Journal of Applied Finance & Banking, vol. 3, no. 6, 2013, 1-24
ISSN: 1792-6580 (print version), 1792-6599 (online)
Scienpress Ltd, 2013

Further Evidence of Deficiencies in Classical Finance

Erhard Reschenhofer[1] and Kevin Windisch[2]

Abstract

Reviewing the basics of mean–variance portfolio optimization and the capital asset pricing model, this paper discusses the plausibility of some of the underlying assumptions. It is pointed out that a positive in-sample relationship between the expected return of an asset and its covariance with the market portfolio can be a statistical artifact because it can be explained without using any economic arguments. In an empirical analysis of two sets of assets consisting of individual stocks and indices, respectively, no indication of any out-of-sample relationship is found. In the absence of such a relationship or any other additional information about the expected returns, simple averages of past returns must be used as input for portfolio-optimization procedures. Empirical evidence is presented which suggests that portfolio optimization is of little practical value in this case and, in addition, that the use of robust estimators can hardly make any difference.

JEL classification numbers: C58, G11, G17
Keywords: Portfolio optimization, Risk-return relationship, Robust estimation

1 Introduction

Modeling the returns of a set of assets as random variables with different means and variances, classical portfolio optimization tries to find a mixture of these assets which either minimizes the variance for a given expected return or maximizes the expected return for a given variance. In general, it is much easier to reduce the variance than to increase the expected return. While the variance of a mixture of assets will practically always be smaller than the minimum of the individual variances, the mean of the mixture can never exceed the maximum of the individual means (see Figure 1). A large variance reduction can already be achieved by using equal weights for all assets. Any further reduction requires knowledge of the covariances between the individual returns. Although the covariances are not only unknown but are also changing over time, they can be

[1]University of Vienna.
[2]University of Vienna.

Article Info: *Received :* July 9, 2013. *Revised :* August 20, 2013.
Published online : November 1, 2013

forecasted with some accuracy. But this means only that variance minimization is a worthwhile exercise, it does not mean that the theoretical concept of a risk-return trade-off is of any practical use. The problem is that the prediction of future returns is much harder than that of future risks. However, additional economic assumptions, which include the unrealistic possibility of unrestricted borrowing and lending at a riskfree rate [12, 15] or, alternatively, the equally unrealistic possibility of unrestricted shortselling [3], imply a linear relationship between the expected value of the return of an asset and the covariance between this return and the market return. If this hypothetical relationship was true, it could be used to obtain forecasts of expected returns from forecasts of covariances.
This paper confronts the fundamental principles of classical finance with data. The bar is not set very high. Nobody expects a perfect agreement between theory and empirical results. Previous empirical studies have already revealed striking discrepancies (for an overview, see [8]). The only remaining question is whether the theory can at least be used to improve the trading performance in a statistically significant and economically relevant manner. If the answer is no, we should first try slight modifications before we abandon classical finance and turn to more complex asset pricing models such as the multi-factor CAPM [6, 7] and the downside-beta CAPM [2, 4, 9, 10] or to a completely different approach such as behavioral finance (for a survey, see [1]). At least, the use of conventional methods for the estimation of the variances and covariances should be questioned. Robust estimators are possibly more appropriate because of the apparent deviations from normality (for a survey of robust portfolio strategies, see [5]). Although the use of robust estimators has the positive side effect that the portfolio turnover is reduced, further stabilization measures might be necessary, e.g., smoothing of the portfolio weights. The results of empirical studies [11, 16, 17] suggest that the use of robust methods may improve portfolio performance. However, the significance of these results is difficult to evaluate because the performance is usually reported only for the whole observation period and, occasionally, also for two subperiods. There is no continuous assessment. Moreover, the observation periods are often very short (typically about ten years or less).

In view of the vast amount of available financial data, it seems to be always possible to find certain assets and certain time periods which support or challenge a given hypothesis or model. To avoid this obvious danger of a data-snooping bias, the sole criterion for the selection of our data was the availability of a long history of daily prices at Yahoo!Finance. Both indices and individual stocks are used. No efforts are made to reduce the variance by searching for sets of most dissimilar assets or to increase the precision of estimates of model parameters such as the betas by replacing real assets by synthetic ones which are just collections of similar assets.

Section 2 gives a short review of the basics of mean–variance portfolio optimization, which is based on the work of Markowitz [13, 14], and the capital asset pricing model (CAPM), which is based on the work of Sharpe [15] and Lintner [12]. In this section, it is also pointed out that some unrealistic assumptions are dispensable. Section 3 presents the empirical results. The focus is on the comparison of classical methods and robust methods. Also of particular interest is the in-sample relationship between the expected return of an asset and its covariance with the market portfolio. Section 4 concludes.

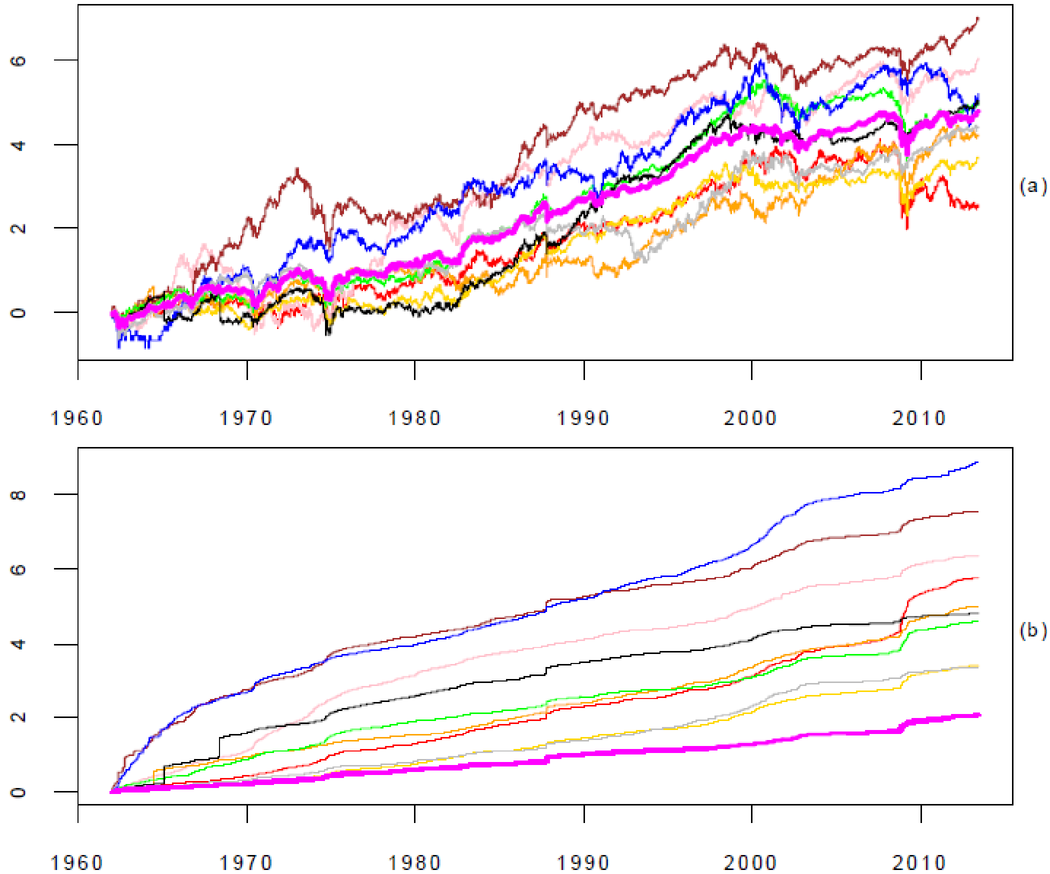

Figure 1: Comparison of cumulative returns (a) and cumulative squared returns (b) of nine components of the DJIA with the cumulative average returns (a) and cumulative squared average returns (b) from 02-01-1962 to 31-05-2013.

Data: AA (red), BA (pink), CAT (orange), DD (gold), DIS (brown), GE (green), HPQ (blue), IBM (gray), KO (black), Average (magenta)

2 Review of the Basics of Classical Finance

2.1 Portfolio Optimization

Suppose we are given K assets with stochastic returns $R_1,\ldots,R_K$. The return of a portfolio which is a mixture of these assets is determined by the portfolio weights $w_1,\ldots,w_K$, i.e.,

$$R_w = \sum_{k=1}^{K} w_k R_k = (w_1,\ldots,w_K)(R_1,\ldots,R_K)^T = w^T R \tag{1}$$

The mean and the variance of R_w are given by

$$\mu_w = E(R_w) = \sum_{k=1}^{K} w_k \underbrace{E(R_k)}_{\mu_k} = (w_1,\ldots,w_K)(\mu_1,\ldots,\mu_K)^T = w^T \mu \tag{2}$$

and

$$\sigma_w^2 = Var(R_w) = \sum_{j=1}^{K}\sum_{k=1}^{K} w_j w_k \underbrace{Cov(R_j, R_k)}_{\sigma_{ij}} = w^T \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1K} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{K1} & \sigma_{K2} & \cdots & \sigma_{KK} \end{pmatrix} w = w^T \Sigma w\,, \tag{3}$$

respectively. The method of Lagrange multipliers can be used to solve the Markowitz [13] problem of finding weights $w_1,...,w_K$ that minimize the portfolio variance (or, equivalently, half the portfolio variance) for a desired expected return

$$w^T \mu = \mu_0 \tag{4}$$

subject to the constraint

$$\sum_{k=1}^{K} w_k = w^T 1 = 1\,. \tag{5}$$

Setting the partial derivatives of the Lagrange function

$$L(w, \lambda_1, \lambda_2) = \tfrac{1}{2} w^T \Sigma w + \lambda_1(\mu_0 - w^T \mu) + \lambda_2(1 - w^T 1)$$

to zero, we obtain

$$\Sigma w - \lambda_1 \mu - \lambda_2 1 = 0\,,$$

$$\mu_0 - w^T \mu = 0\,,$$

$$1 - w^T 1 = 0\,,$$

which is a system of K+2 linear equations in the K+2 unknowns $w_1,...,w_K,\lambda_1,\lambda_2$. The first K equations yield

$$w = \lambda_1 \Sigma^{-1} \mu + \lambda_2 \Sigma^{-1} 1$$

and the last two

$$\begin{pmatrix} \mu_0 \\ 1 \end{pmatrix} = \begin{bmatrix} w^T \mu \\ w^T 1 \end{bmatrix} = \begin{pmatrix} \lambda_1 \mu^T \Sigma^{-1} \mu + \lambda_2 1^T \Sigma^{-1} \mu \\ \lambda_1 \mu^T \Sigma^{-1} 1 + \lambda_2 1^T \Sigma^{-1} 1 \end{pmatrix} = \begin{pmatrix} \mu^T \Sigma^{-1} \mu & 1^T \Sigma^{-1} \mu \\ \mu^T \Sigma^{-1} 1 & 1^T \Sigma^{-1} 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} p & r \\ r & q \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

Thus,

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} p & r \\ r & q \end{pmatrix}^{-1} \begin{pmatrix} \mu_0 \\ 1 \end{pmatrix} = \frac{1}{pq - r^2} \begin{pmatrix} q & -r \\ -r & p \end{pmatrix} \begin{pmatrix} \mu_0 \\ 1 \end{pmatrix} = \tfrac{1}{d} \begin{pmatrix} q\mu_0 - r \\ -r\mu_0 + p \end{pmatrix}$$

and

$$\begin{aligned} \sigma_w^2 = w^T \Sigma w &= (\lambda_1 \mu^T \Sigma^{-1} + \lambda_2 1^T \Sigma^{-1}) \Sigma (\lambda_1 \Sigma^{-1} \mu + \lambda_2 \Sigma^{-1} 1) \\ &= \lambda_1^2 \mu^T \Sigma^{-1} \mu + \lambda_1 \lambda_2 \mu^T \Sigma^{-1} 1 + \lambda_1 \lambda_2 1^T \Sigma^{-1} \mu + \lambda_2^2 1^T \Sigma^{-1} 1 \\ &= \lambda_1 (\lambda_1 \mu^T \Sigma^{-1} + \lambda_2 1^T \Sigma^{-1}) \mu + \lambda_2 (\lambda_1 \mu^T \Sigma^{-1} + \lambda_2 1^T \Sigma^{-1}) 1 \\ &= \lambda_1 w^T \mu + \lambda_2 w^T 1 \\ &= \lambda_1 \mu_0 + \lambda_2 \\ &= \tfrac{1}{d}(q\mu_0 - r)\mu_0 + \tfrac{1}{d}(p - r\mu_0)\,. \end{aligned}$$

Rewriting the equation

$$d\sigma_w^2 = q\mu_0^2 - 2r\mu_0 + p = q(\mu_0 - \tfrac{r}{q})^2 + \underbrace{p - \tfrac{r^2}{q}}_{d/q}$$

as

$$\frac{\sigma_w^2}{\frac{1}{q}} - \frac{(\mu_0 - \frac{r}{q})^2}{\frac{d}{q^2}} = 1, \tag{6}$$

we see that it defines a hyperbola with center (0,r/q) and vertices $(\pm\sqrt{1/q}, r/q)$.

A portfolio implying a point above the vertex on the right branch of the hyperbola is called an efficient frontier portfolio because no other portfolio with the same expected return can have a smaller variance (see Figure 1). The minimum variance portfolio is that efficient frontier portfolio with the smallest variance (see Figure 1). The tangency portfolio is that efficient frontier portfolio which minimizes the Sharpe [15] ratio

$$\frac{\mu_w - r_f}{\sigma_w}, \tag{7}$$

where r_f is the (deterministic) return of the (hypothetical) risk-free asset (see Figure 1). If short selling is not allowed, the nonnegativity constraints

$$w_1, \ldots, w_K \geq 0 \tag{8}$$

must be imposed. Under these constraints, portfolio optimization is more difficult and requires the use of numerical methods.

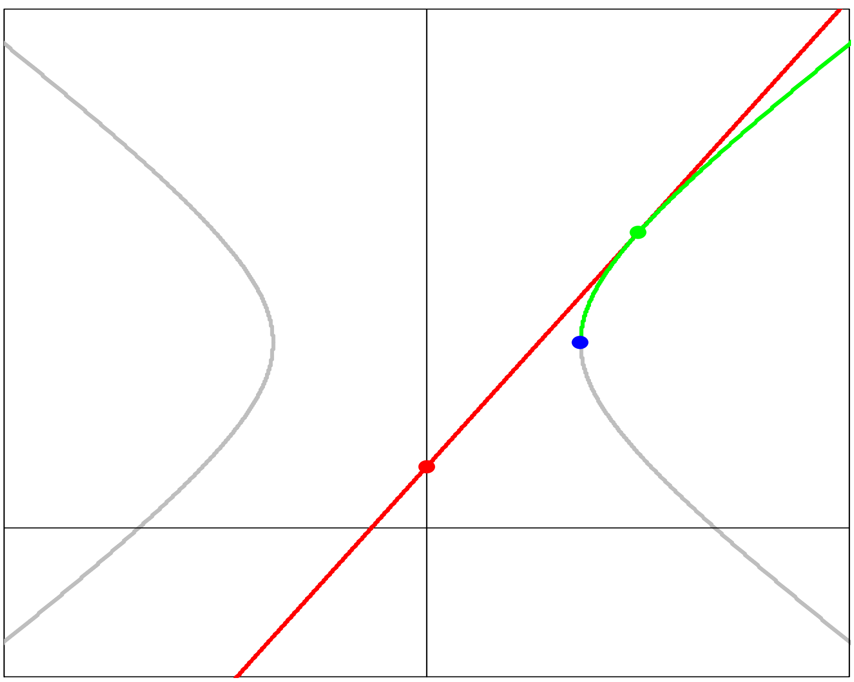

Figure 2: Markowitz hyperbola (6) in the σ–μ plane with efficient frontier (green line), tangency portfolio (green point), and minimum variance portfolio (blue point). The red line is the tangent to the hyperbola from the red point which has σ-coordinate 0 and μ-coordinate r_f.

2.2 The Capital Market Line

We consider portfolios that are mixtures of a market-portfolio with stochastic return R_m and a risk-free asset with interest rate r_f. The return of a portfolio with weights λ and 1-λ is given by

$$\underline{R}_\lambda = \lambda R_m + (1-\lambda) r_f \tag{9}$$

and its mean and variance by

$$\underline{\mu}_\lambda = E(\underline{R}_\lambda) = \lambda E(R_m) + (1-\lambda) r_f = \lambda \mu_m + (1-\lambda) r_f \tag{10}$$

and

$$\underline{\sigma}_\lambda^2 = \lambda^2 Var(R_m) = \lambda^2 \sigma_m^2, \tag{11}$$

respectively. It follows from

$$\lambda = \frac{\underline{\sigma}_\lambda}{\sigma_m}$$

that

$$\underline{\mu}_\lambda = r_f + \frac{\mu_m - r_f}{\sigma_m} \underline{\sigma}_\lambda = g(\underline{\sigma}_\lambda) \tag{12}$$

The slope of the capital market line g is the Sharpe ratio of the market-portfolio and its intercept is the risk-free interest rate.

2.3 The Capital Asset Pricing Model

The derivation of the CAPM relies on the critical assumption that for any given risk σ the highest possible expected return is $g(\sigma)$. We consider portfolios that are mixtures of a risky asset with stochastic return R_k and a market portfolio with stochastic return R_m. The mean and the variance of the return R_λ of a portfolio with weights λ and 1-λ are given by

$$\mu_\lambda = \lambda E(R_k) + (1-\lambda) E(R_m) = \mu_m + \lambda(\mu_k - \mu_m) \tag{13}$$

and

$$\begin{aligned}\sigma_\lambda^2 &= \lambda^2 Var(R_k) + 2\lambda(1-\lambda) Cov(R_k, R_m) + (1-\lambda)^2 Var(R_m) \\ &= \lambda^2 \sigma_k^2 + 2\lambda(1-\lambda)\sigma_{km} + (1-\lambda)^2 \sigma_m^2 \\ &= \lambda^2 \underbrace{(\sigma_k^2 - 2\sigma_{km} + \sigma_m^2)}_{A_k} + 2\lambda \underbrace{(\sigma_{km} - \sigma_m^2)}_{B_k} + \sigma_m^2 .\end{aligned} \tag{14}$$

The capital market line g intersects the function

$$\mu = f_1(\sigma) = \mu_m + \underbrace{\frac{-B_k + \sqrt{B_k^2 - A_k(\sigma_m^2 - \sigma^2)}}{A_k}}_{\lambda_1(\sigma)} (\mu_k - \mu_m), \quad \sigma \geq \sigma_m^2 - \frac{B_k^2}{A_k} \tag{15}$$

at the point (σ_m, μ_m) if $B_k \geq 0$ and the function

$$\mu = f_2(\sigma) = \mu_m + \underbrace{\frac{-B_k - \sqrt{B_k^2 - A_k(\sigma_m^2 - \sigma^2)}}{A_k}}_{\lambda_2(\sigma)}(\mu_k - \mu_m), \quad \sigma \geq \sigma_m^2 - \frac{B_k^2}{A_k} \tag{16}$$

if $B_k \leq 0$. The capital asset line g can only be a tangent to f_j at the point (σ_m,μ_m) if the respective function is increasing. The function f_1 is increasing if $\mu_k > \mu_m$ and f_2 is increasing if $\mu_k < \mu_m$ (see Figure 3). Thus, the critical assumption of the optimality of the capital market line already implies a positive relationship between B_k and $\mu_k - \mu_m$, i.e.,

$$B_k > 0 \Rightarrow \mu_k - \mu_m > 0, \quad B_k < 0 \Rightarrow \mu_k - \mu_m < 0$$

or, equivalently,

$$\beta_k = \frac{\sigma_{km}}{\sigma_m^2} > 1 \Rightarrow \mu_k - \mu_m > 0, \quad \beta_k < 1 \Rightarrow \mu_k - \mu_m < 0.$$

Equating the slopes of g and f_1 or f_2 at the point (σ_m,μ_m) shows that this relationship is linear. It follows from

$$\sigma^2 = A_k\lambda_j^2(\sigma) + 2B_k\lambda_j(\sigma) + \sigma_m^2,$$

$$2\sigma = 2A_k\lambda_j(\sigma)\lambda_j'(\sigma) + 2B_k\lambda_j'(\sigma),$$

and

$$2\sigma_m = 2A_k\underbrace{\lambda_j(\sigma_m)}_{=0}\lambda_j'(\sigma_m) + 2B_k\lambda_j'(\sigma_m) = 2(\sigma_{km} - \sigma_m^2)\lambda_j'(\sigma_m)$$

that

$$\frac{\mu_m - r_f}{\sigma_m} = f_j'(\sigma_m) = \lambda_j'(\sigma_m)(\mu_k - \mu_m) = \frac{\sigma_m}{\sigma_{km} - \sigma_m^2}(\mu_k - \mu_m)$$

and

$$(\beta_k - 1)(\mu_m - r_f) = \frac{\sigma_{km} - \sigma_m^2}{\sigma_m^2}(\mu_m - r_f) = \mu_k - \mu_m.$$

The CAPM is obtained by rewriting the last equation as

$$\beta_k(\mu_m - r_f) = \mu_k - r_f. \tag{17}$$

2.4 The "Broken" CAPM

Modifications of the CAPM which avoid the unrealistic assumption of borrowing at the risk-free rate have already been proposed in the early 1970s [3]. In contrast, Subsection 2.3 has implicitly assumed that borrowing at the risk-free rate is possible and shorting does not entail additional costs. Both assumptions can be avoided by using only proper weights, i.e., $0 \leq \lambda \leq 1$. If $\beta < 1$, the CAPM can still be obtained by equating the slope of g to the derivative of f_2 from the left (see Figure 3.c). If $\beta > 1$, the derivative of f_1 from the right must be used and the borrowing rate r_b must be used instead of the risk-free rate r_f in the definition of the capital market line (see Figure 3.a). Of course, the concept of a broken capital market line with different rates for lending and borrowing does not automatically compromise its optimality property because the expected returns must in the case of negative weights also be adjusted when the costs of shorting are taken into account. In practice, the "broken" CAPM

$$\mu_k = \begin{cases} \mu_k = r_f + \beta_k(\mu_m - r_f) & \text{if } \beta_k < 1 \\ \mu_k = r_b + \beta_k(\mu_m - r_b) & \text{if } \beta_k > 1 \end{cases} \tag{18}$$

will not make a big difference because the discrepancy between r_b and r_f is usually small compared to the size of μ_m.

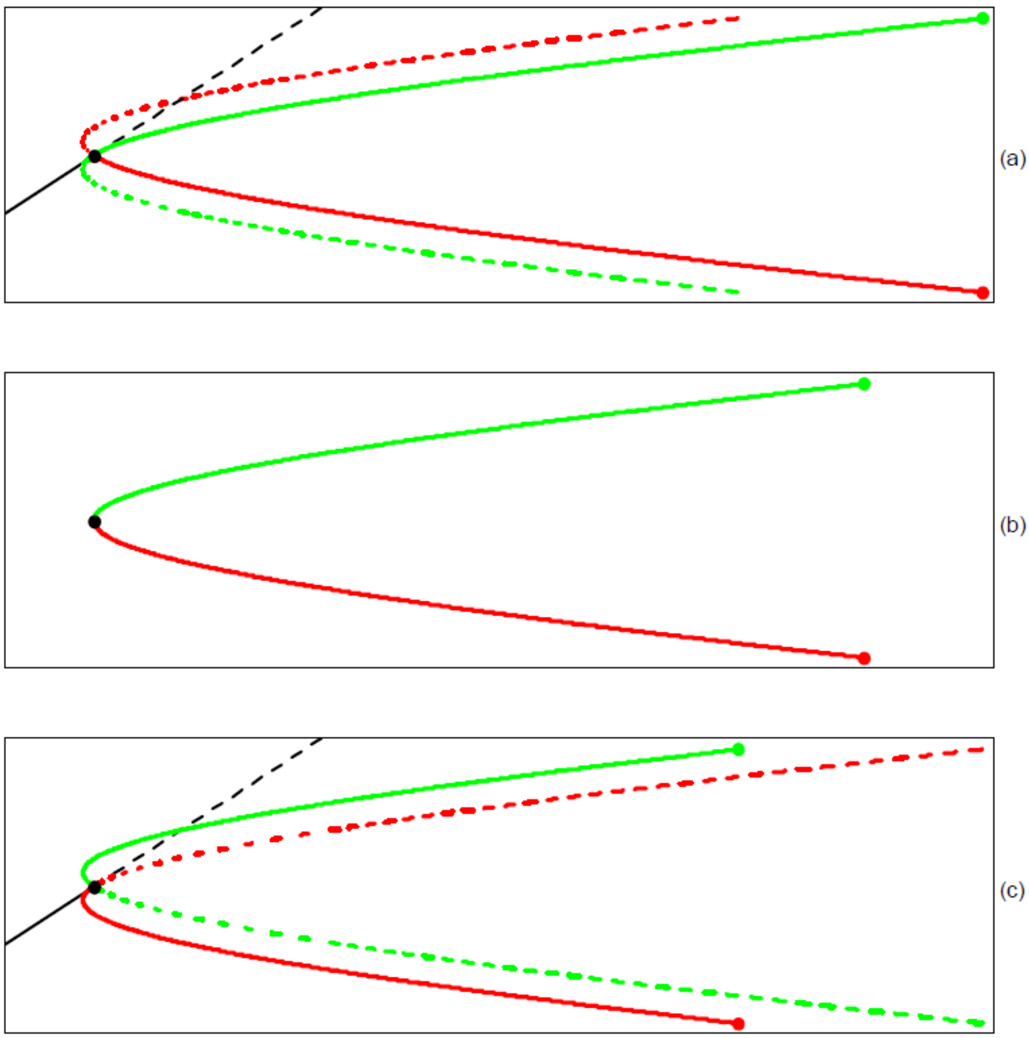

Figure 3:

(a) The capital asset line g is a tangent to f_1 at (σ_m,μ_m) if $B_k > 0$ and $\mu_k > \mu_m$.

(b) The capital asset line g cannot be a tangent to $f_1 = f_2$ at (σ_m,μ_m) if $B_k = 0$.

(c) The capital asset line g is a tangent to f_2 if $B_k < 0$ and $\mu_k < \mu_m$.

3 Empirical Results

Two sets of assets are used. The first set consists of those nine components of the Dow Jones Industrial Average the prices of which are available at Yahoo!Finance since January 2, 1962. The selected components are Alcoa (AA), Boeing (BA), Caterpillar (CAT), Du Pont (DD), Walt Disney (DIS), General Electric (GE), Hewlett-Packard (HPQ), IBM (IBM), and Coca-Cola (KO). The second set consists of nine major US indices the prices of which are available at Yahoo!Finance since June 4, 1996. The selected indices are Nasdaq Bank (^IXBK), Nasdaq Biotechnology (^NBI), Nasdaq Insurance (^IXIS), Nasdaq Telecommunications (^IXUT), Nasdaq Transportation (^IXTR), AMEX Gold Bugs (^HUI), AMEX Oil (^XOI), AMEX Pharmaceutical (^DRG), and PHLX Semiconductor (^SOX). For both sets, the sample period ends on May 31, 2013.

3.1 Empirical Evidence on the CAPM

Figure 4.a shows a plot of

$$\sum_{s=1}^{t} \hat{B}_k(s) = \sum_{s=1}^{t} \left(R_k(s) R_a(s) - R_a^2(s) \right)$$

against time $t = 1, ..., n$, where $R_k(s)$ is the return of the kth stock at time s and $R_a(s)$ is the return of the equally weighted portfolio at time s. The average of the K=9 DJIA components is used instead of a broad market index in order to avoid a survival bias. This will not be necessary when the set of indices is analyzed. In the latter case, it makes more sense to use the S&P 500 as a proxy for the market. Overall, there seems to be no obvious relationship between the return R_k (Figure 1.a) and the parameter B_k (Figure 4.a). However, using the sign of $\hat{B}_k(s)$ to switch between $R_k(s)$ and $R_a(s)$ is an apparently successful strategy. For each k, the switching strategy outperforms both the average (Figure 4.b) and the respective stock (Figure 4.c). Additional strong evidence in favor of such a relationship is obtained when all stocks are used simultaneously. The strategy which always switches to the stock with the kth largest value has just the kth best performance (Figure 6.a).

Because of the relative stability of $\hat{B}_k$, it might be expected that past values are also related to present and future returns. Figure 5 and Figures 6.b-c show that this is definitely not the case. The out-of-sample performance of switching strategies based on sums of past returns is extremely poor. The explanation for this apparent paradox is purely statistical. Suppose that R_k and R_a are positively correlated random variables with practically identical means but different variances. If positive returns occur more frequently than negative returns, the conditional means of R_k and R_a given $R_k R_a$<0 will be smaller than the unconditional means and probably be negative. It is then safer to choose the asset with the smaller variance. Figure 7 shows that the conditional mean of R_k given $R_k R_a$<0 is indeed negative for each k. Clearly, this relationship holds only for a fixed pair of random variables and is therefore of no use for the prediction of subsequent returns. Because of its technical nature it has absolutely no economic relevance and must be interpreted with great caution.

However, the use of the more robust statistics:

$$\widehat{B}_k(t) = |R_k(t)|\, sign(R_k(t)R_a(t)) - |R_a(t)|$$

or

$$\widetilde{B}_k(t) = |R_k(t) - R_a(t)|$$

instead of $\hat{B}_k(t)$ brings a spark of hope. Overall, strategies which invest in stocks with larger estimates appear to outperform the equally weighted portfolio out-of-sample (Figure 8). However, the evidence is not very strong and is practically non-existent in extended subperiods. Moreover, an analogous analysis of the second data set yields a disappointing result (Figure 14). In general, the evidence obtained from the indices is less conclusive (Figures 9-14) than that obtained from the stocks (Figures 1, 4-8). This may be due to the smaller sample size as well as to the presence of indices with divergent behavior (Figure 9.a). A further important difference is that the S&P500 is used as a proxy for the market instead of the average of the indices.

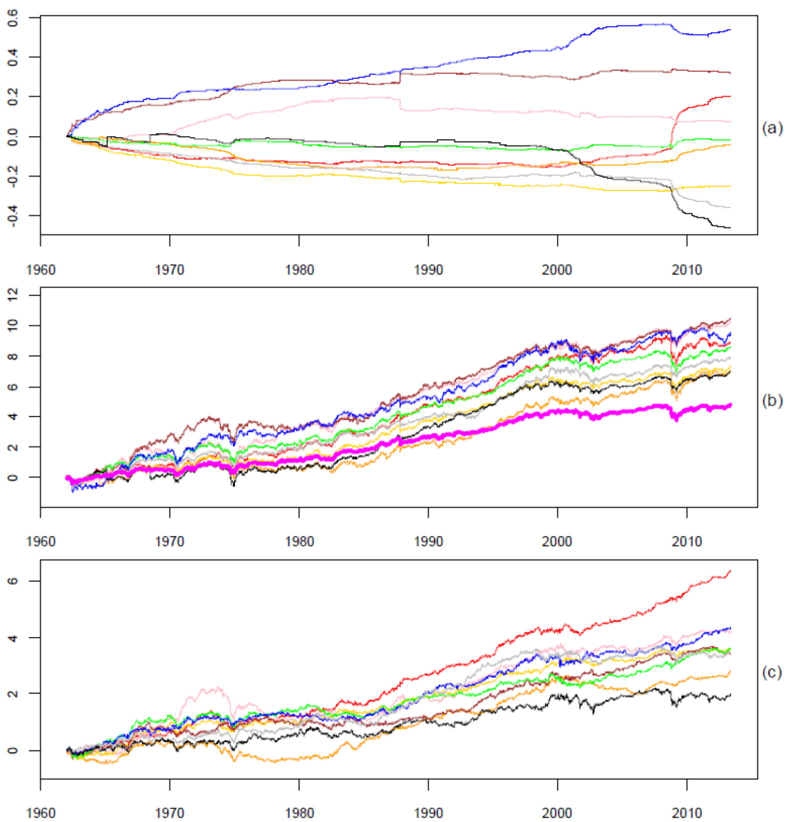

Figure 4: In-sample dependence of R_k on $\hat{B}_k$

(a) Cumulative sums of $\hat{B}_k$

(b) Performance of switching between R_k and R_a based on the sign of $\hat{B}_k$

(c) Performance of switching between R_k and R_a relative to R_k

Data: AA (red), BA (pink), CAT (orange), DD (gold), DIS (brown), GE (green), HPQ (blue), IBM (gray), KO (black), Average (magenta)

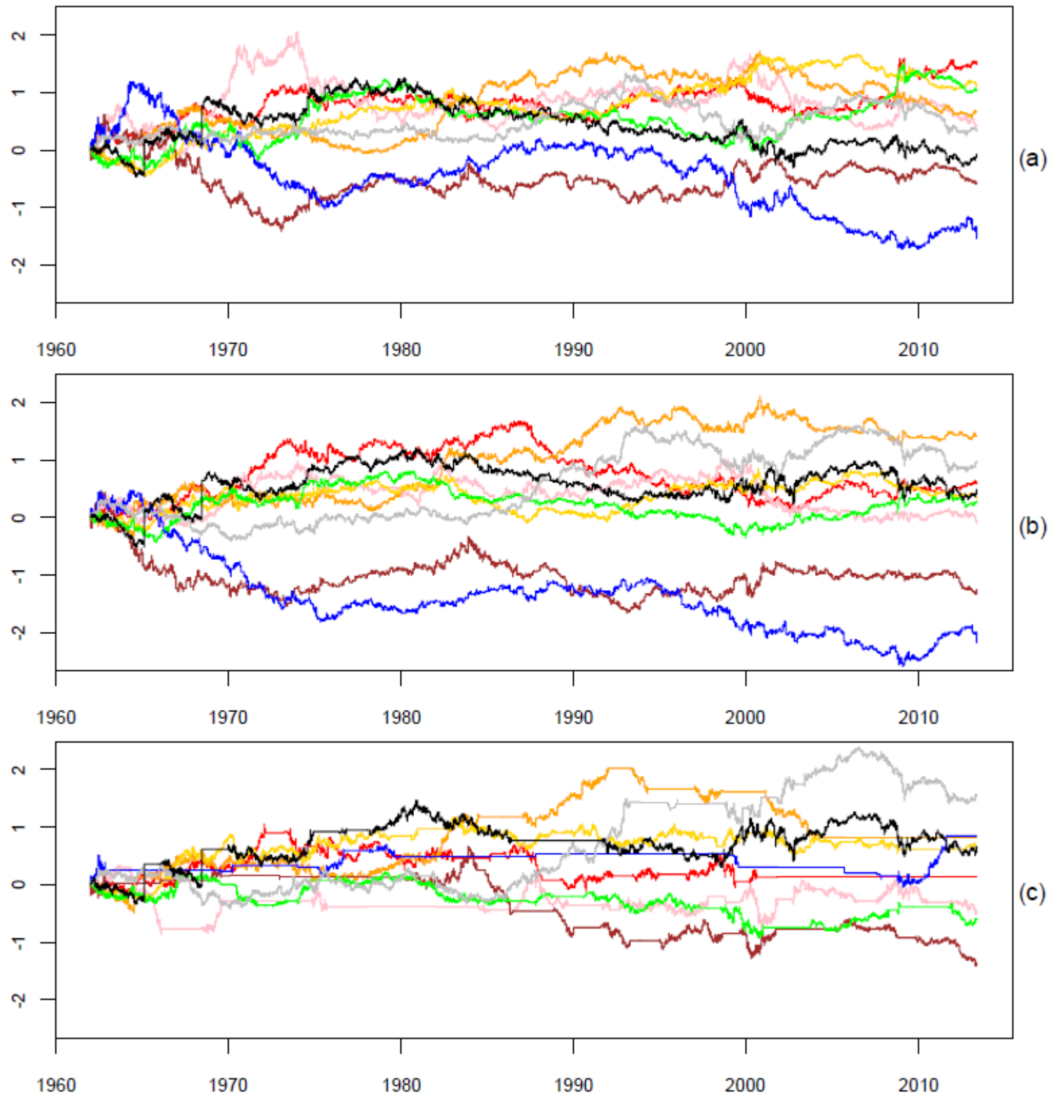

Figure 5: Out-of-sample performance relative to that of average (equally weighted portfolio)

(a) Switching between R_k and R_a based on sign of $\hat{B}_k(t-1)$

(b) Switching between R_k and R_a based on sign of $\hat{B}_k(t-5)+...+\hat{B}_k(t-1)$

(c) Switching between R_k and R_a based on sign of $\hat{B}_k(t-250)+...+\hat{B}_k(t-1)$

Data: AA (red), BA (pink), CAT (orange), DD (gold), DIS (brown), GE (green), HPQ (blue), IBM (gray), KO (black)

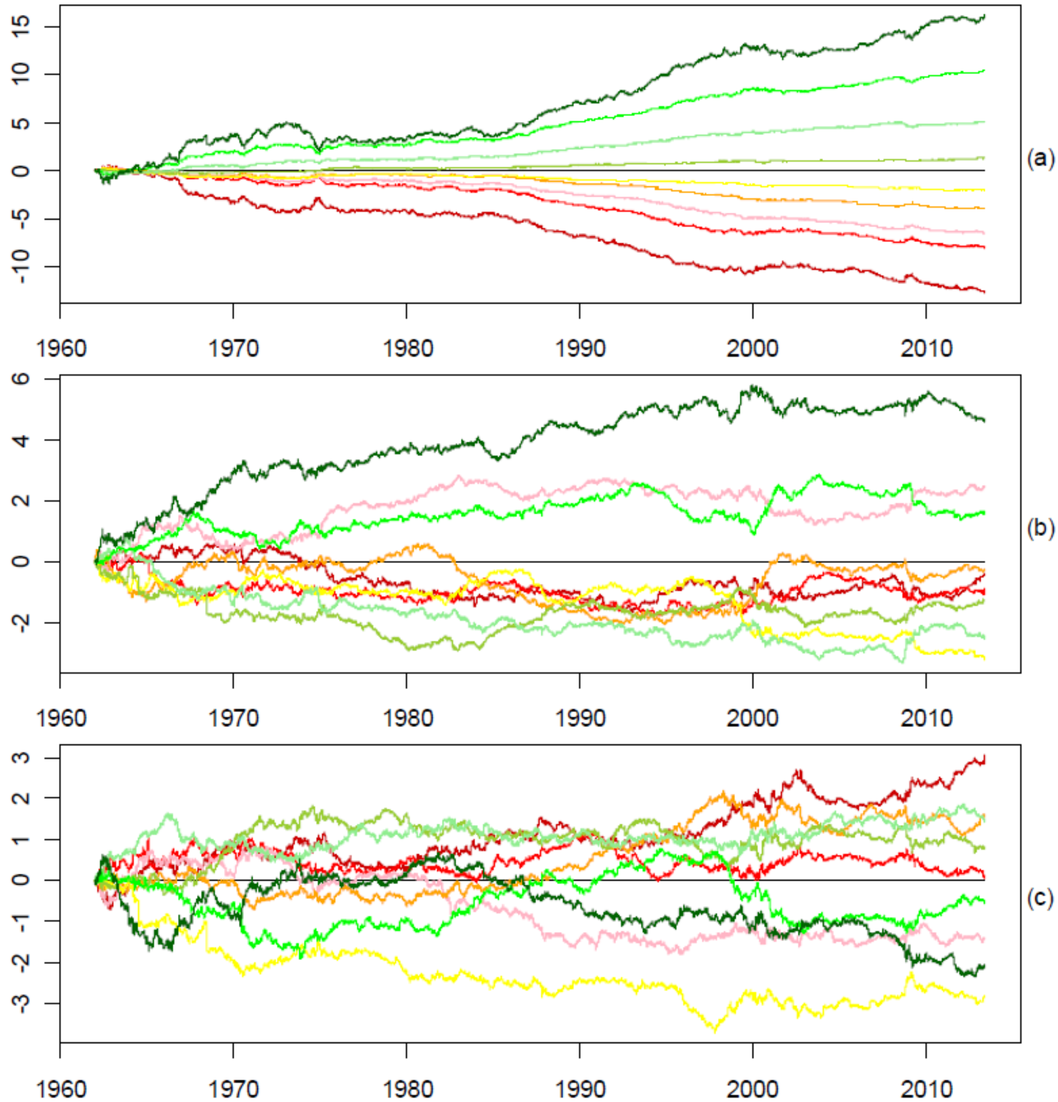

Figure 6: In-sample (a) and out-of-sample (b-c) performance relative to that of average (a) Switching to R_k if $\hat{B}_k$ is largest (darkgreen), 2nd largest (green), 3rd (lightgreen), 4th (yellowgreen), 5th (yellow), 6th (orange), 7th (pink), 8th (red), 9th (darkred) ... (b) Switching to R_k if $\hat{B}_k(t-1)$ is largest (darkgreen), ... (c) Switching to R_k if $\hat{B}_k(t-5)+...+\hat{B}_k(t-1)$ is largest (darkgreen),...

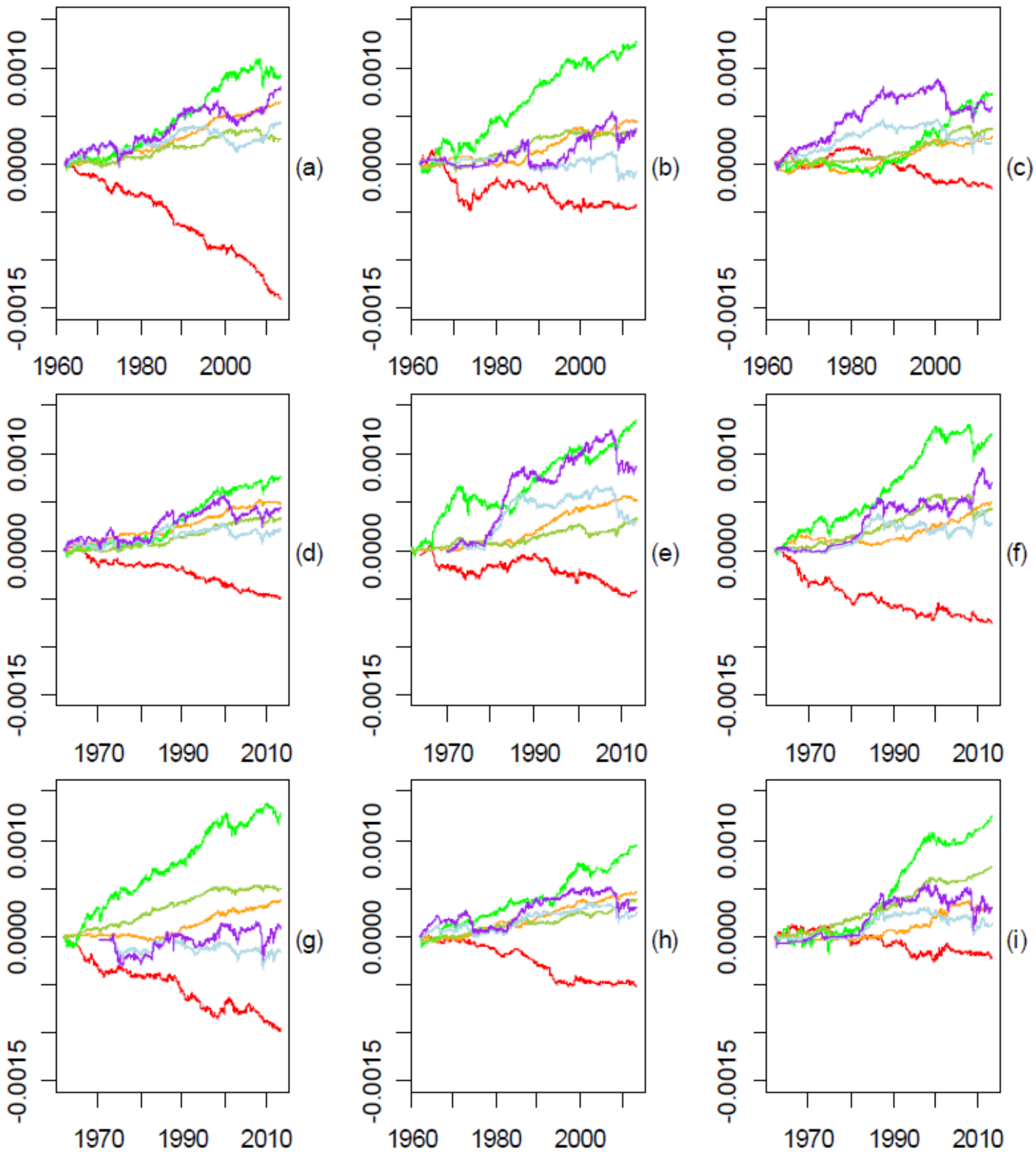

Figure 7: Conditional mean of R_k given that $R_k R_a < 0$ (red), $R_k R_a > 0 \wedge |R_k| > |R_a|$ (green), $R_k R_a > 0 \wedge |R_k| < |R_a|$ (lightblue), conditional mean of R_a given that $R_k R_a < 0$ (orange), $R_k R_a > 0 \wedge |R_k| > |R_a|$ (yellowgreen), $R_k R_a > 0 \wedge |R_k| < |R_a|$ (purple)

Data: AA (a), BA (b), CAT (c), DD (d), DIS (e), GE (f), HPQ (g), IBM (h), KO (i)

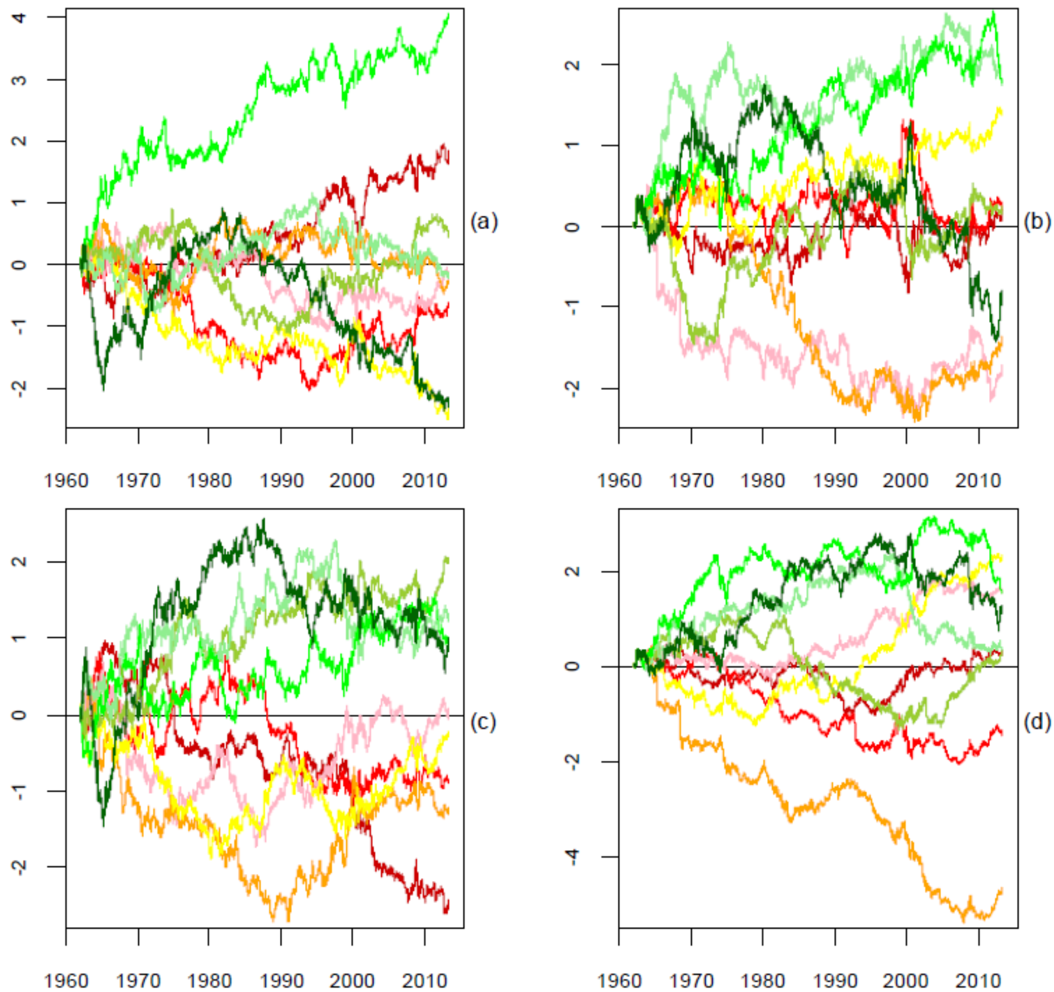

Figure 8: Relative out-of-sample performance of robust switching between DJIA components: (a) Switching to R_k if $\widehat{B}_k(t-5)+...+\widehat{B}_k(t-1)$ is largest (darkgreen), 2nd largest (green), 3rd (lightgreen), 4th (yellowgreen), 5th (yellow), 6th (orange), 7th (pink), 8th (red), 9th (darkred) (b) Switching to R_k if $\widehat{B}_k(t-250)+...+\widehat{B}_k(t-1)$ is largest (darkgreen), ... (c) Switching to R_k if $\widetilde{B}_k(t-5)+...+\widetilde{B}_k(t-1)$ is largest (darkgreen), ... (d) Switching to R_k if $\widetilde{B}_k(t-250)+...+\widetilde{B}_k(t-1)$ is largest (darkgreen), ...

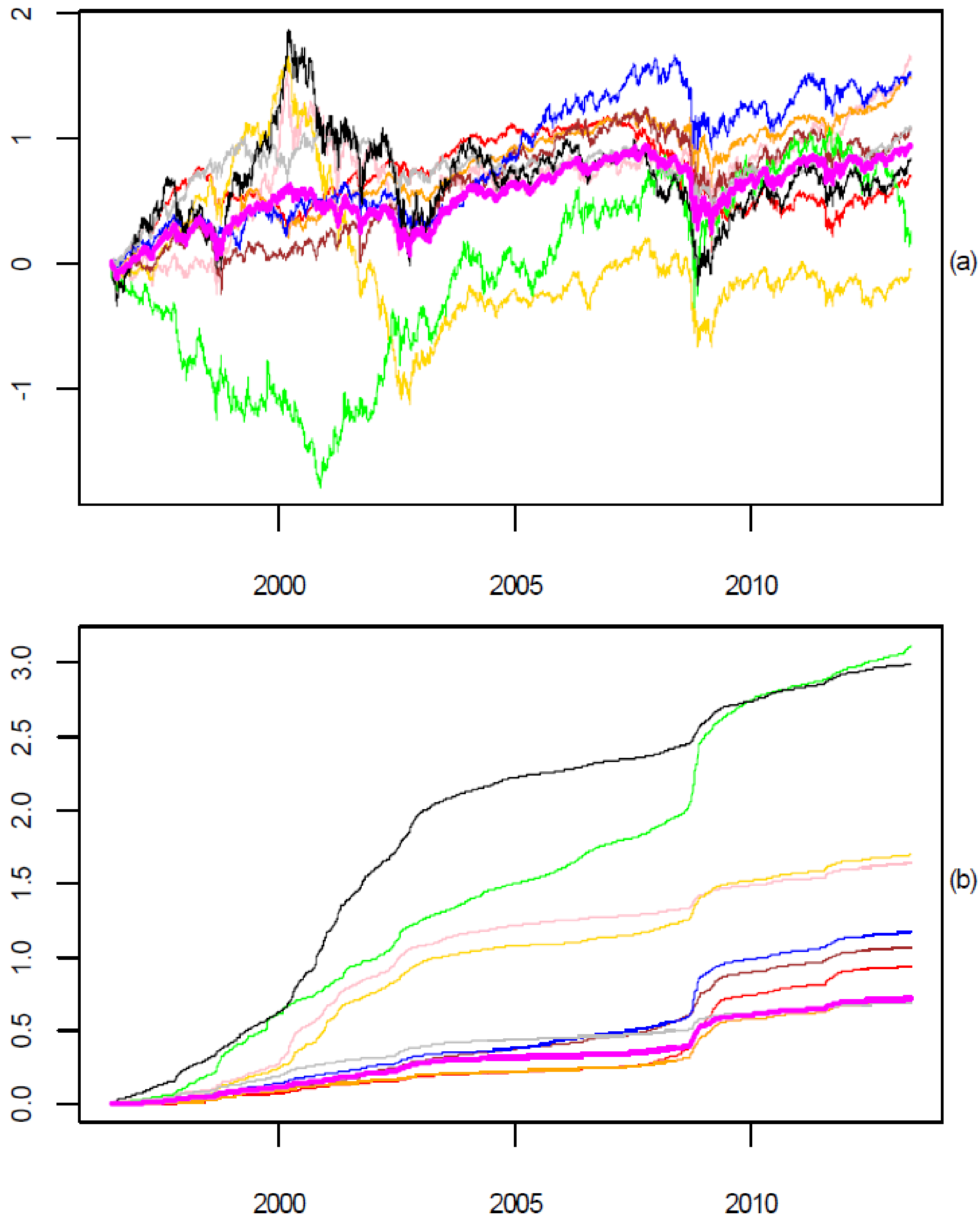

Figure 9: Comparison of cumulative returns (a) and cumulative squared returns (b) of nine major US indices with the cumulative S&P500 returns (a) and cumulative squared S&P500 returns (b) from 04-06-1996 to 31-05-2013.
Data: ^IXBK (red), ^NBI (pink), ^IXIS (orange), ^IXUT (gold), ^IXTR (brown), ^HUI (green), ^XOI (blue), ^DRG (gray), ^SOX (black), S&P500 (magenta)

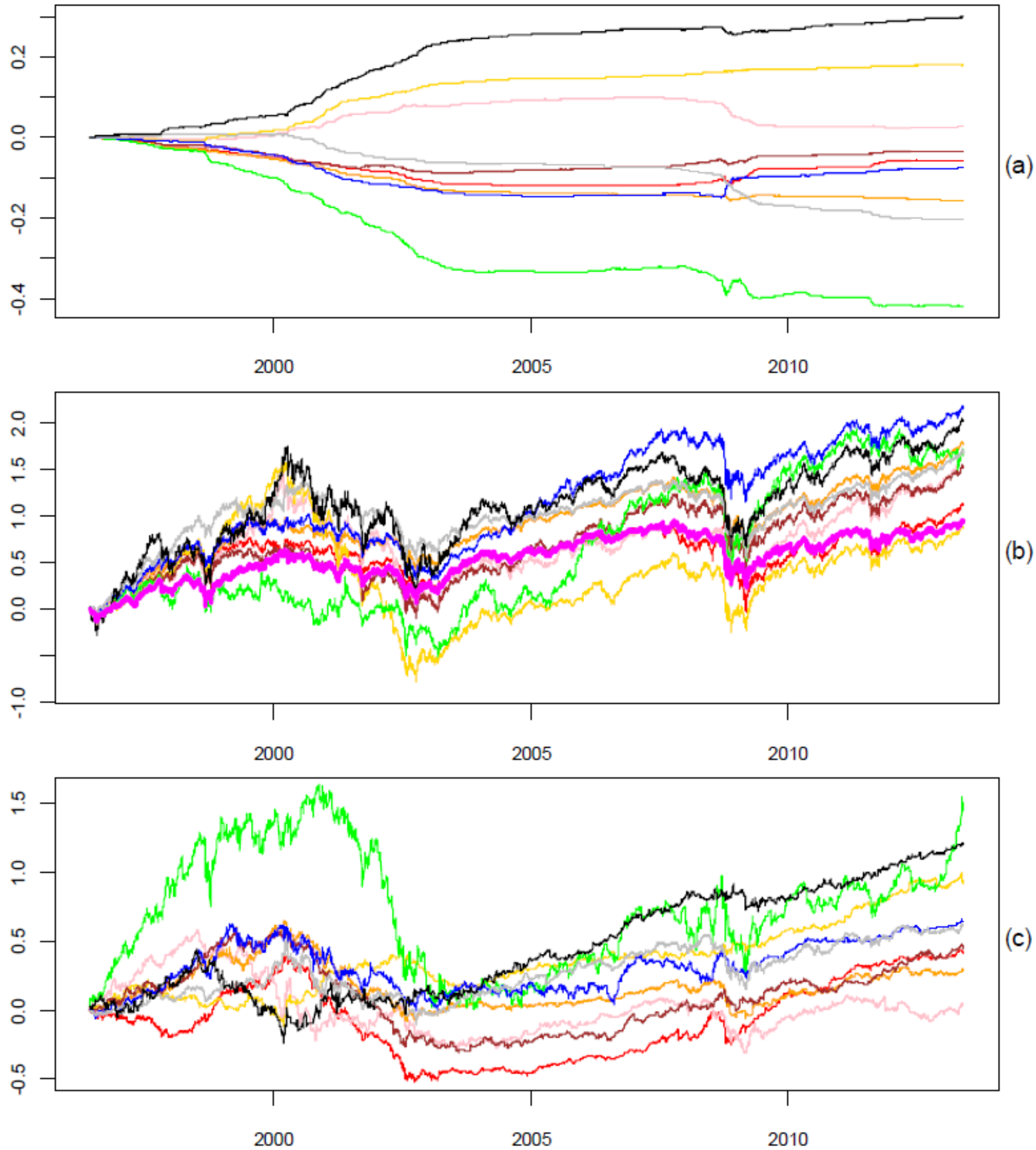

Figure 10: In-sample dependence of R_k on $\hat{B}_k$

(a) Cumulative sums of $\hat{B}_k$

(b) Performance of switching between R_k and R_m based on the sign of $\hat{B}_k$

(c) Performance of switching between R_k and R_m relative to R_k

Data: ^IXBK (red), ^NBI (pink), ^IXIS (orange), ^IXUT (gold), ^IXTR (brown), ^HUI (green), ^XOI (blue), ^DRG (gray), ^SOX (black), S&P500 (magenta)

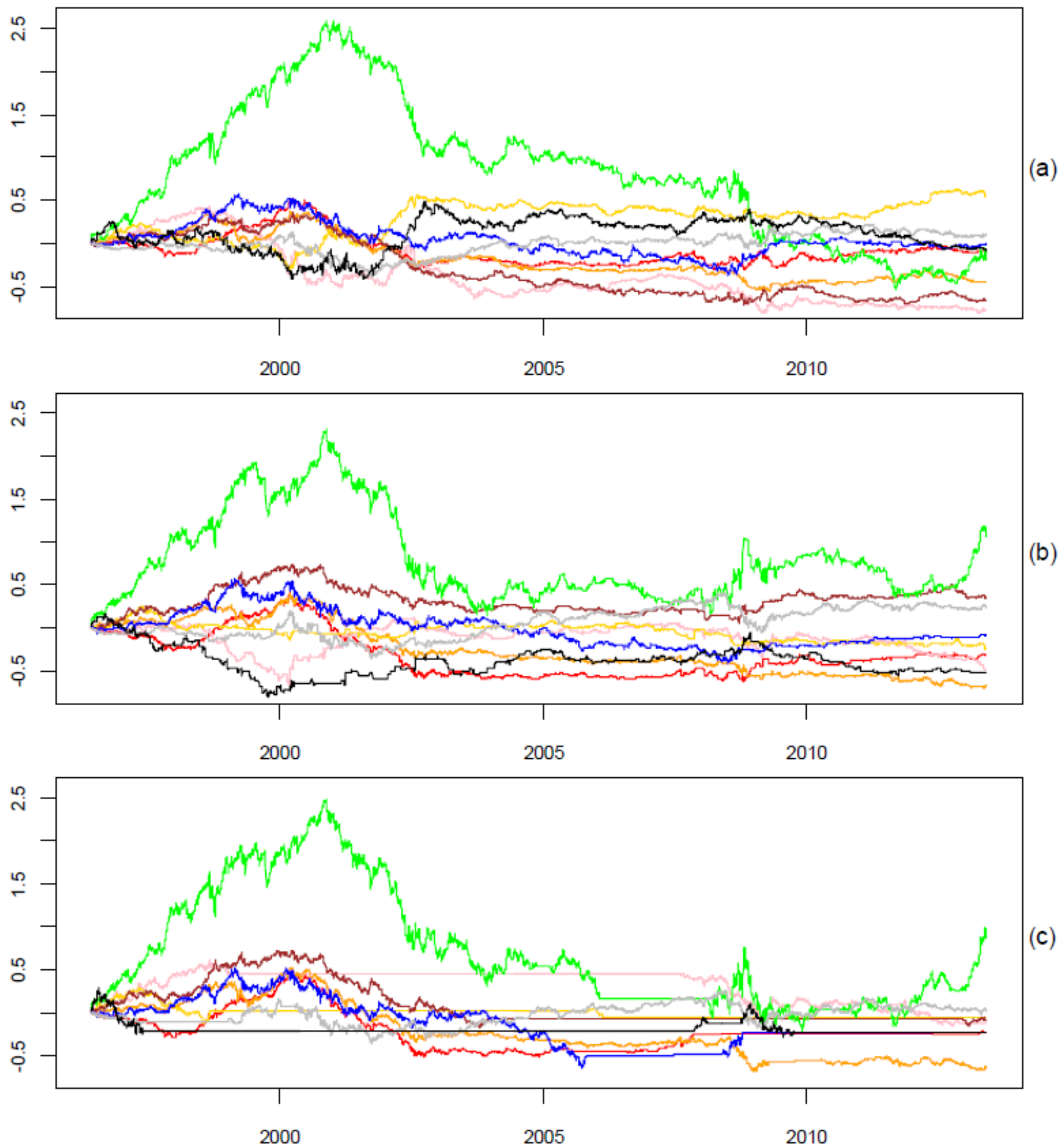

Figure 11: Out-of-sample performance relative to that of S&P500

(a) Switching between R_k and R_m based on sign of $\hat{B}_k(t-1)$

(b) Switching between R_k and R_m based on sign of $\hat{B}_k(t-5)+...+\hat{B}_k(t-1)$

(c) Switching between R_k and R_m based on sign of $\hat{B}_k(t-250)+...+\hat{B}_k(t-1)$

Data: ^IXBK (red), ^NBI (pink), ^IXIS (orange), ^IXUT (gold), ^IXTR (brown), ^HUI (green), ^XOI (blue), ^DRG (gray), ^SOX (black)

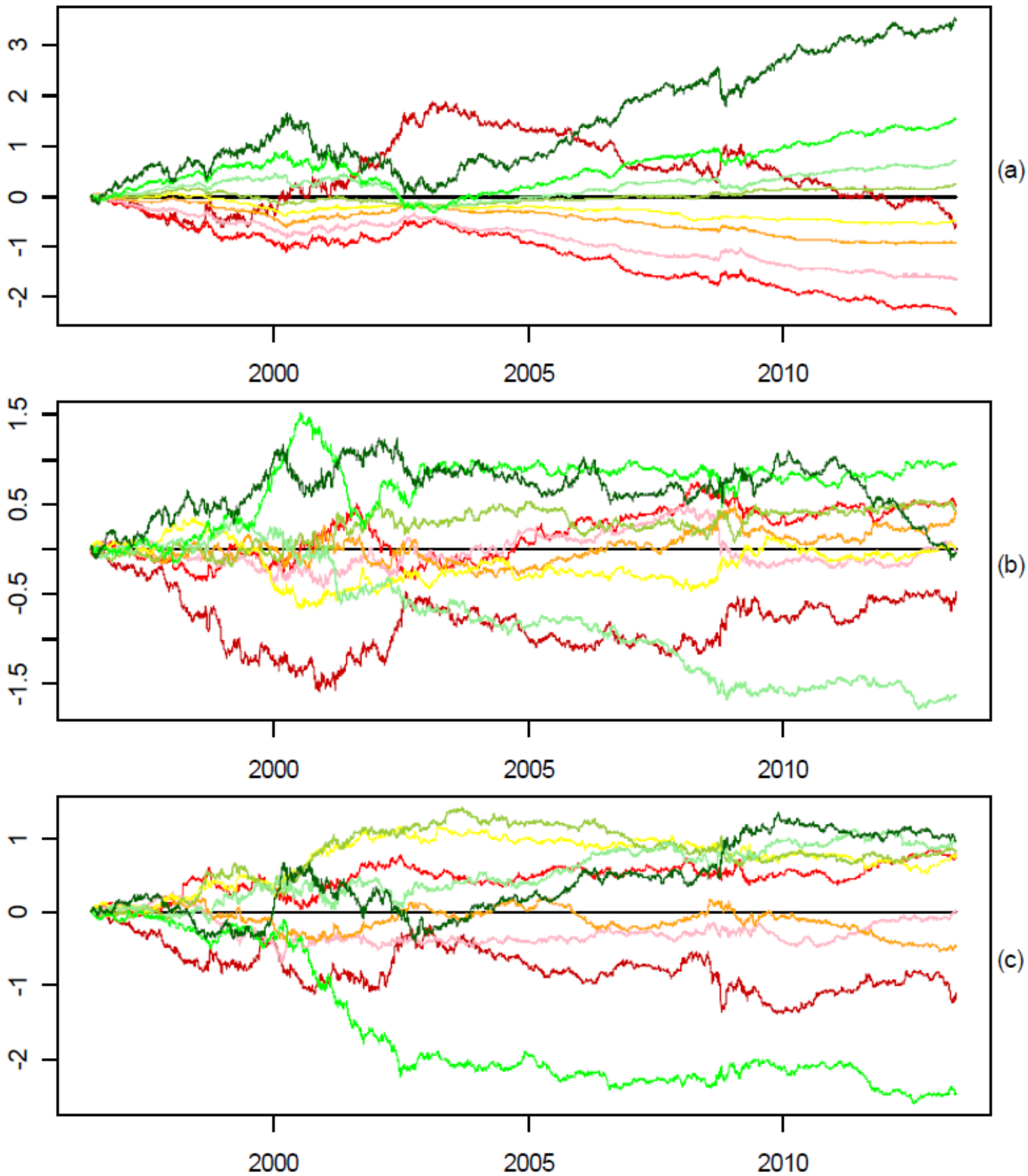

Figure 12: In-sample (a) and out-of-sample (b-c) performance relative to that of S&P500 (a) Switching to R_k if $\hat{B}_k$ is largest (darkgreen), second largest (green), 3rd (lightgreen), 4th (yellowgreen), 5th (yellow), 6th (orange), 7th (pink), 8th (red), 9th (darkred) (b) Switching to R_k if $\hat{B}_k(t-1)$ is largest (darkgreen), ... (c) Switching to R_k if $\hat{B}_k(t-5)+...+\hat{B}_k(t-1)$ is largest (darkgreen), ...

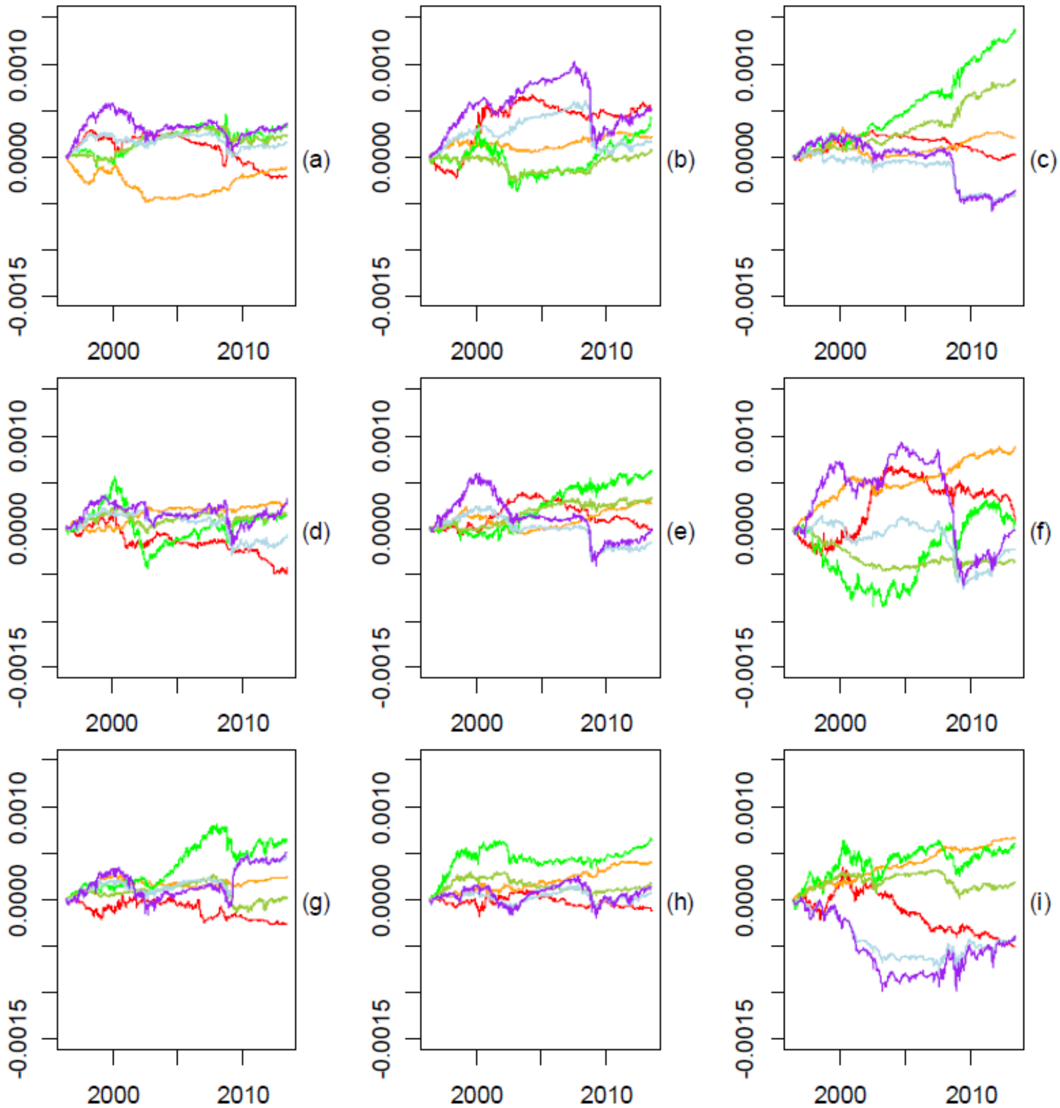

Figure 13: Conditional mean of R_k given that $R_k R_a < 0$ (red), $R_k R_a > 0 \wedge |R_k| > |R_a|$ (green), $R_k R_a > 0 \wedge |R_k| < |R_a|$ (lightblue), conditional mean of R_a given that $R_k R_a < 0$ (orange), $R_k R_a > 0 \wedge |R_k| > |R_a|$ (yellowgreen), $R_k R_a > 0 \wedge |R_k| < |R_a|$ (purple)

Data: ^IXBK (a), ^NBI (b), ^IXIS (c), ^IXUT (d), ^IXTR (e), ^HUI (f), ^XOI (g), ^DRG (h), ^SOX (i)

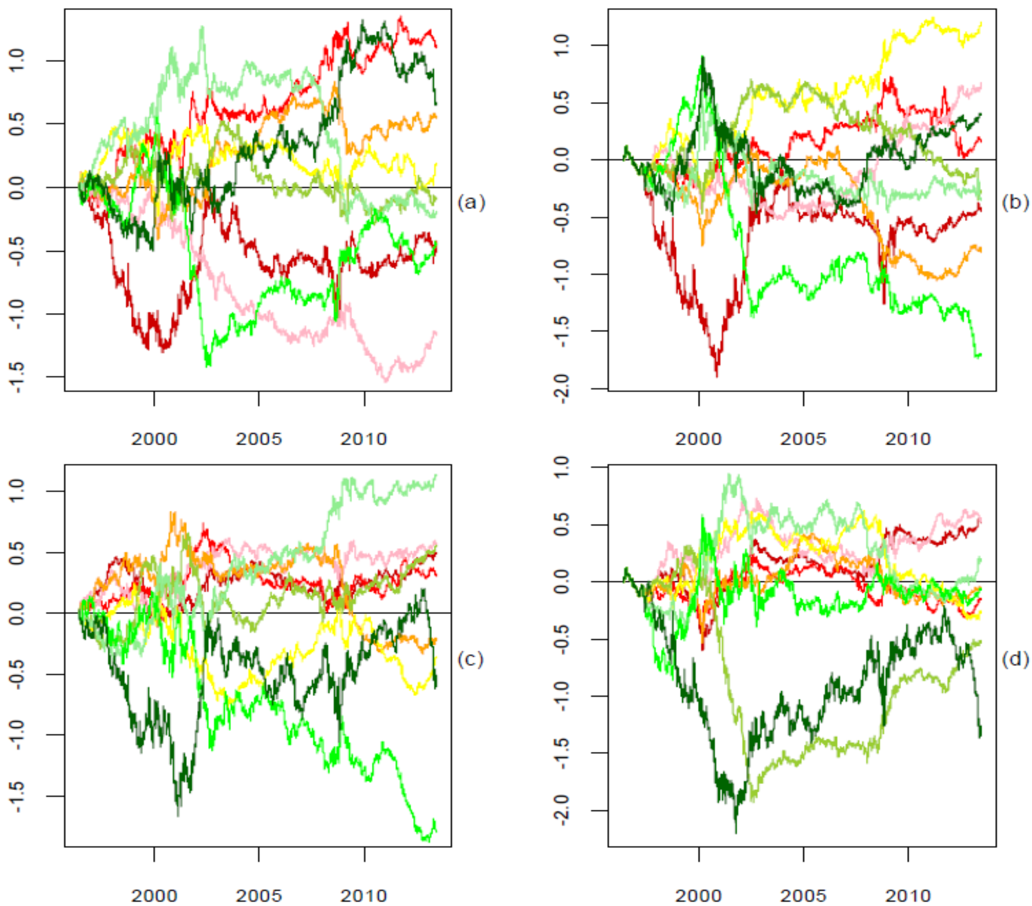

Figure 14: Relative out-of-sample performance of robust switching between US indices (a) Switching to R_k if $\widehat{B}_k(t-5)+...+\widehat{B}_k(t-1)$ is largest (darkgreen), 2nd largest (green), 3rd (lightgreen), 4th (yellowgreen), 5th (yellow), 6th (orange), 7th (pink), 8th (red), 9th (darkred) (b) Switching to R_k if $\widehat{B}_k(t-250)+...+\widehat{B}_k(t-1)$ is largest (darkgreen), ... (c) Switching to R_k if $\widetilde{B}_k(t-5)+...+\widetilde{B}_k(t-1)$ is largest (darkgreen), ... (d) Switching to R_k if $\widetilde{B}_k(t-250)+...+\widetilde{B}_k(t-1)$ is largest (darkgreen), ...

3.2 Empirical Evidence on the Performance of Portfolio Optimization Procedures

Because of the failure of the CAPM to provide suitable forecasts of future returns, simple historical means are used instead. In a rolling analysis, the returns of the last 200 trading days are used as input for the R/Rmetrics "fPortfolio" package [18] to find the optimal portfolio weights (long only) for the next trading day. The risk-free rate is either set to zero (r_f=0) or obtained from the 13-week treasury bill rate (r_f=r_{13}). Figure 15.a shows the relative performance of the minimum variance portfolio as well as the tangency portfolios with risk-free rates r_f=0 and r_f=r_{13}, respectively. Only the latter tangency portfolio can outperform the equally weighted portfolio over an extended period of time. However, its miserable performance in the last decades wrecks all hopes. In another try to corroborate the theoretical results, the portfolio variance is minimized for different return targets (first, second and third quartile of historical 200-day means). Contrary to expectations, the lowest target return yields the highest return and the highest target return yields the lowest return (Figure 15.b). Finally, in a last attempt, the portfolio-optimization procedure is robustified by using alternative covariance estimators. Unfortunately, the results do not get any better. They are still the wrong way round (Figure 15.c). The fact that the observed performance differences are possibly not even significant is only small comfort. Similarly, there is also no indication that the performance depends on the specification of the target return in the case of the second data set (Figures 16.b-c).

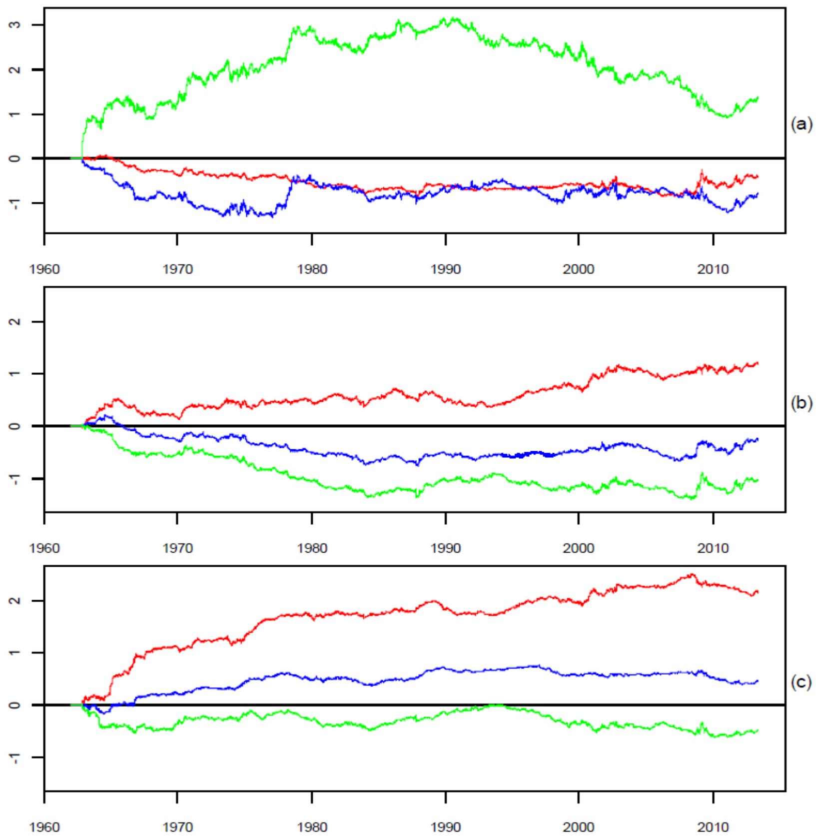

Figure 15: Relative out-of-sample performance of optimized portfolios of DJIA components (a) Minimum variance portfolio (red), tangency portfolios with r_f=0 (blue) and r_f=r_{13} (green) (b) Minimum variance for given low (red), medium (blue) and high (green) target return (c) Minimum variance for given low (red), medium (blue) and high (green) target return based on Spearman's rank estimator

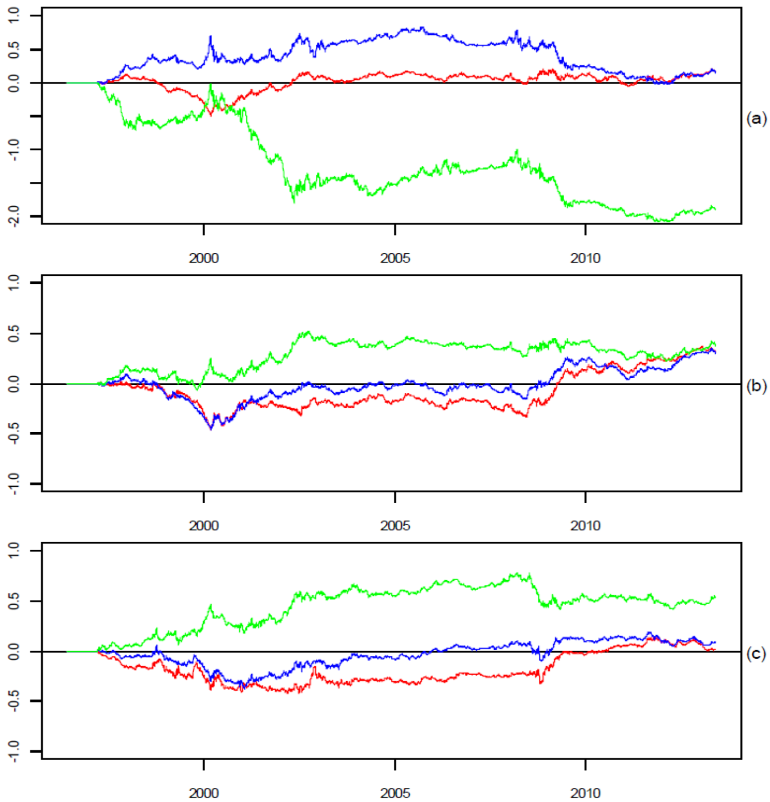

Figure 16: Relative out-of-sample performance of optimized portfolios of major US indices (a) Minimum variance portfolio (red), tangency portfolios with $r_f=0$ (blue) and $r_f=r_{13}$ (green) (b) Minimum variance for given low (red), medium (blue) and high (green) target return (c) Minimum variance for given low (red), medium (blue) and high (green) target return based on Spearman's rank estimator

4 Conclusion

Mean–variance portfolio optimization and the CAPM are the pillars of classical finance. Both focus on the first and second moments of the returns of financial assets. Despite the usually large sample sizes in financial applications, it is practically impossible to obtain precise estimates of these quantities because they change over time. Naturally, the non-normality of the returns seriously complicates the estimation of the variances and covariances. In the case of the means, it is their smallness (compared to the size of the variances) that makes the estimation so difficult. While the former problem can possibly be overcome by using robust estimation methods, additional information is required for the latter one. Thus, the validity of the CAPM which claims that there is a linear relationship between the expected return of an asset and its covariance with the market portfolio is not only important for its own sake but is also of vital importance for the practical implementation of portfolio-optimization procedures.

On the one hand, the discussion in 2.4 shows that some unrealistic assumptions usually required for the derivation of the CAPM can be relaxed and, on the other hand, it is argued in 3.1 that a relationship of the CAPM-type can as well be explained by statistical effects alone. The empirical results obtained from two sets of assets consisting of individual stocks and indices, respectively, suggest that the CAPM is either wrong or of no practical value (not even when robust methods are used and transaction costs are disregarded), and, in addition, that portfolio optimization is of little use for managing the trade-off between risk and return. Only the trivial task of reducing the risk appears to be doable.

Of course, nicer results could easily be obtained by fiddling around with different assets, time periods, estimation methods, parameters (such as risk-free rates and return targets) and - most conveniently - sampling frequencies (anything can be "proven" with monthly data). However, the evidence obtained in this way would not be of any significance.

References

[1] N. Barberis and R. Thaler, A survey of behavioral finance, in G. Constantinides, M. Harris and R. Stulz (editors) *Handbook of the Economics of Finance* **1B**, Elsevier, Amsterdam, 2003, 1053–1128.

[2] V. Bawa and E. Lindenberg, Capital market equilibrium in a mean - lower partial moment framework, *Journal of Financial Economics*, **5**, (1977), 189-200.

[3] F. Black, Capital market equilibrium with restricted borrowing, *The Journal of Business*, **45**, (1972), 444-455.

[4] J. Estrada, The Cost of Equity in Emerging Markets: A Downside Risk Approach, *Emerging Markets Quarterly*, (Fall 2000), 19-30.

[5] F.J. Fabozzi, D. Huang and G. Zhou, Robust portfolios: contributions from operations research and finance, *Annals of Operations Research*, **176**, (2010), 191–220.

[6] E.F. Fama and K.R. French, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics*, **33**, (1993), 3–56.

[7] E.F. Fama and K.R. French, Multifactor explanations of asset pricing anomalies, *The Journal of Finance*, **51**, (1996), 55-84.

[8] E.F. Fama and K.R. French, The capital asset pricing model: theory and evidence, *Journal of Economic Perspectives*, **18**, (2004), 25-46.

[9] V. Harlow and R. Rao, Asset pricing in a generalized mean - lower partial moment framework: theory and evidence, *Journal of Financial and Quantitative Analysis*, **24**, (1989), 285-311.

[10] W. Hogan and J. Warren, Toward the development of an equilibrium capital-market model based on semivariance, *Journal of Financial and Quantitative Analysis*, **9**, (1974), 1-11.

[11] L. Huo, T-H. Kim and Y. Kim, Robust estimation of covariance and its application to portfolio optimization, *Finance Research Letters*, **9**, (2012), 121-134.

[12] J. Lintner, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, The *Review of Economics and Statistics*, **47**, (1965), 13–37.

[13] H.M. Markowitz, Portfolio Selection, *The Journal of Finance*, **7**, (1952), 77-91.

[14] H.M. Markowitz, *Portfolio selection: efficient diversification of investments*, John Wiley & Sons, New York, 1959.

[15] W.F. Sharpe, Capital asset prices: a theory of market equilibrium under conditions of risk, *The Journal of Finance*, **19**, (1964), 425–442.

[16] T. Tsagaris, A. Jasra and N. Adams, Robust and adaptive algorithms for online portfolio selection, *Quantitative Finance*, **12**, (2012), 1651-1662.
[17] R.E. Welsch and X. Zhou, Application of robust statistics to asset allocation models, *REVSTAT – Statistical Journal*, **5**, (2007), 97–114.
[18] D. Würtz, Y. Chalabi, W. Chen and A. Ellis, *Portfolio Optimization with R/Rmetrics*, Rmetrics eBook, Rmetrics Association and Finance Online, Zurich, 2009.