18 stock components. The weights are static, set as the relative trading amount for each component throughout the entire 40-year long trading period. Logarithmic scale is applied for the relative frequencies to illustrate the normality hypothesis for low values of R; fat tails of the distribution can be observed at both ends of the graph, which are known to exhibit slow, power-law decay.

shows the histogram of the dimensionless logarithmic returns for the static index constructed as described above. The index is determined as the weighted sum,

$$I(t) = \sum_{i=1}^{N} w_i P_i(t), \quad N = 18, \quad t = 1, \dots T. \tag{3}$$

The distribution of trading volumes in currency units for the market is shown in Fig. 3. Since the source data from Yahoo Finance report stock volumes as the number of shares traded, this value is multiplied by the average of the opening and closing prices for each stock, which is accurate enough for the determination of relative trading amounts that serve as the index weights.

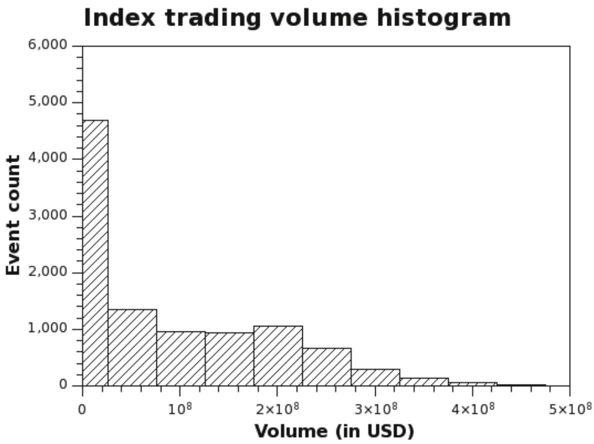


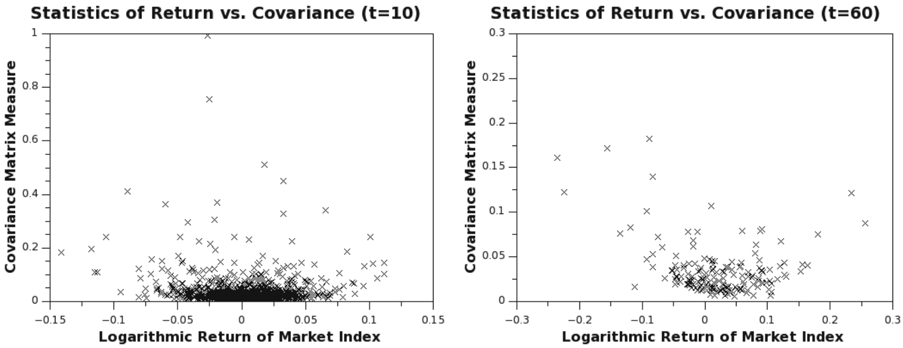
Fig. 3. Histogram of index portfolio trades. Although lower volumes are more frequent, there is a systematic bias towards higher amounts as the market prices inflate over the 40-year period.

3 Covariance Matrix Structure

The sample covariance matrix element for two time series of log returns $R_i(t)$ and $R_j(t)$ is given as

$$C_{ij} = \frac{1}{T-1} \sum_{t=1}^{T} \left(R_i(t) - \bar{R}_i\right) \left(R_j(t) - \bar{R}_j\right), \tag{4}$$

where $\bar{R}_i$, $\bar{R}_j$ are the sample means computed for $t = 1, \dots, T$. We define two covariance measures, $C_{off} = \sum_{i>j} C_{ij}/(N(N-1))$, and C, which is the plain sum of all elements of the covariance matrix. Figure 4 shows the relation of



(a) Time step of 10 trading days. (b) Time step of 60 trading days.

Fig. 4. Off-diagonal aggregate covariance for the logarithmic returns of all 18 index components. The relation to the market index volatility (as measured by the aggregate daily logarithmic return of the index) is shown for two time scales. For $\Delta t = 60$, larger values of logarithmic return of the index correlate to the larger covariance measure (sum of the off-diagonal elements of the covariance matrix).

the log returns of the market index to the covariance measure C_{off}. Two time scales are considered, 10 days and 60 days. In the non-overlapping intervals $I_j = \langle j\Delta T, (j+1)\Delta T - 1\rangle$, the covariance matrices are computed, C_{off} determined, and compared to the market index return, $R_j = \log[I((j+1)\Delta T - 1)/I(j\Delta T)]$, which is the same as the sum of the daily index returns $\log[I(t+1)/I(t)]$ throughout this time interval. The off-diagonal part of the covariance, for low values of R, clusters at the bottom of the figure, although highly correlated events can also be seen at the upper part. These correspond to large volatility events, when the stocks move in unison, but the rises and falls of the index compensate each other. At larger values of R, nevertheless, the covariance measure tends to increase, which is common to both time scales indicated in Fig. 4. We do not observe, nevertheless, any stylized fact that would reduce the cluster onto a single line. The results are similar if the second covariance measure, C, is employed (not shown in the Figure).

4 Market Portfolio Dimension

In order to examine the structure of the covariance matrix more thoroughly, we apply the Principal Component Analysis, PCA. Since C_{ij} is a positive semidefinite matrix, the total volatility of the dataset can be decomposed using the positive eigenvalues λ_i and the orthogonal eigenvectors $\boldsymbol{v}^{(i)}$ of the covariance matrix. Selecting the first principal component, we study the share of the dataset variance it comprises, i.e. the relative measure

$$\rho = \frac{\lambda_1}{\sum_{i=1}^{N} \lambda_i}, \quad \lambda_1 \geq \lambda_2 \geq \ldots, \lambda_N, \quad (N = 18). \tag{5}$$

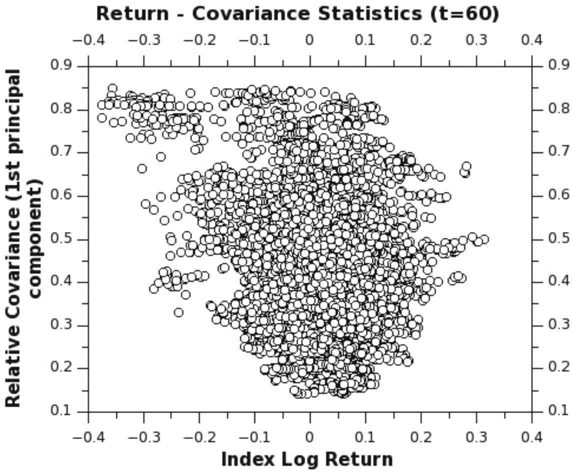


Fig. 5. Statistical distribution of index returns and implied market dimension as measured by the share of the first principal component in the covariance matrix on the total variance of the entire data set. Notice that the quite uniform population for low values of the logarithmic return changes as the return values grow. For larger absolute values of R the first principal component explains a larger share of the market covariance matrix, which corresponds to the collapse of inner market dimension and highly correlated price movement during a period of a bubble or a crash. The trend is more pronounced in market crash situations as shown in the left upper part of the graph. Unlike from Fig. 4, in which the samples do not overlap, moving window on the time series is used here to increase the size of the statistics.

For large values of ρ, near the upper bound of 1, the log returns of all the stocks move in unison, and the intrinsic market dimension reduces as the systematic risk prevails and the price/return movements become uniform. Figure 5 indeed shows this regime is more common when the market index return R is large. Similar to the case of Fig. 4, a significant data spread exists, nevertheless, for low values of R. There is an asymmetry in Fig. 5, namely the values near $\rho = 1$ are more populated for large negative values of R, showing the prevalence of correlated movements in the case of market panic as compared to the case of market euphoria. Similar findings have been found previously in the literature [5]. These findings roughly persist when we alter the definition of the market index, be it an index of prices or index of returns, with the weights defined using the coordinates of the first principal component eigenvector.

5 Concluding Remarks

We have studied a model financial portfolio universe of 18 major stock titles for which (1) 40-year trading history is available and (2) which are at present the formal constituents of the DJIA. Stylized features in the dataset were observed, such as the fat tails in the histogram of the log returns of market index, clustering of the volatility, or the correlation of high covariance measures with large negative values of the index log return. Using the covariance matrix no precursor of market turbulence could be found. Statistical correlation between the situations when the market moves in unison and the aggregate index returns are high was observed, but the emergence of universal features and scaling laws for various-length subsamples of the time series remains to be elucidated.

References

1. Bouchaud, J.P., Potters, M.: Theory of Financial Risks: From Statistical Physics to Risk Management. Cambridge University Press, Cambridge (2000)
2. Sornette, D.: Why Stock Markets Crash: Critical Events in Complex Financial Systems. Princeton University Press, Princeton (2002)
3. Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L., Stanley, H.E.: Universal and Nonuniversal Properties of Cross Correlations in Financial Time Series. Phys. Rev. Lett. 83, 1471–1474 (1999)
4. Chib, S., Nardari, F., Shephard, N.: Analysis of High Dimensional Multivariate Stochastic Volatility Models. Journal of Econometrics 134, 341–371 (2006)
5. Preiss, T., Kenett, D.Y., Stanley, H.E., Helbing, D., Ben-Jacob, E.: Quantifying the Behavior of Stock Correlations Under Market Stress. Scientific Reports 2, 752, 1–5 (2012)
6. Conlon, T., Ruskin, H.J., Crane, M.: Cross-correlation Dynamics in Financial Time Series. Physica A 388, 705–714 (2009)
7. Hamao, Y., Masulis, R.W., Ng, V.: Correlations in Price Changes and Volatility Across International Stock Markets. Rev. Financ. Stud. 3, 281–307 (1990)
8. Fenn, D.J., Porter, M.A., Williams, S., McDonald, M., Johnson, N.F., et al.: Temporal Evolution of Financial-Market Correlations. Phys. Rev. E 84, 026109, 1–13 (2011)
9. Balogh, E., Simonsen, I., Nagy, B., Neda, Z.: Persistent Collective Trend in Stock Markets. Phys. Rev. E 82, 066113, 1–9 (2010)
10. Ren, F., Zhou, W.-X.: Dynamic Evolution of Cross-Correlations in the Chinese Stock Market. PLoS ONE 9, e97711, 1–15 (2014)
11. Kritzman, M., Li, Y.Z., Page, S., Rigobon, R.: Principal Components as a Measure of Systemic Risk. J. Portf. Manag. 37, 112–126 (2011)