Geometric Visualization of Dynamic Asset Allocation

Wei-Fang Niu[ǂ*], Kehluh Wang[*], Jenher Jeng[ǂ*]

G5 Capital Management, Ltd.[ǂ]

Graduate Institute of Finance, National Chiao Tung University[*]

November, 2009

Abstract

In wake of the currently ongoing financial crisis, how to monitor and track systemic risks of financial markets has become the most important task for the practitioners and researchers. Due to flourishing innovations of multi-asset-linking complex derivatives and the gradually integrated global cross-asset trading networks, a well diversified portfolio which can adapt to trends and cycles of capital flows and spot clues of systemic risks need be constructed in the conceptual framework of dynamic asset allocation (DAA). To support the methodological development of DAA by computational inventions, the following steps are suggested. First, a universe of attainable assets should be specified and classified. Next, a measure for the distances between pairs of assets is to be established. Finally, multidimensional scaling techniques are implemented to shape an asset-allocation map as a geometric visualization tool for monitoring the changing landscape of the capital market and its risk structure of asset allocation. Based on the canonical map, DAA strategies then can be easily built by visual senses or quantitative simulations with performances being more intuitively and systematically tracked. Empirical results on several major markets also revealed the effectiveness and depth of this approach. This approach can not only provide a console to implement DAA strategies for constructing portfolios, but also open a window to explore the new concept of DAA in the progress of modern portfolio theory advancing from capital asset pricing model (CAPM) and arbitrage pricing theory (APT).

Keywords: diversification, portfolio allocation, multidimensional scaling

1. Introduction

Diversification is the ultimate principle for asset allocation. The elementary part of this paper is devoted to provide a quantitative approach for the construction of well diversified portfolios. The key is the introduction of the multidimensional scaling algorithm while a proper distance measure can be defined. The benefits of diversification are also investigated under this scheme. And the technique may also initiate a methodology for further exploration of the financial markets.

In the modern portfolio theory of Markowitz (1952, 1959, and 1987) and Sharpe (1963, 1964), diversification is generally mentioned in the same breath with the implementation of the market portfolio. Measuring the level of diversification thus involves only comparisons of the variances or the compositions with the market portfolio (Goetzmann and Kumar, 2008). For example, based on random sampling and considering the excess volatility relative to an equally weighted index, Campbell et al. (2001) inferred that 20 stocks are essentially necessary for the 1963-85 period but 50 stocks for the 1986-97 period. Then the diversification puzzle is that average investors hold very limited amounts of stocks.

The rise of the exchange traded funds (ETF's) and index funds in recent years can be viewed as the practice of diversification. However, not all of these types of funds are managed with full replications. In fact, index tracking can be also carried out through sampling (Schoenberg, 2004). And there can be found literatures addressing the problem of index tracking with only a small set of stocks, for example, Corielli and Marcellino (2005). This clearly contradicts the assertion about diversification from the context of the MPT.

In fact, the contradiction comes from the differences between random sampling of stocks and active selection with respect to the correlation structure of the market. Intuitively and technically, the concept of diversification should be referred to "active" investments in securities that are uncorrelated or even negatively correlated. With the number of assets being fixed, a portfolio composed of negatively correlated assets is certainly better diversified than the one composed of positively correlated assets only. And furthermore, since the correlation structure of the market may evolve over time (Engle, 2009), the degree of diversification for a portfolio certainly may also change temporally.

On the other hand, price comovements among securities in one market generally exist (Barberis et al., 2005). This implies that, while considering the correlation between two stocks, their correlations with a third stock could also provide information. So on deciding whether a portfolio is well diversified, a universe of securities containing the securities selected should be specified a priori.

Thus, to dynamically achieve diversification, what really matter is if the persistency of the correlation structure of the market really exists and if an adequate measure of correlation between two assets can be found. While the measure is determined, a correspondent metric as distance generally can be derived and calculated. Then the multi-dimensional scaling (MDS) algorithm can be the natural choice to map all the assets into a low-dimension space. Then allocation strategies or asset clustering can be more intuitively shaped by the "map". For example, Groenen and Franses (2000) used this approach to explore the time-varying correlations across 17 stock markets.

So a practical procedure for constructing a well diversified portfolio would be as follows. First, a universe of assets should be identified. This specifies the largest range of assets to be selected. Next, define a measure for the distances between pairs of assets. A simple choice may be the standard deviation of the spread of the (standardized) returns of assets, which just corresponds to the linear correlation. Obviously, more sophisticated measures that help capture the characteristics of distribution tails may be also used. Finally, by applying multidimensional scaling or other similar techniques a market map can be drawn for constructing allocation strategies. This step indeed is a dimension reduction procedure to reduce the $n(n-1)/2$ correlations into k-tuple coordinates of n objects.

Once the orientations of the stocks are distinguished, two quantitative criteria are proposed and can be applied for the sampling scheme to construct portfolios. Both the two criteria and their combinations lead to allocations of geometric symmetry on the map.

Several allocation strategies are designed with the map. Datasets from the US, UK, Japan and Taiwan from 2004 to the beginning of 2009 are used as examples. With fixed number of stocks selected, it is found that the diversified portfolios would not necessarily have lower variances of returns. However, these diversified portfolios will track the market index more closely and may be more efficient than those ill diversified.

The remainder of this paper proceeds as follows. A brief review and implementation details about MDS are presented in Section 2; empirical investigations and further applications are in Section 3; and Section 5 contains the conclusions.

2. <u>M</u>ulti-<u>D</u>imensional <u>S</u>caling

Multidimensional scaling (MDS) is a set of techniques often used in information visualization for exploring similarities or dissimilarities in data. With measurements of dissimilarity, d_{ij}, for the (i,j) pairs of n objects as inputs, this method can be used to project these objects into a low dimension space $\mathcal{X}$, the MDS space of dimensionality m:

$$f : d_{ij} \rightarrow d_{ij}^{(m)}(\mathcal{X}).$$

Thus, viewing each stock as one object and taking a proper measure of distance, the correlation structure of the market can be easily visualized. It is noted that employing MDS to map n stocks into a k-dimension space in fact involves a dimension reduction from $n(n\text{-}1)/2$ to $n{\cdot}k$ and thus also a denoise process. This also echoes to the idea that information about correlations between two stocks can be enhanced through the networks of all stocks.

2.1 The algorithm

The input for MDS is a distance matrix with element d_{ij} to represent the dissimilarities between the i-th and j-th objects. The basic properties required for these distances are listed as follows:

(Nonnegativity) $\quad d_{ii} = d_{ii} = 0 \le d_{ij}$,

(Symmetry) $\quad d_{ij} = d_{ji}$,

(Triangle inequality) $\quad d_{ij} \le d_{ik} + d_{kj}$.

Let n be the number of distinct objects and p be the pre-specified dimensionality of the space. The objective of MDS is to find an $n\times m$ matrix X whose raw vectors represent the coordinates in the new space for each object. A stress function is usually used as the object of minimization:

$$stress = L(X) = \frac{\sum_{i<j} \left(d_{ij} - d_{ij}^{(m)}(\mathcal{X})\right)^2}{\sum_{i<j} d_{ij}^{(m)}(\mathcal{X})^2},$$

where $d_{ij}^{(m)}(\mathcal{X})$ represents the distance for the (i,j) pair of objects in $\mathcal{X}$.

Note that the stress is a measure of badness of fit. In general, the higher the number of objects n, the higher the stress could be. And the higher the dimensionality m, the lower the stress will be.

The minimization of the stress generally cannot be carried out algebraically. Numerical procedures can be found in many mathematics or statistics software. And it should be noted that the MDS solutions are indeed invariant with respect to location shifts and rotations since these operations do not alter the distances between pairs of points.

2.2 Implementing MDS to the stock markets

It is interesting and tricky to determine how dissimilar two stocks are. A commonly acceptable concept may be related to the correlation of the two price returns of the stocks. So the simplest way is to view the returns as random variables, real valued functions on some specified sample space, and to calculate the standard errors of the differences of the two returns.

However, due to that some stocks tend to be more volatile and some others less volatile, a standardization procedure should be taken to prevent that volatile stocks are always far away others.

For simplicity, in this paper the procedure following the Riskmetrics manual is taken, that is, the exponentially weighted moving average model (EWMA) or the IGARCH model is used for the variance of the return of stock i at time t:

$$r_{i,t} = \sqrt{h_{i,t}}\, z_{i,t}\,,$$

$$h_{i,t} = \lambda h_{i,t-1} + (1-\lambda) r_{i,t-1}^2\,,$$

where $r_{i,t}$ represents the return of stock i at time t and $h_{i,t}$ its variance. Then the distance $d_{ij,t}$ between two stocks at time t would be

$$d_{ij,t} = \sqrt{\frac{1}{k}\sum_{l=0}^{k}\left(z_{i,t-l} - z_{j,t-l}\right)^2}\,,$$

where k is the pre-specified window length.

Although the EWMA approach may seriously ignore some important properties of the financial time series and thus many critics arose, it indeed provides as a quick first-step and acceptable filter for the unit innovations. The distance measure with the standard error of the difference of the returns is also the simplest but easy to implement.

Obviously all the setups above for calculating distance measures can be modified. Extensions of distance measures such as L^p norm, the difference of joint entropy and mutual information, and the Hellinger distance are possible alternatives while specific forms of joint distributions are assumed. And prior information such as sectors, large or small cap, and PE ratios can be also used in the definitions of distances.

2.3 Clustering securities with MDS: simulation studies

In this section, properties of the maps by implementing MDS to time series from some simple data generation processes are investigated through simulations.

First, assume that stock prices processes geometric Brownian motions with constant drift, volatility and correlations between each other. It is intuitive that the map will look like a circle or an ellipse. Figure 1 illustrates the results for different sizes of n.

Next suppose that the returns follows a factor model and these factors follow an AR(1) model, that is,

$$r_{it} = \sum_{j=1}^{k_i} b_{ij} F_{jt} + \varepsilon_{it}$$

where F_{jt} denote the non-correlated risk factors with AR(1) structure, b_{ij} denote the factor loadings and ε_{it} denote the uncorrelated errors.

For a total number of stocks n = 120, suppose that the first 20 stocks corresponds to the first factor only, the second 20 stocks to the first two factors, and so on. That is $k_1= \cdots = k_{20}=0$, $k_{21}= \cdots = k_{40}=1$, $k_{101}= \cdots = k_{120}=5$. In Figure 2, it is seen that these 120 stocks are successfully clustered into 6 groups. This essentially demonstrates the ability of classification of MDS.

Furthermore, consider the data generated from a GARCH process

$$r_{i,t} = \sqrt{h_t} \cdot \Delta \cdot \Gamma \cdot \varepsilon_t,$$

$$h_{i,t} = \alpha_{i,0} + \alpha_{i,1} r_{i,t-1}^2 + \beta_1 h_{i,t-1},$$

where ε_t's are random vectors with standard normal distribution or t distribution, Δ and Γ are $n \times n$ matrices designed to produce clustering effect and satisfy δ_{ij}=1 or -1,

$\gamma_{ij} > 0$, and $\sum_j \gamma_{ij}^2 = 1$. The setting states that there are n risk factors and i-th stock has s positive response to the j-th factor while δ_{ij}=1, and vice versa.

Figure 3a shows a map when there is only one cluster, that is δ_{ij}=1 for all i and j. The left panel corresponds to normally distributed innovations and the right panel t_5 distribution. As expected, it is seen that the points in the right map are more wildly spreaded since the t distribution has a fatter tail.

While there are 10 stocks have negative responses and there are two clusters, Figure 3b shows the results. The 10 stocks are better clustered from the others when the innovations are generated from the normal distribution. For the right panel in which data are generated from the t distribution, the 10 stocks are then mixed up with other stocks.

The maps with real data from the US and Taiwan in 2008 are shown in Figure 4. Clear clustering can be seen for the US data and some mix-up exists for the Taiwan data. These results illustrate how MDS may be applied for clustering.

3. Constructing portfolio with MDS

3.1 Data

In this paper, daily returns of stocks from four countries are used for illustrations. The stock markets and the benchmarks are listed in Table 1. The whole datasets consists of daily close prices of the index and the component stocks.

Stock prices of the Taiwan stock market are collected from TEJ database. The dataset consist of 1219 daily prices of fifty companies from April 1, 2004 to March 2, 2009. Stock prices of the other three stock markets are collected from Datastream database. These data consist of 1219 daily prices of many different companies from April 1, 2004 to December 2, 2008.

On each day, a MDS map can be constructed with distances calculated with standardized innovations of the past 60 trading days.

3.2 Strategies for diversification

When the relations among all the stocks can be represented by a map, two criteria on diversification for building portfolios are proposed here.

Definition. *Diversification with respect to the universe*. With respect to a universe of assets $\mathcal{U}$ and the corresponding MDS space $\mathcal{X}$, a portfolio is said to be with respect to the universe if the sum of the distances of all pairs of selected and unselected assets is minimized.

Definition. *Diversification within selected assets*. With respect to a universe of assets $\mathcal{U}$ and the corresponding MDS space $\mathcal{X}$, a portfolio is said to be diversified within selected components if the sum of the distances among pairs of the selected assets is maximized.

Clearly the two criteria correspond to different schemes on diversification of assets. Diversification with respect to the universe asserts the sampled stocks to perform close to the average of all of the stocks, while diversification within selected assets pursues least inter-dependence among selected stocks.

However, it should be stressed that, while the stocks are distributed symmetric about the origin (the centroid of the points on the map), diversification with respect to the universe would result in stocks concentrated around to the origin and diversification within selected assets would lead to stocks that are located at the boundary. Both of the selections will be symmetric about the origin. Thus the two concepts, diversification and symmetry, are elaborately linked.

The assertion above indeed has been originated from the MDS framework. Since the inputs contain no information for absolute orientation, the only meaningful reference point on the map would be the origin (and infinities). While the stocks are symmetrically distributed, a well diversified portfolio must be symmetrically distributed too.

So four sampling schemes based on the MDS map are employed for the construction of portfolios as shown in the Figure 5. The first one is composed of stocks with the distances from the center ranked at equally spaced quantiles. The second one corresponds to the selection of the stocks near the center of the map. The third one consists of stock on the boundary of the map. The fourth one is comprised of stocks at the positive of the x-axis. For simplicity, these portfolios are denoted respectively P1, P2, P3 and P4.

The portfolio P2 with stocks near the center indeed follows the first criterion, and P3 corresponds to an approximation to the second criterion. The first one, which meets common intuition about diversification, is in fact a hybrid realization of both of the two criteria. And obviously P4 shall be poorly allocated since it is asymmetric about the origin.

3.3 Empirical performances of the portfolios

The backtests proceed as follows. First standardized innovations for each stock on each day are filtered as described in Section 2.2. Distances between two stocks on each day are then calculated with standardized innovations in the past 60 trading days. After running the MDS algorithm, 10 stocks are selected for the four portfolio mentioned above for each market. Each portfolio is constructed with weights inversely proportional to the volatilities:

$$w_{it} \propto \frac{1}{\sqrt{h_{it}}},$$

and the return of the portfolio on day t will be

$$R_t = \frac{1}{\sum_i \frac{1}{\sqrt{h_{it}}}} \sum_i \frac{1}{\sqrt{h_{it}}} \cdot r_{it}.$$

It is 21 trading days for each portfolio to be held. At expiry, a new map is drawn and another 10 stocks are sampled for the next 21 trading days. Summary statistics for the returns of these portfolios are also shown in Table 2.

Three diversity measures for the returns are considered: standard deviation, range and the difference between 1st and 3rd quartiles. It is easily seen that these measure for P4 are not necessarily the largest across all the markets in different periods. In fact, they tend to be lower than those of other portfolios. However, P4 almost always has the lowest correlations with the market index except for FTSE in the period after July of 2007.

For further investigations on the performances of these portfolios, the whole sampling period is divided into 46 frames of length of 21 trading days and the performance in each frame will be viewed as an independent sample. Summary statistics for volatilities of the four portfolios in 46 frames are tabulated in Table 3. Note that the correlations are calculated with the logarithms of the volatilities. Pairs plot for the logarithms of volatilities for the index and four portfolios in Taiwan 50 and S&P 100 are shown in Figure 6. Linear relations can be easily seen for all pairs of volatilities and correlations of P4 with the index are generally lower than others.

A multiple comparison procedure based on Tukey's Honestly Significant Differences (HSD) test is executed to test the hypotheses that all of the portfolios have the same levels of volatilities regarding the changes of volatilities of the index. Specifically, for the volatility of portfolio i at period t, the following linear model is considered

$$\log(v_{it}) \sim \alpha_i + \beta \log(v_{0t}),\ i=1, ,2 ,3, 4$$

where $\log(v_{0t})$ is the volatility of the index.

The results are shown in Table 4 and meet the previous findings. The volatilities of P4 are generally not the largest. Even more, they are significantly the lowest for the S&P 100 case. This phenomenon is obviously worth more investigations.

Next, the efficiencies of the portfolios are investigated by the spread of the M^2 risk-adjusted performance measures (Modigliani and Modigliani, 1997) to the index returns, that is,

$$\frac{\sigma_m}{\sigma_P}\left(R_P - r_f\right) + r_f - R_m,$$

where R_m, σ_m, R_P, and σ_P are respectively the returns , standard deviations of the index and the index, and r_f is the risk free rate.

The boxplots for the spreads are shown in Figure 7 and the results for multiple comparison are in Table 5. Interestingly it is seen that the efficiencies of the first three portfolios differ across markets. However, except NEKKEI for which all the four portfolios are very similar, the ill diversified forth portfolio P4 is generally dominated by or at most comparable with some of the other three portfolios the rest. This seems to imply that benefits via diversifications essentially exist but optimal strategies may depend on which markets.

4. Conclusions and Extensions

In this paper, a systematic approach to diversification of assets with MDS is proposed. The dimension reduction technique converts the problem from dealing with correlations of pairs of stocks to accessing points in a low dimension space. Thus visualization of the correlation structure is possible and dynamic quantitative allocation strategies can be also made.

From this context, diversification can be elaborately related to symmetry. Diversified portfolios do not necessarily have lower variances, but they track the index or a market portfolio more closely and thus may perform efficiently relative to certain risk levels.

Further applications can be easily found, for example the index tracking problem. Higher levels of correlations with the index both for returns and volatilities implies that the symmetrically allocated portfolios match the index better and should be considered favorably. However, to meet with the challenges, more elaborate distance measures should be introduced to adequately address the tail behaviors of the return distributions.

Furthermore, since the correlations of the stocks are represented as coordinates in a low dimension space, the structural change can be also dynamically monitored through this tool for more nonlinear modeling purposes as the industrial or even systemic risk factors evolves, for example, the AR(1)-modeling framework can be extended to a more general setting (Jeng 2008). This method thus provides as a basis to geometrically shape and monitor the frequently seen qualitative descriptions of fundamental analysis for market tructures which are constantly reshaped by macroeconomic growth-value prospects (investment styles), new investment opportunities and innovations of financial instruments and investment vehicles. Also, this method can provide easy and quick cross-market and cross-horizon comparisons for obtaining more complete and deeper market insights.

What is the most important is that this paper serves as a demonstration for the application of data mining tools in finance. Beyond the variances of the individual stocks and the correlation structures, there certainly will be more quantities useful for the description of the market conditions. It is then the data that reveal the nature of the markets.

References

1. Barberisa, N., Shleifer, A. and Wurgler, J., 2005, Comovement, *Journal of Financial Economics*, 75, 283–317.

2. Bera, A. K. and Kim, S., 2002, Testing constancy of correlation and other specifications of the BGARCH model with an application to international equity returns, *Journal of Empirical Finance*, 9, 2, 171-195.

3. Borg, I. and Groenen, P. 1997, Modern Multidimensional Scaling: Theory and Applications, Springer Series in Statistics, Springer.

4. Campbell, J. Y., Lettau, M., Malkiel , B. G. and Xu, Y., 2001, Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk, *Journal of Finance*, 56,1, 1-43.

5. Corielli, F. and Marcellino, M., 2005, Factor based index tracking, *Journal of Banking & Finance*, 30, 2215–2233.

6. Cox, T. F. and Cox, M. A. A., 2001, Multidimensional Scaling, Chapman & Hall/ CRC.

7. Dwyer, T. and Gallagher, D. R ., 2004, Visualising changes in fund manager holdings in two and a half-dimensions, *Information Visualization*, 3, 4, 227.

8. Engle, R., 2009, Anticipating correlations, Princeton.

9. Gabi, G., 2005, Semi-correlations as a tool for geographical and sector asset allocation, *The European Journal of Finance*, 11, 3, 271-281.

10. Goetzmann, W. N. and Kumar, A., 2008, Equity portfolio diversification, *Review of Finance*, 12, 433-463.

11. Goyal, A. and Santa-Clara, P., 2003, Idiosyncratic risk matters! *Journal of finance*, LVIII, 975-1007.

12. Groenen, P. J. F. and Franses, P. H., 2000, Visualizing time-varying correlations across stock markets, *Journal of Empirical Finance*, 7, 155–172.

13. Hung, S. Y., Liang, T. P. and Liu, V.W. C.,1996, Integrating arbitrage pricing theory and artificial neural networks to support portfolio management, *Decision Support Syst.*, 18, 301–316.

14. Jeng, Jenher, 2008, General adaptive principals for tracking nonlinear dynamics, Conference Paper, *IASC 2008*, Japan.

15. John, G. H., Miller, P. and Kerber, R.,1996, Stock selection using rule induction, *IEEE Expert*, 11, 52-58.

16. Kruskal, J. B., 1964, Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, *Psychometrika*, 29, 1–27.

17. Markowitz, H. M., 1952, Portfolio Selection, *Journal of Finance*, l7, 77-91.

18. Markowitz, H. M.,1959, Portfolio Selection: Efficient diversification of investments, New York: Jonh Wiley & Sons.

19. Markowitz, H. M.,1987, Mean-variance analysis in portfolio choice and capital markets, New York: Basil Blackwell.

20. Miller, R. G., 1981, Simultaneous Statistical Inference. Springer.

21. Modigliani, F. and Modigliani, L. 1997, Risk-adjusted performance, *The Journal of Portfolio Management*, 23, 2, 45-54.

22. Potvin, J., Soriano, P. and Vallée, M.,2004, Generating trading rules on the stock

markets with genetic programming, *Comput. Oper. Res.* 31, 7, 1033-1047.

23. RiskMetrics Group, 1999, Risk management: a practical guide.

24. Ross, S. A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory*, 13, 341-360.

25. Schoenfeld, S. A., 2004, Active Index Investing, Hooboken, New Jersey: Wiley.

26. Statman, M., 2004, The diversification puzzle, *Financial Analysts Journal*, 60, 4, 44-53.

27. Tegarden, D., 1997, Business information visualization, *Communications of the Association for Information Systems* 1, 4.

Table 1. Stock markets and the benchmarks.

Country	Benchmark
US	S&P 100
UK	FTSE 100
Japan	NIKKEI 225
Taiwan	TSEC Taiwan50

Table 2. Summary statistics for performances of portfolios in the four markets.

		Taiwan 50					S&P 100					FTSE					NIKKEI				
		index	P1	P2	P3	P4	index	P1	P2	P3	P4	index	P1	P2	P3	P4	index	P1	P2	P3	P4
Whole period	Mean	-0.01%	0.06%	0.04%	0.02%	0.06%	-0.03%	0.02%	0.00%	0.00%	-0.03%	-0.01%	0.03%	0.04%	0.00%	0.05%	-0.01%	0.01%	0.00%	0.00%	-0.01%
	Max	6.45%	6.39%	5.49%	6.76%	5.78%	11.24%	14.20%	12.04%	12.45%	13.53%	9.84%	7.99%	7.69%	14.36%	6.72%	14.15%	18.70%	16.55%	15.69%	17.78%
	3rd Quartile	0.79%	0.83%	0.80%	0.86%	0.92%	0.51%	0.63%	0.65%	0.66%	0.57%	0.61%	0.72%	0.74%	0.79%	0.82%	0.83%	0.99%	1.03%	1.00%	0.89%
	1st Quartile	-0.75%	-0.72%	-0.67%	-0.74%	-0.76%	-0.48%	-0.53%	-0.57%	-0.60%	-0.58%	-0.57%	-0.62%	-0.59%	-0.67%	-0.60%	-0.77%	-0.92%	-0.82%	-0.89%	-0.92%
	Min	-6.50%	-6.91%	-6.04%	-6.51%	-6.25%	-8.78%	-9.05%	-8.17%	-10.41%	-8.76%	-8.85%	-8.04%	-6.70%	-11.46%	-7.05%	-11.41%	-11.90%	-13.00%	-11.92%	-12.46%
	STD	1.57%	1.61%	1.42%	1.66%	1.54%	1.49%	1.58%	1.48%	1.69%	1.44%	1.45%	1.52%	1.43%	1.87%	1.45%	1.84%	2.06%	2.06%	1.96%	2.13%
	Range	12.95%	13.30%	11.53%	13.27%	12.02%	20.02%	23.25%	20.22%	22.86%	22.30%	18.69%	16.03%	14.39%	25.82%	13.77%	25.56%	30.61%	29.55%	27.61%	30.23%
	Difference between 1st and 3rd quartiles	1.54%	1.55%	1.47%	1.59%	1.67%	1.00%	1.16%	1.22%	1.26%	1.15%	1.17%	1.34%	1.34%	1.46%	1.42%	1.60%	1.91%	1.85%	1.89%	1.81%
	Correlation with the index	1	0.919	0.901	0.913	0.831	1	0.948	0.923	0.947	0.922	1	0.905	0.878	0.881	0.891	1	0.942	0.936	0.930	0.924
Before 7/31/2007	mean	0.08%	0.16%	0.11%	0.11%	0.13%	0.03%	0.07%	0.09%	0.05%	0.03%	0.05%	0.08%	0.10%	0.08%	0.11%	0.09%	0.10%	0.11%	0.12%	0.09%
	max	3.08%	3.42%	3.71%	4.05%	4.25%	2.04%	2.44%	3.14%	2.37%	3.07%	2.64%	3.60%	4.43%	5.55%	3.68%	3.58%	4.74%	5.59%	4.55%	3.03%
	3rd Quartile	0.71%	0.78%	0.72%	0.79%	0.79%	0.41%	0.52%	0.55%	0.54%	0.46%	0.51%	0.63%	0.63%	0.69%	0.71%	0.69%	0.87%	0.84%	0.89%	0.74%
	1st Quartile	-0.47%	-0.44%	-0.48%	-0.44%	-0.48%	-0.30%	-0.35%	-0.36%	-0.44%	-0.38%	-0.38%	-0.42%	-0.34%	-0.48%	-0.41%	-0.47%	-0.53%	-0.54%	-0.55%	-0.47%
	min	-4.96%	-4.22%	-4.71%	-4.14%	-4.95%	-3.57%	-3.11%	-3.81%	-3.01%	-3.29%	-3.15%	-4.21%	-3.83%	-4.37%	-3.86%	-4.14%	-5.55%	-3.37%	-4.82%	-3.84%
	STD	1.05%	1.04%	1.00%	1.06%	1.15%	0.65%	0.74%	0.78%	0.75%	0.74%	0.74%	0.91%	0.92%	0.99%	0.98%	1.07%	1.20%	1.16%	1.22%	1.06%
	Range	8.04%	7.65%	8.41%	8.19%	9.20%	5.60%	5.56%	6.95%	5.38%	6.36%	5.79%	7.81%	8.26%	9.92%	7.53%	7.73%	10.29%	8.96%	9.37%	6.87%
	Difference between 1st and 3rd quartiles	1.19%	1.22%	1.20%	1.23%	1.28%	0.72%	0.87%	0.91%	0.98%	0.84%	0.89%	1.05%	0.97%	1.17%	1.12%	1.16%	1.40%	1.38%	1.45%	1.21%
	Correlation with the index	1	0.881	0.847	0.880	0.755	1	0.874	0.839	0.858	0.863	1	0.899	0.804	0.883	0.850	1	0.888	0.886	0.903	0.875
After 7/31/2007	mean	-0.14%	-0.10%	-0.06%	-0.11%	-0.03%	-0.12%	-0.07%	-0.12%	-0.08%	-0.12%	-0.09%	-0.06%	-0.05%	-0.13%	-0.03%	-0.15%	-0.12%	-0.15%	-0.15%	-0.16%
	max	6.45%	6.39%	5.49%	6.76%	5.78%	11.24%	14.20%	12.04%	12.45%	13.53%	9.84%	7.99%	7.69%	14.36%	6.72%	14.15%	18.70%	16.55%	15.69%	17.78%
	3rd Quartile	0.99%	1.12%	1.09%	1.45%	1.19%	0.72%	0.94%	0.99%	0.97%	0.89%	0.97%	1.07%	1.06%	1.17%	0.96%	1.14%	1.34%	1.35%	1.19%	1.41%
	1st Quartile	-1.32%	-1.33%	-1.12%	-1.41%	-1.30%	-1.17%	-1.18%	-1.18%	-1.32%	-1.03%	-1.12%	-1.22%	-1.17%	-1.37%	-0.96%	-1.36%	-1.55%	-1.52%	-1.65%	-1.87%
	min	-6.50%	-6.91%	-6.04%	-6.51%	-6.25%	-8.78%	-9.05%	-8.17%	-10.41%	-8.76%	-8.85%	-8.04%	-6.70%	-11.46%	-7.05%	-11.41%	-11.90%	-13.00%	-11.92%	-12.46%
	STD	2.09%	2.16%	1.86%	2.24%	1.97%	2.22%	2.30%	2.11%	2.47%	2.07%	2.10%	2.12%	1.96%	2.70%	1.96%	2.54%	2.85%	2.85%	2.65%	3.03%
	Range	12.95%	13.30%	11.53%	13.27%	12.02%	20.02%	23.25%	20.22%	22.86%	22.30%	18.69%	16.03%	14.39%	25.82%	13.77%	25.56%	30.61%	29.55%	27.61%	30.23%
	Difference between 1st and 3rd quartiles	2.31%	2.44%	2.21%	2.86%	2.49%	1.90%	2.12%	2.16%	2.29%	1.93%	2.08%	2.29%	2.24%	2.54%	1.92%	2.50%	2.89%	2.87%	2.84%	3.28%
	Correlation with the index	1	0.932	0.922	0.924	0.864	1	0.958	0.940	0.959	0.934	1	0.909	0.903	0.881	0.914	1	0.955	0.948	0.938	0.937

Table 3. Summary statistics for volatilities (daily) of the four portfolios in 46 frames. The correlations, slopes and intercepts of regressions are calculated with the logarithms of the volatilities with those of the index.

		Index	P1	P2	P3	P4
Taiwan 50	Mean	0.0149	0.0145	0.0131	0.0150	0.0143
	Max	0.0452	0.0363	0.0252	0.0380	0.0276
	Min	0.0058	0.0067	0.0057	0.0065	0.0063
	Quartile3	0.0177	0.0181	0.0166	0.0183	0.0171
	Quartile1	0.0088	0.0090	0.0088	0.0098	0.0099
	Correlation		0.9085	0.8910	0.8988	0.8820
S&P 100	Mean	0.0114	0.0124	0.0119	0.0136	0.0119
	Max	0.0473	0.0428	0.0437	0.0559	0.0488
	Min	0.0041	0.0042	0.0049	0.0046	0.0040
	Quartile3	0.0133	0.0138	0.0120	0.0157	0.0134
	Quartile1	0.0054	0.0063	0.0068	0.0064	0.0062
	Correlation		0.7808	0.7858	0.7866	0.7576
FTSE	Mean	0.0118	0.0131	0.0123	0.0147	0.0125
	Max	0.0485	0.0398	0.0406	0.0578	0.0446
	Min	0.0033	0.0042	0.0035	0.0047	0.0042
	Quartile3	0.0153	0.0189	0.0158	0.0184	0.0149
	Quartile1	0.0061	0.0071	0.0075	0.0074	0.0078
	Correlation		0.7810	0.7855	0.7497	0.7249
NIKKEI	Mean	0.0153	0.0168	0.0170	0.0166	0.0171
	Max	0.0555	0.0646	0.0635	0.0570	0.0682
	Min	0.0046	0.0054	0.0054	0.0048	0.0058
	Quartile3	0.0170	0.0190	0.0198	0.0169	0.0202
	Quartile1	0.0090	0.0101	0.0102	0.0104	0.0095
	Correlation		0.8180	0.8584	0.8392	0.8808

Table 4. Results for the multiple comparisons for the volatility levels of the index and four portfolios in the 46 frames by Tukey's HSD method. The dash lines at the right sides represent that the levels of the logarithms of volatilities do not differ significantly.

Taiwan 50			S&P 100		
=0.1	=0.05	=0.01	=0.1	=0.05	=0.01
P3	P3	P3	P3	P3	P3
P4	P4	P4	P1	P1	P1
P1	P1	P1	P2	P2	P2
P2	P2	P2	P4	P4	P4
FTSE			**NIKKEI**		
=0.1	=0.05	=0.01	=0.1	=0.05	=0.01
P3	P3	P3	P2	P2	P2
P1	P1	P1	P3	P3	P3
P4	P4	P4	P1	P1	P1
P2	P2	P2	P4	P4	P4

Table 5. Results for the multiple comparisons for the spread of M^2 to the index return for the four portfolios in 46 frames. The dash lines at the right sides represent insignificances of level differences.

Taiwan 50			S&P 100		
=0.1	=0.05	=0.01	=0.1	=0.05	=0.01
P1	P1	P1	P2	P2	P2
P4	P4	P4	P3	P3	P3
P2	P2	P2	P4	P4	P4
P3	P3	P3	P1	P1	P1
FTSE			**NIKKEI**		
=0.1	=0.05	=0.01	=0.1	=0.05	=0.01
P4	P4	P4	P2	P2	P2
P3	P3	P3	P3	P3	P3
P2	P2	P2	P1	P1	P1
P1	P1	P1	P4	P4	P4

Figure 1. Maps for normally distributed returns with constant drift, volatility and correlations.

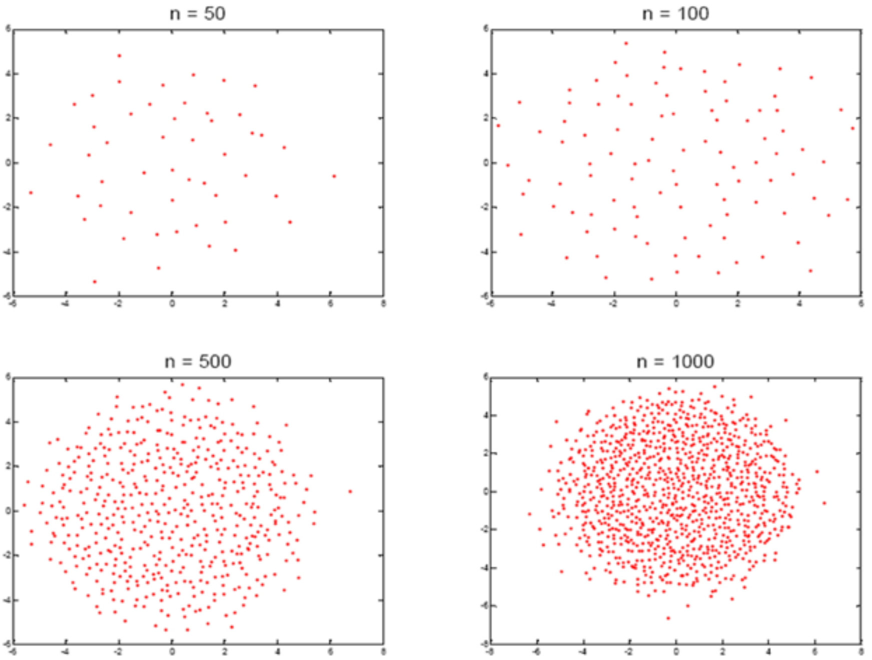


Figure 2. Maps for returns with AR(1) factors and 6 clusters.

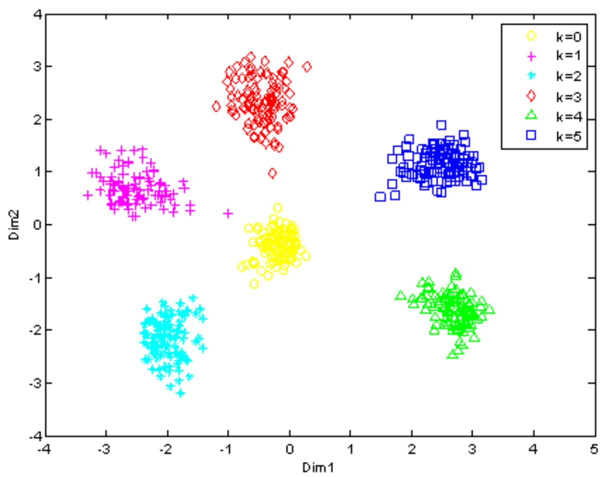
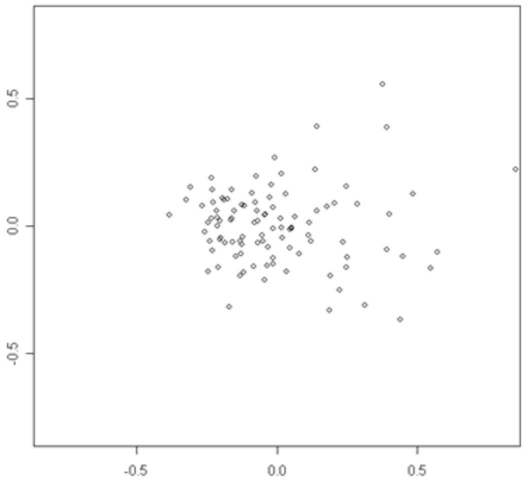


Figure 3a. Maps for returns from a GARCH model (one cluster). Left penal: normally distributed innovations. Right Penal: t_5 distributed innovations.



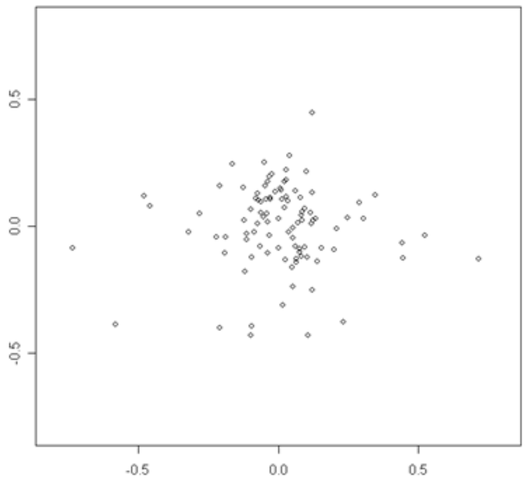
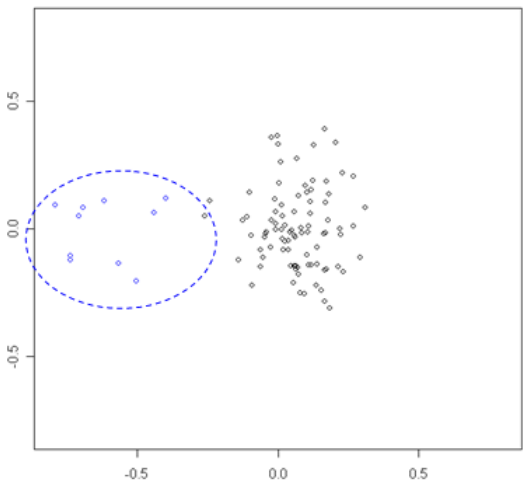


Figure 3b. Maps for returns from a GARCH model (two cluster). Left penal: normally distributed innovations. Right Penal: t_5 distributed innovations. The dash lines indicate the location of stocks in the second clusters.



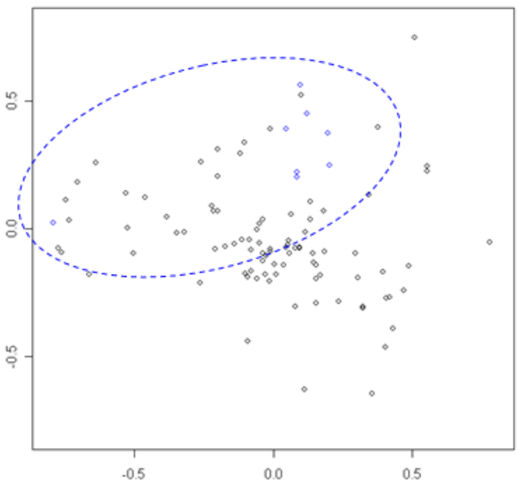


Figure 4. Clustering effects in Taiwan 50 and S&P 100..

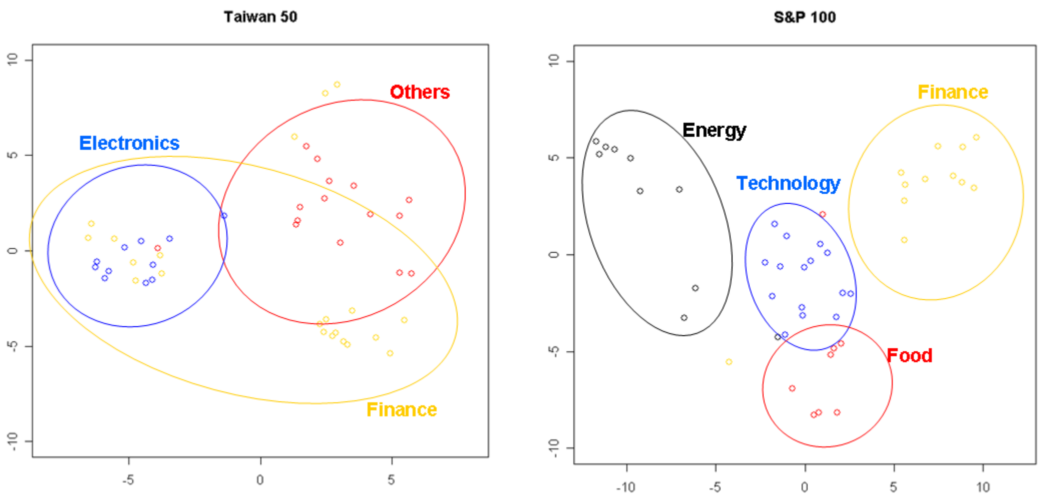


Figure 5. Sampling strategies for constructing portfolios.

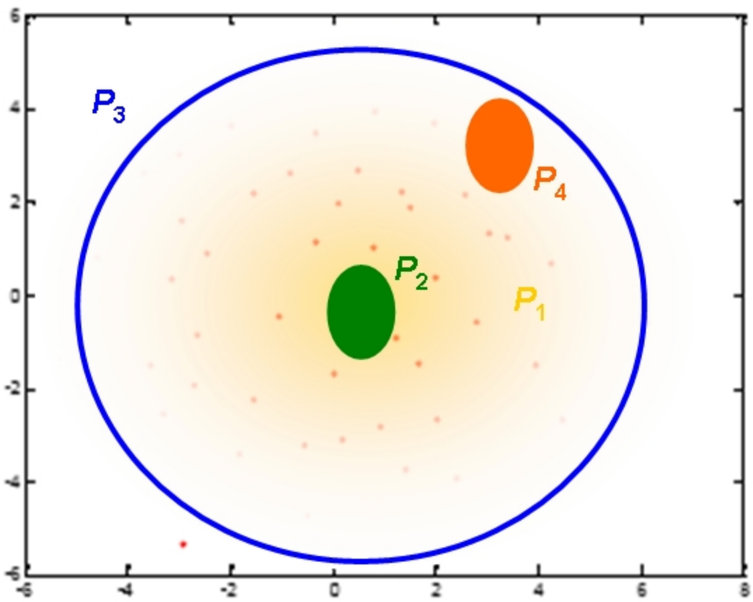
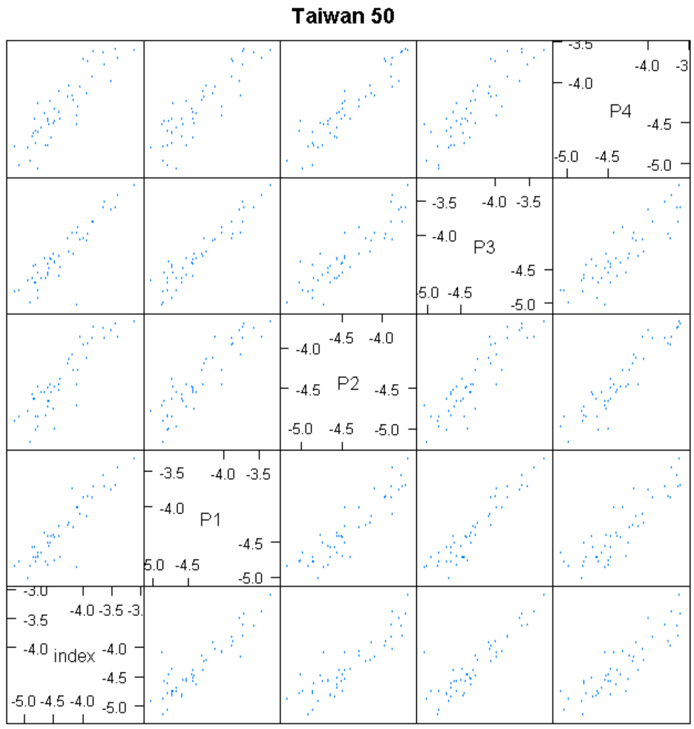


Figure 6. Pairs plot for the logarithms of volatilities for the index and four portfolios in Taiwan 50 and S&P 100 in 46 frames.



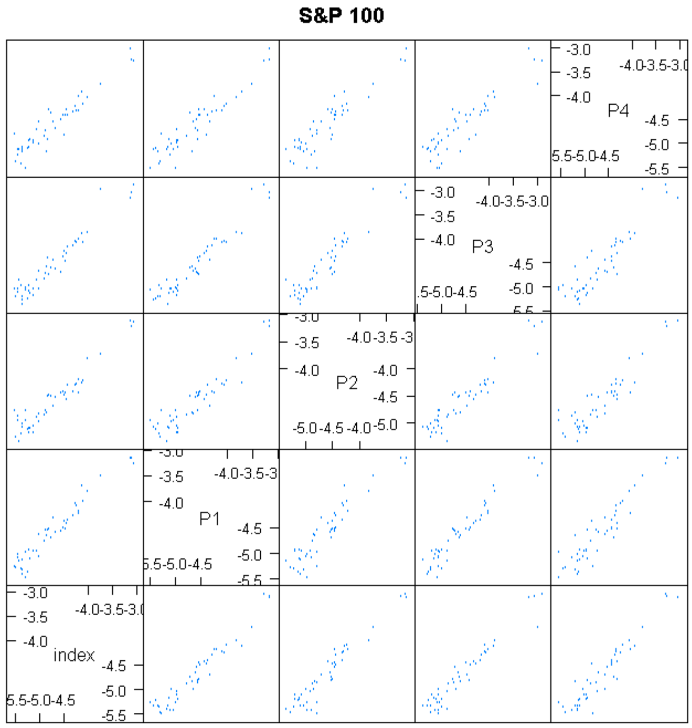


Figure 7. Boxplots for the spread of M^2 to the index return for the four portfolios in 46 frames.