International Journal of Business and Management Study - IJBMS
Volume 3 : Issue 1 [ISSN 2372-3955]
Publication Date : 18 April, 2016

Naïve and Rational Investments?

The Efficiency of Equal Weighted Indices

Guido Abate

***Abstract*—This study is aimed at finding an empirical evaluation of the rationality of naïve diversification when applied to indexed investments, linked to the most representative benchmarks of the US and EMU stock markets. An investor, in the MPT and CAPM frameworks, is assumed to be rational when he/she chooses the most efficient portfolio. The empirical study employs four measures of efficiency: the Gibbons, Ross and Shanken test; the Wald test, implemented through a bootstrap simulation; the Generalized Method of Moments test, implemented through a block bootstrap simulation; and Kandel and Stambaugh's relative efficiency measure. Results provide strong evidence of the superior efficiency of equal weighting if compared to float- and capitalization- weighting. As a consequence, these findings suggest that naïve diversification can be regarded as rational behavior for indexed investors.**

***Keywords*—naïve diversification; 1/N heuristic; diversification heuristic; portfolio efficiency; indexed investing; equal weighting.**

I. Introduction

One of the several heuristics typical of human behavior is "naïve" or "1/N" diversification, i.e. the choice of equal weight investment strategies. This asset allocation technique consists of the subdivision of the investor's wealth into equal fractions, each one invested into a different asset. The apparent simplicity of this asset allocation strategy has led the financial community to label it "naïve diversification", a definition which implies a judgment of limited, if not absent, rationality.

Markowitz [1], enunciating the Modern Portfolio Theory (MPT), has postulated that rational investors should diversify their investments and, at the same time, maximize their utility function, a function which is directly proportional to the expected return of the portfolio and inversely proportional to the variance of returns. Portfolios with the least variance, given a certain expected return, are defined as "efficient portfolios" and lie on the "efficient frontier". Risky assets are always present with a positive weight in efficient portfolios: long investments are the only ones allowed, in accordance with the limitations imposed by real-life regulations, at least in the retail segment of the financial market. Every investor should choose an asset allocation equal to that of an efficient portfolio, which coincides with the tangency point between his/her utility function and the efficient frontier on the mean-standard deviation plane. MPT is, then, a prescriptive theory, according to which investors, in order to behave rationally, are assumed to follow its advice about mean-variance allocation.

Guido Abate
University of Brescia
Italy

Following this approach, Sharpe [2], Lintner [3] and Mossin [4] have developed the Capital Asset Pricing Model (CAPM), a positive theory that is assumed to be a realistic description of investors' behavior. According to the CAPM, if its underlying hypotheses are met, all the investors, who are assumed to be rational, hold the same portfolio of risky assets, the "market portfolio", along with a portion, positive or negative (according to the investor's risk aversion), of the risk-free asset. From the condition of equilibrium and the hypothesis of uniform beliefs it follows that the market portfolio must be capitalization weighted. The investment weights in each asset must therefore be strictly positive and proportional to the ratio of its capitalization and the total capitalization of the universe of investible assets.

As a consequence, according to both MPT and CAPM, rationality and efficiency are interlocked concepts: investors are rational if and only if they select an efficient set of assets. It is useful to underline that the concept of optimal, i.e. the most efficient, portfolio varies between the MPT and the CAPM: the former allows for the presence of several efficient risky portfolios, while the latter postulates the existence of only one. This discrepancy is often overlooked by asset managers, mostly due to the high level of personalization necessary to advise investors in the MPT framework. On the other hand, the development of indexed investment vehicles, such as exchange traded funds (ETFs), has allowed retail investors to try to apply the CAPM, provided they are able to choose the benchmark most suitable to approximate the market portfolio.

II. Naïve Diversification and Indexed Investments

Investors usually follow apparently sub-optimal allocation models, one of which is equal weighting. This allocation technique violates the hypothesis of investors' rationality, since theoretically it should lead to sub-efficient portfolios. This analysis, on the other hand, aims at demonstrating empirically that the efficiency of equal weighted portfolios can be the highest attainable by indexed investors in the environment of financial markets, and consequently it provides an investment choice coherent with the postulate of rationality.

Fisher and Statman [5] [6] have studied the investment choices made by "normal" or "behavioral" investors. While optimized "à la Markowitz" portfolios include extreme positions on some assets and often exclude whole asset classes, investors show a layer-by-layer approach to portfolio construction. Each layer is aimed at a specific goal. For example, there may be a "downside protection layer", insuring against large losses, or an "upside potential layer", aimed at high returns. These investment decisions uncover utility functions with varying degrees of risk aversion, which are higher for investments that are perceived as being protective and lower for speculative or lottery-like investments. Layers are then bound together, without any optimization, usually with the same weight and ignoring the covariance between components. Along with this strictly behavioral approach, the authors underline the influence of the advice of consultants, who usually follow ERISA standards, which in turn may promote naïve diversification.

Benartzi and Thaler [7] have been the first to link the 1/*N* asset allocation rule to the "diversification heuristic", as defined for the first time by Read and Lowenstein [8], but discovered empirically by Simonson [9]. While these earlier researchers focused their attention on the optimal consumption decisions, Benartzi and Thaler evaluated the impact of the 1/*N* heuristic on defined contribution saving plans. They discovered that investors act by following different criteria related to whether they are allowed to choose sequentially or simultaneously. In the former case, they are able to evaluate their decisions step by step and modify their behavior according to their utility functions. In the latter case, investors are required to choose the allocation of their wealth in a single moment, as hypothesized by every one-period model, such as MPT and CAPM. The result of this constraint is that individuals tend to spread their wealth evenly across the investment options irrespective of their expected return, volatility and correlations. In other words, investors act according to the 1/*N* heuristic.

A further reason for "not-choosing" any specific asset allocation may be detected in the Regret Theory, developed by Loomes and Sudgen [10]. Within this framework, investors' utility is affected not only by the outcome of the chosen action, but also by the outcomes of those actions that have not been implemented. If the regret associated with the scenarios that have not been chosen is strong enough to overcome the expected utility of the single allocation which the investor has been able to select, then he/she may decide to avoid any decision about the mix of his/her assets and follow a mechanical strategy, such as equal weighting. It may seem surprising that Harry Markowitz himself has acknowledged that his retirement plan follows the most simple of equal weight allocations, for reasons that may be traced to the Regret Theory: "I should have computed the historical co-variances of the asset classes and drawn an efficient frontier. Instead, I visualized my grief if the stock market went way up and I wasn't in it - or if it went way down and I was completely in it. My intention was to minimize my future regret. So I split my contributions 50/50 between bonds and equities." [11], p. 118.

III. Efficiency Tests

Black, Jensen and Scholes [12] have provided the first test of the empirical validity of CAPM and, at the same time, the first example of what could be considered to be an implicit efficiency test. This judgment is justified by Roll's critique [13], according to which the market portfolio postulated by CAPM is unobservable in the real world. The corollary of this assertion is that CAPM itself is not empirically verifiable, because its tests would necessarily jointly test both the CAPM and the efficiency of the appropriate proxy for the market portfolio. This critique, which apparently undermines the use of CAPM in empirical analysis, is, on the contrary, the theoretical basis for the following tests of index efficiency. In order to conceive these tests in the CAPM framework, it is necessary to limit the degrees of freedom from two to one: once the empirical validity of the CAPM is assumed as verified, the only object of the test is the efficiency of the index used as a proxy.

Under these assumptions, Gibbons, Ross and Shanken [14] have developed their multivariate test, under the null $H_0 : \hat{\boldsymbol{\alpha}} = \mathbf{0}$, with $\hat{\boldsymbol{\alpha}}$ N×1 vector of intercepts of the regression of excess returns of the panel of *N* components of index *P*, proxy of the market portfolio, on the excess returns of *P* itself:

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i r_{p,t} + \hat{\varepsilon}_{i,t} . \quad (1)$$

Residuals (N×T matrix $\hat{\boldsymbol{\varepsilon}}$) are distributed as a normal with mean zero and diagonal covariance matrix $\hat{\boldsymbol{\Sigma}}$ (dimensions N×N), since the residuals are uncorrelated by hypothesis. The hypothesis of normality, imposed by the authors, would not be strictly necessary in order to evaluate the test statistic, but Shanken [15] has underlined its sensibility to conditional heteroskedasticity of $\hat{\boldsymbol{\varepsilon}}$.

The evaluation of the statistical significance of the intercepts is carried out through recourse to a Wald test (*WT*) using the following notation:

$$WT = T\left[1+\frac{\hat{\mu}_P^2}{\hat{\sigma}_P^2}\right]^{-1} \underset{1\times N}{\hat{\boldsymbol{\alpha}}'}\ \underset{N\times N}{\hat{\boldsymbol{\Sigma}}^{-1}}\ \underset{N\times 1}{\hat{\boldsymbol{\alpha}}} . \quad (2)$$

The mean of the excess returns of *P* is indicated by $\hat{\mu}_P$ and their variance by $\hat{\sigma}_P^2$.

On the other hand, the Wald test suffers from a serious practical shortcoming: its distribution is known only asymptotically and, consequently, any empirical analysis based upon this test must resort to simulation techniques. In order to overcome this problem, Gibbons, Ross and Shanken have proposed the following correction to the Wald test, thanks to which its small sample distribution is known and, when H_0 holds true, is:

$$\frac{WT}{T}\frac{(T-N-1)}{N} \sim F(N, T-N-1) . \quad (3)$$

In other words, it is possible to carry out a linear transformation of the Wald test that is distributed as an *F* with *N* and *T*–*N*–1 degrees of freedom. This *F* distribution is non-central when H_0 cannot be accepted, because its non-centrality parameter is zero when $\hat{\boldsymbol{\alpha}} = \mathbf{0}$. The Gibbons, Ross and Shanken (GRS) test is then:

$$GRS = \frac{(T-N-1)}{N}\left[1+\frac{\hat{\mu}_P^2}{\hat{\sigma}_P^2}\right]^{-1} \underset{1\times N}{\hat{\boldsymbol{\alpha}}'}\ \underset{N\times N}{\hat{\boldsymbol{\Sigma}}^{-1}}\ \underset{N\times 1}{\hat{\boldsymbol{\alpha}}} . \quad (4)$$

This formulation of the GRS test allows for its decomposition into factors of clear economic interpretation. The ratio of the squared mean and variance of *P* is nothing else than the squared Sharpe ratio of *P* (SR_P). Less evident is the meaning of the quadratic form $\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}$. In fact, this matrix product is the summation of the ratios of the squared alphas and variances of residuals (only if, as assumed by the model, $\hat{\mathbf{\Sigma}}$ is diagonal). Recalling that the appraisal ratio (*AR*) is defined as the ratio between alpha and the standard deviation of residuals, we can rewrite the quadratic form into:

$$\hat{\boldsymbol{\alpha}}'\hat{\mathbf{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}} = \sum_{i=1}^{N}\frac{\alpha_i^2}{\sigma_i^2} = \sum_{i=1}^{N} AR_i^2 . \quad (5)$$

Therefore the GRS test can be reformulated using measures typical of asset management performance evaluation:

$$GRS = \frac{(T-N-1)}{N}\cdot\frac{\sum_{i=1}^{N} AR_i^2}{(1+SR_P^2)} . \quad (6)$$

According to the *F* distribution, the probability that the portfolio *P* is efficient increases as the *GRS*-stat approaches zero. Given (6), the efficiency of *P* is:

- directly proportional to the square of its Sharpe ratio;
- inversely proportional to the sum of the squared appraisal ratios of its components.

The direct link to the Sharpe ratio is in perfect accordance with the CAPM: given that the Sharpe ratio of *P* is the slope of the capital allocation line passing through *P*, the higher the slope, the greater the degree of efficiency of an asset. Also the inverse relation to the appraisal ratio is linked to the CAPM theory. Given that the presence of positive or negative intercepts is not envisaged by that model, the presence of significant AR_i would be in contrast with the notion of efficiency of *P*. Recalling that AR_i is measured as the ratio between the intercept α_i and the standard deviation of residuals of the CAPM regression, a value of AR_i near zero implies that either the intercept is small or it is not statistically significant due to its volatility.

Empirical analyses carried out on the GRS test by Gibbons, Ross and Shanken [14], Campbell, Lo and MacKinlay [16] and Sentana [17] have shown that its power, i.e. the probability that the test will reject the null hypothesis when the null hypothesis is false, is sensitive to sample size. Power increases with length *T*, but declines as the total number of assets *N* grows: Campbell, Lo and MacKinlay [16] suggest keeping *N* not larger than 10.

Under the assumption that residuals are i.i.d., Gibbons, Ross and Shanken [14] show that:

$$SR_M^2 = \hat{\boldsymbol{\mu}}'\hat{\mathbf{V}}^{-1}\hat{\boldsymbol{\mu}} \quad (7)$$

and

$$\hat{\boldsymbol{\alpha}}'\hat{\mathbf{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}} = \sum_{i=1}^{N} AR_i^2 = SR_M^2 - SR_P^2 . \quad (8)$$

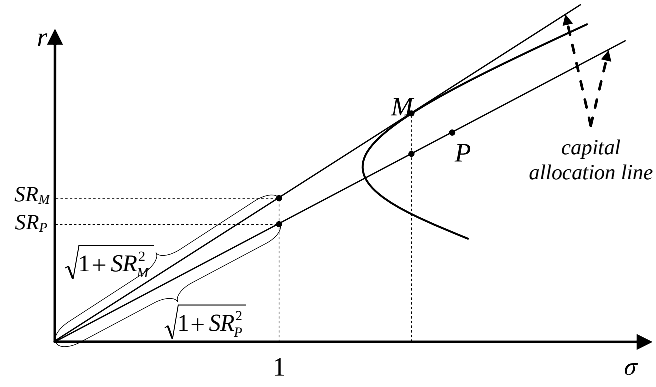

Figure 1. Graphical representation of GRS and Wald tests.

Replacing these equivalences in the original GRS test formula, we get:

$$GRS = \frac{(T-N-1)}{N}\cdot\frac{SR_M^2 - SR_P^2}{1+SR_P^2} = = \frac{(T-N-1)}{N}\cdot\left(\frac{1+SR_M^2}{1+SR_P^2}-1\right) . \quad (9)$$

The last factor of (9) can be rewritten as:

$$\frac{1+SR_M^2}{1+SR_P^2} - 1 = \left(\frac{\sqrt{1+SR_M^2}}{\sqrt{1+SR_P^2}}\right)^2 - 1 . \quad (10)$$

This new formulation shows that the GRS-stat is proportional to the ratio of the lengths of the hypotenuses of two right-angled triangles. The ratio converges to 1, and thus GRS to zero, as the Sharpe ratio of *P* approaches the Sharpe ratio of the market portfolio. Graphically, the GRS test can be represented as in figure 1. Given that the GRS test is derived from the Wald test, the same interpretation can also be applied to the latter.

The GRS test is based upon a finite sample of data, which is an important advantage, given that it makes it possible to analyze a historical time series of returns, but suffers from its assumption of the normality of returns. In order to model the presence of the heteroskedasticity of residuals, we have resorted to the Generalized Method of Moments test [18]. The first and principal shortcoming of this test is that its distribution is known only asymptotically, and thus requires the use of simulation techniques. The most common notation used for this test is the Wald-like one, as reported by Chou and Zhou [19]:

$$J_1 = T\hat{\boldsymbol{\alpha}}'[\boldsymbol{\eta}[\mathbf{D}'_T\,\mathbf{S}_T^{-1}\mathbf{D}_T]^{-1}\boldsymbol{\eta}']\hat{\boldsymbol{\alpha}} . \quad (11)$$

The details of formula (11) are provided in Appendix A.

The GMM test, in other words, is a Wald-like test, in which co-variances of residuals are correlated with the returns of the components of index *P*: the GMM test works in a framework of conditional heteroskedasticity.

Given the fact that the distribution of J_1 is known only asymptotically, it is necessary to utilize sampling techniques. This solution can lead to sub-optimal results in the case of serial correlation of residuals. Returns, in fact, show cross-section correlation and time series correlation and therefore, if resampling were to be applied in each time t of length equal to one, only the former type of correlation could be simulated. In fact, while cross-section correlation can always be simulated through sampling techniques, one-time sampling ignores any time-series correlation present in the original returns. In order to overcome these limitations, it is necessary to apply a heuristic technique such as the block bootstrap.

Block bootstrap consists of the joint extraction of blocks of consecutive residuals of returns, each block having a predefined length b. It is precisely this length that allows for the simulation of autocorrelation, even though only within each one of the blocks. It should be noted that, when the bootstrap jumps to a new block, it can be sampled from another non-consecutive point in the data series and thus it may be uncorrelated with the former block. As a consequence, the choice of b is subject to the following conflicting issues:

- if b is smaller, a lower relevance is given to autocorrelation;

- if b is larger, there can be fewer possible permutations based upon the available panel of data, which is necessarily limited.

While the tests analyzed so far are aimed at evaluating the statistical significance of the intercepts of component assets when subject to a CAPM-like regression on portfolio P, the measure of relative efficiency by Kandel and Stambaugh [20] follows a different approach. These authors, in fact, evaluate relative efficiency, measured with respect to the efficient frontier. The aim is to compare the excess return of P to that of x, i.e. the efficient portfolio with the same volatility of P. In order to implement this comparison, the excess return of the minimum variance portfolio g (see figure 2) is subtracted from both the returns of P and x. In formal terms, ψ_P, i.e. the measure of relative efficiency of portfolio P, is defined as:

$$\psi_P = \frac{\mu_P - \mu_g}{\mu_x - \mu_g}. \tag{12}$$

The perfect efficiency of P is measured when ψ_P is equal to +1. In this case P and x are coincident and thus P lies on the efficient frontier.

This approach, despite its apparent simplicity, offers some advantages if compared to the tests based upon the theoretical assumptions of CAPM: through Kandel and Stambaugh's measure it is always possible to sort portfolios according to their level of efficiency, even if the empirical validity of CAPM were not verified.

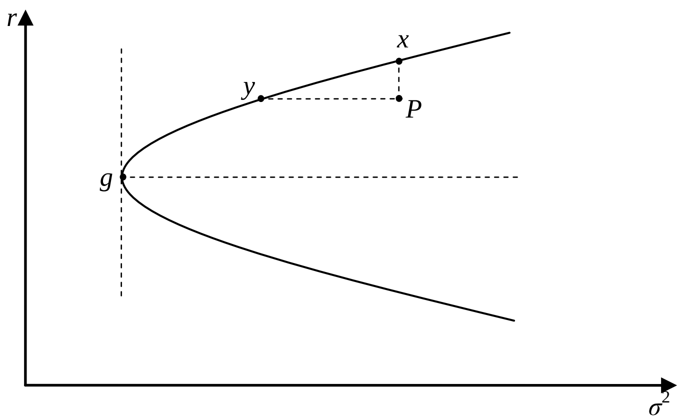

Figure 2. Kandel and Stambaugh's efficiency measure.

This measure of efficiency had been conceived two years before the introduction of the resampling technique by Michaud [21], a simulation method which attempts to overcome the serious problems of error maximization typical of the usual construction of the efficient frontier, which causes extreme allocations in only a few assets. In this empirical study, the estimation of portfolios g and x has been implemented through the resampling of 1,000 scenarios for each frontier.

IV. Empirical analysis

A. *Composition of the Sample*

Indices representative of the US and EMU stock markets are constructed following several different techniques and, therefore, these asset classes have been chosen as the sample utilized by this empirical analysis. This sample of indices has been selected in order to compare the equal weighting and the most typical construction techniques employed by index providers, with the aim of identifying the most efficient one, which therefore should be followed by rational indexed investors. This analysis will make it possible to understand whether a supposedly non-rational asset allocation method, such as equal weighting, might be regarded as more efficient than the other weighting schemes and consequently a more rational choice for indexed investors. This aim can be reached thanks to the presence of equal weighted indices in the available sample: an investor interested in a certain asset class and who follows the "1/N rule", in fact, can apply this strategy easily by buying such an index. The list of indices and their characteristics are reported in table 1.

The selection of indices follows quali-quantitative criteria. In fact, the chosen indices are among the benchmarks most used by practitioners, and they have to comply with the following features:

- having a track record of monthly returns, available in the database Morningstar Direct, since at least January 2003;

- being composed, completely or partially, of the stocks that belong to the beta-sorted portfolios described in detail in the next paragraph.

International Journal of Business and Management Study - IJBMS
Volume 3 : Issue 1 [ISSN 2372-3955]
Publication Date : 18 April, 2016

TABLE I. PANEL A: US STOCK INDICES

Index	Weighting[a]	Reinvestment of income[b]
DJ Composite Average PR	Price	No
DJ Composite Average TR	Price	Gross
DJ Industrial Average PR	Price	No
DJ Industrial Average TR	Price	Gross
MSCI USA GR	Float	Gross
MSCI USA NR	Float	Net
MSCI USA PR	Float	No
Russell 1000 CEW TR	Equal	Gross
Russell 1000 SEW TR	Sector Equal	Gross
Russell 1000 PR	Float	No
Russell 1000 TR	Float	Gross
Russell 3000 CEW TR	Equal	Gross
Russell 3000 PR	Float	No
Russell 3000 TR	Float	Gross
S&P 500 Equal Weighted TR	Equal	Gross
S&P 500 NR	Float	Net
S&P 500 PR	Float	No
S&P 500 TR	Float	Gross
Value Line New Arithmetic PR	Equal	No
Wilshire 5000 Equal Weight PR	Equal	No
Wilshire 5000 Total Market PR	Float	No

TABLE I. PANEL B: EMU STOCK INDICES

Index	Weighting[a]	Reinvestment of income[b]
EURO STOXX 50 GR	Float, cap 10%	Gross
EURO STOXX 50 NR	Float, cap 10%	Net
EURO STOXX 50 PR	Float, cap 10%	No
EURO STOXX 50 EW NR	Equal weight	Net
EURO STOXX 50 EW PR	Equal weight	No
EURO STOXX GR	Float, cap 20%	Gross
EURO STOXX NR	Float, cap 20%	Net
EURO STOXX PR	Float, cap 20%	No
FTSEurofirst 80 TR	Float	Gross
FTSEurofirst 300 Eurozone PR	Float	No
MSCI EMU GR	Float	Gross
MSCI EMU NR	Float	Net
MSCI EMU PR	Float	No
S&P Euro PR	Float	No
S&P Euro TR	Float	Gross

a. Weighting scheme of the index:
- "Price": the index is weighted according to the price of each component;
- "Float": the index is capitalization weighted, adjusted to take into account the free float of each component;
- "Equal": the index is equal weighted, in other words components have the same weight at each rebalancing date;
- "Sector equal": the economic sectors of the components have the same weights at each rebalancing date, and at the same date components are equal weighted within each sector.

b. The column "Reinvestment of income" indicates whether dividends, bonus shares, etc. are regarded as reinvested in the index net ("Net") or gross ("Gross") of taxes. If no income is included in the calculation of the index, it is reported as "Price".

The chosen proxy of risk-free rate on the US Dollar is the return of the Citigroup Treasury Bill 1 Month USD, an index calculated as the monthly mean of the T-Bills issued with a maturity of four weeks. The proxy of Euro risk-free rate is, instead, the return of the Citigroup EUR EuroDeposit 1 Month EUR, an index calculated as the monthly average of the bid rates on Eurodeposits denominated in Euro with a maturity of one month.

The tests used for this analysis are subject to potential biases depending upon the sample dimensions, as already stated with reference to the GRS test. In order to invert the covariance matrix, it is necessary that the number of assets N be smaller than the time length T. With regard to Kandel and Stambaugh's measure, the impact is less relevant from a strictly statistical point of view, but it is more important according to the practice of asset management: the larger N is, the higher the probability of including assets with extreme performances. Because of this, there is a higher probability that the efficient frontier is composed only of assets with risk-return profiles that it would be impossible to replicate out of sample.

B. *Beta-sorted Portfolios*

Given these premises, it has been necessary to solve the problem of reducing the N/T ratio by limiting its numerator. Black, Jensen and Scholes [12] provide a well-known aggregation method: beta-sorted portfolios. The first step in their construction is the estimation of the vector of the slopes β_i of the OLS regressions of the returns of the N assets on the returns of the portfolio of which they are components. Subsequently, these assets are ordered according to their slope and subdivided into an arbitrary number of quantiles Q. In the present study, following Gibbons, Ross and Shanken [14] and Campbell, Lo and MacKinlay [16], Q has been set as being equal to ten.

Each beta-sorted portfolio is composed of the assets of its corresponding quantile and its returns are equal to the arithmetic mean of the returns of its components. Thus the slope of the portfolio is equal to the systematic risk of the assets included in the portfolio and its intercept is equal to the mean intercept. In order to implement the analysis, it has been necessary to select the time series of the returns of the components related to two indices, representative respectively of the US and EMU stock market, on the database Datastream. Such indices have been identified in the Standard & Poor's 500, composed of the 500 largest companies listed on US stock markets, and the Euro Stoxx 50, a free-float weighted average of the 50 supersector leaders from the Euro Stoxx index.

Given that these indices are subject to quarterly revision, both with regard to their components and to their weights, it would have been incorrect to use only the stocks included on the last available date. For these reasons, the following procedure has been implemented for each index:

- the list of components for each quarter since December 2002 until September 2010 has been downloaded from Datastream;

- 63 monthly total returns of the stocks of each list have been downloaded from Datastream, of which the first 60 months (in sample) has been used for the estimate of the betas and 3 (out of sample) for the construction of beta-sorted portfolios;

- for each rolling window of 63 months, the first 60 monthly returns of the components have been regressed on the index only when they have at least 24 months in sample and two out of sample;

- stocks have been sorted according to their beta and aggregated into ten beta-sorted portfolios;

- the monthly return of each beta-sorted portfolio has been calculated as the arithmetic mean of the returns of its components in the out of sample months, covering the period January 2003-December 2010.

From this procedure it can be implied that the composition of each portfolio varies in time, modifying itself quarterly: in practice, beta-sorted portfolios can be regarded as mutual funds, endowed with their own autonomous identity but with a variable internal composition. The quarterly recalibration of beta-sorted portfolios has an important advantage: it allows for the relocation of stocks in different portfolios according to the variation in their beta, if it occurs, even though it may be implemented with a temporal lag of three months at worst. Thus, with regard to the variable composition of the beta-sorted portfolios, their risk profile is kept constant, because stocks are transferred to a different portfolio when their beta migrates to another quantile.

C. *Implementation of the Empirical Analysis*

Out of sample series of returns of the beta-sorted portfolios have been used as a panel of components for all the indices of the "US Stock Market" and "EMU Stock Market" asset classes, regardless of whether such beta-sorted portfolios are or are not composed of the same assets included in each of the indices subject to this analysis. This choice, besides being caused by a lack of data about the time series of every index, is founded also on theoretical bases: in order to identify the most efficient construction methods for a market benchmark, it is useful to compare each stock index to the same sample of assets.

Given that all the tests employed, except for the GMM test, hypothesize the presence of normal returns, the possible deviations from the Gaussian distribution have been analyzed through the Jarque-Bera test (table 2). This statistical measure, which distributes asymptotically as a chi-square with two degrees of freedom, suffers from serious bias if the sample is limited. Because of this, the analysis has been made in Matlab, a program that estimates the p-value of the JB test according to a table of critical values computed through Monte Carlo simulations.

Given the outcome of the JB tests, shown in table 2, it is possible to infer that the GMM test will be the most significant, since it is the only one extraneous to the hypothesis of normality. The GRS test has been used in its original notation, instead of in its decomposition into the Sharpe ratio and appraisal ratio, because this latter method is too biased in the presence of correlated residuals. The Wald test has been implemented through a bootstrap simulation, in order to model an empirical distribution that can overcome the problems related to limited samples. For each index and the ten beta-sorted portfolios, 10,000 scenarios have been simulated, following a procedure identical to that described for the block bootstrap, except for the fact that the length of each block has been fixed to one period. In fact, the Wald test assumes that residuals are i.i.d., and thus the presence of serial correlation is excluded a priori by this model. The p-value of the test is equal to the number of simulated WT^* larger than WT, divided by 10,000.

Unlike the other tests, Kandel and Stambaugh's measure does not impose limits on the number of assets N or on the length T, but the ten beta-sorted portfolios have been used for the construction of the resampled frontiers. This choice has been dictated both by coherence with the other efficiency indicators used in this study and by statistical reasons. Moreover, the grouping of stocks into portfolios limits the impact of outlier returns, further reducing error maximization. In detail, the procedure for each asset class follows these steps:

- in 10,000 scenarios, each one 96 months long, the monthly excess returns of the ten beta-sorted portfolios and of all the indices have been jointly simulated;

- in each scenario and for each index, the efficient frontiers, composed of the ten beta-sorted portfolios and one index a time, have been estimated;

- for each index, the resampled frontier has been calculated and, through a cubic spline interpolation, the efficient portfolio x has been identified;

- finally, for each index, the value of ψ_P has been calculated.

The empirical distribution of the GMM test J_1 statistics has been estimated through the block bootstrap of 10,000 scenarios, using blocks of a length of six months each. This length has been defined according to the autocorrelation of excess returns in the indices of the sample: in their large majority (19 out of 22 in the US stock market, 14 out of 15 in the EMU stock market) autocorrelation is statistically significant up to the fourth lag. The use of blocks with a length of six periods, then, is a compromise that makes it possible to capture, within each block, an autocorrelation of:

- first order, in five periods out of six;

- fourth order, in two periods out of six.

D. *Findings of the Empirical Analysis*

The evaluation of efficiency has reached results that are substantially concordant among themselves and useful in identifying optimal construction techniques, even though the analysis has been based upon tests which are quite different from each other, both because of their theoretical bases and because of their construction. The underlying theoretical framework is, in fact, the CAPM for the GRS, Wald and GMM tests, and the MPT for the Kandel and Stambaugh measure. Moreover, while the GRS employs the historical time series of returns, the other tests require the use of simulation techniques, such as the bootstrap (Wald test), the block bootstrap (GMM test) and resampling (Kandel and Stambaugh). Table 3 shows the results of the tests and, where possible, their percentage of p-value, i.e. of the probability that the hypothesis of efficiency cannot be rejected.

It can be observed that all the indices are efficient in the time-span considered, which is characterized by an initial growth in stock prices and then by a subsequent time of strong turbulence on markets. Along with this overall judgment, it can be useful to analyze the ranking obtained using the results of the tests. Given that there are four tests, it is not always possible to reach an univocal judgment. In order, then, to construct a unitary ranking it is possible to follow a multi-criteria analysis approach, typical of decision theory, such as the PROMETHEE (Preference Ranking Organization Method for Enrichment of Evaluations) created by Brans [22]. With this technique, rankings are made using the net outranking flow $\varphi(i) = \varphi^{+}(i) - \varphi^{-}(i)$, where $\varphi^{+}(i)$ and $\varphi^{-}(i)$ are respectively the positive and negative outranking flows, which express how much an alternative is respectively outranking and outranked by all the others. It can be observed that all the indices are efficient in the time-span considered, which is characterized by an initial growth in stock prices and then by a subsequent time of strong turbulence on markets. Along with this overall judgment, it can be useful to analyze the ranking obtained using the results of the tests. Given that there are four tests, it is not always possible to reach an univocal judgment. In order, then, to construct a unitary ranking it is possible to follow a multi-criteria analysis approach, typical of decision theory, such as the PROMETHEE (Preference Ranking Organization Method for Enrichment of Evaluations) created by Brans [22]. With this technique, rankings are made using the net outranking flow $\varphi(i) = \varphi^{+}(i) - \varphi^{-}(i)$, where $\varphi^{+}(i)$ and $\varphi^{-}(i)$ are respectively the positive and negative outranking flows, which express how much an alternative is respectively outranking and outranked by all the others.

As is shown in table 4, panel A, five of the six indices built through the technique of equal weighting are in the first five places in the efficiency ranking in the US market. Within this sub-sample we can observe, among the Russell indices, the presence of two methods of equal weighting:

- constituent equal weight (CEW);

- sector equal weight (SEW).

TABLE II. PANEL A: DESCRIPTIVE STATISTICS OF US STOCK INDICES

Index	Mean	Standard Deviation	Asymmetry	Kurtosis	JB stat	p-value[a]	Normality
DJ Composite Average PR	0.48%	4.21%	-0.80	4.53	19.51	0.36%	No
DJ Composite Average TR	0.68%	4.20%	-0.79	4.50	19.12	0.38%	No
DJ Industrial Average PR	0.26%	4.11%	-0.70	4.40	15.72	0.61%	No
DJ Industrial Average TR	0.47%	4.10%	-0.69	4.37	15.06	0.67%	No
MSCI USA GR	0.49%	4.39%	-0.85	5.07	28.79	0.14%	No
MSCI USA NR	0.44%	4.40%	-0.86	5.07	28.90	0.13%	No
MSCI USA PR	0.33%	4.40%	-0.86	5.08	29.15	0.13%	No
Russell 1000 CEW TR	0.95%	5.66%	-0.34	6.29	45.24	0.10%	No
Russell 1000 SEW TR	1.07%	5.41%	-0.57	6.58	56.29	0.10%	No
Russell 1000 PR	0.36%	4.47%	-0.88	5.14	30.80	0.11%	No
Russell 1000 TR	0.52%	4.46%	-0.87	5.13	30.43	0.12%	No
Russell 3000 CEW TR	1.01%	6.34%	-0.21	5.14	19.04	0.38%	No
Russell 3000 PR	0.39%	4.55%	-0.88	5.07	29.48	0.13%	No
Russell 3000 TR	0.54%	4.55%	-0.87	5.07	29.10	0.13%	No
S&P 500 Equal Weighted TR	0.83%	5.38%	-0.51	5.86	36.82	0.10%	No
S&P 500 NR	0.42%	4.38%	-0.84	4.97	26.77	0.16%	No
S&P 500 PR	0.31%	4.39%	-0.84	4.93	26.08	0.17%	No
S&P 500 TR	0.47%	4.38%	-0.83	4.97	26.61	0.17%	No
Value Line New Arithmetic PR	1.11%	6.17%	-0.22	5.25	20.92	0.31%	No
Wilshire 5000 Equal Weight PR	1.31%	6.61%	-0.15	4.48	9.09	1.94%	No
Wilshire 5000 Total Market PR	0.41%	4.53%	-0.87	5.03	28.64	0.14%	No

a. Rounded to 0,10% by Matlab if tending to zero.

TABLE II. PANEL B: DESCRIPTIVE STATISTICS OF EMU STOCK INDICES

Index	Mean	Standard Deviation	Asymmetry	Kurtosis	JB stat	p-value[a]	Normality
EURO STOXX 50 GR	0.40%	5.21%	-0.34	4.26	8.22	2.34%	No
EURO STOXX 50 NR	0.33%	5.20%	-0.35	4.21	7.82	2.59%	No
EURO STOXX 50 PR	0.10%	5.18%	-0.38	4.04	6.61	3.51%	No
EURO STOXX 50 EW NR	0.65%	7.17%	-0.40	4.63	13.18	0.89%	No
EURO STOXX 50 EW PR	0.28%	5.62%	-0.06	4.92	14.88	0.69%	No
EURO STOXX GR	0.51%	5.17%	-0.47	4.53	12.93	0.93%	No
EURO STOXX NR	0.45%	5.16%	-0.48	4.49	12.55	1.00%	No
EURO STOXX PR	0.23%	5.13%	-0.52	4.33	11.36	1.23%	No
FTSEurofirst 80 TR	0.44%	5.19%	-0.39	4.36	9.84	1.67%	No
FTSEurofirst 300 Eurozone PR	0.19%	5.08%	-0.52	4.28	10.93	1.34%	No
MSCI EMU GR	0.48%	5.16%	-0.45	4.53	12.61	0.99%	No
MSCI EMU NR	0.42%	5.15%	-0.47	4.48	12.21	1.06%	No
MSCI EMU PR	0.20%	5.12%	-0.50	4.31	10.93	1.34%	No
S&P Euro PR	0.19%	5.12%	-0.48	4.20	9.50	1.79%	No
S&P Euro TR	0.47%	5.15%	-0.44	4.40	10.95	1.33%	No

a. Rounded to 0,10% by Matlab if tending to zero.

The former technique follows the usual concept of equal weighting, which attributes the same weight to each component at each rebalancing date. For example, while in the Russell 1000 (float weight index) the weights of the first 100 stocks sum to 58% of the total, in the equal weight version, by definition, they represent just 10%. However, the technique of sector equal weight has been developed by Russell and today is used only by this index provider. It has been implemented to avoid a biased representation of economic sectors (sector bias). As a consequence, the weight of each component is calculated through two passages:

- the stock market is divided into nine economic sectors, each one included in the index with the same weight;

- within each sector, stocks are equal weighted.

The Russell 1000 SEW TR, fourth in the ranking, appears to be less efficient than the Russell 1000 CEW TR (the most efficient index according to the ranking), regardless of the more advanced construction scheme of the former, but it must be underlined that this result is influenced by the GRS test, which is not completely reliable in the presence of deviations from the Gaussian distribution.

TABLE III. PANEL A: EFFICIENCY LEVELS OF US STOCK INDICES

Index	Kandel & Stambaugh	GRS		Wald		GMM	
		GRS	p-value	WT	p-value[a]	J_1	p-value[a]
DJ Composite Average PR	0.4896	0.7882	95.49%	8.9019	93.09%	8.4437	59.22%
DJ Composite Average TR	0.9363	0.7969	95.42%	9.0008	95.80%	8.6592	60.65%
DJ Industrial Average PR	-0.0472	0.8725	94.80%	9.8545	84.78%	9.6189	57.34%
DJ Industrial Average TR	0.4929	0.7926	95.45%	8.9523	92.68%	8.5761	65.76%
MSCI USA GR	0.4971	0.7962	95.42%	8.9918	92.21%	8.3761	86.59%
MSCI USA NR	0.3779	0.8053	95.35%	9.0953	92.66%	8.5220	82.51%
MSCI USA PR	0.0909	0.8494	94.98%	9.5935	90.20%	9.1762	81.50%
Russell 1000 CEW TR	0.8838	0.7650	95.69%	8.6395	96.04%	8.1204	93.10%
Russell 1000 SEW TR	0.9843	0.8128	95.28%	9.1798	97.36%	9.3989	89.20%
Russell 1000 PR	0.1612	0.8370	95.08%	9.4536	90.39%	8.9910	81.90%
Russell 1000 TR	0.5505	0.7948	95.43%	8.9762	93.79%	8.3464	86.72%
Russell 3000 CEW TR	0.8638	0.7665	95.67%	8.6572	95.33%	7.9410	93.91%
Russell 3000 PR	0.2262	0.8280	95.16%	9.3516	89.89%	8.8369	83.67%
Russell 3000 TR	0.5891	0.7936	95.44%	8.9629	94.08%	8.3128	87.47%
S&P 500 Equal Weighted TR	0.8341	0.7854	95.51%	8.8704	96.64%	8.4389	91.91%
S&P 500 NR	0.3288	0.8104	95.30%	9.1529	90.88%	8.6104	81.56%
S&P 500 PR	0.0403	0.8593	94.90%	9.7055	88.71%	9.3234	78.48%
S&P 500 TR	0.4535	0.7987	95.40%	9.0205	92.89%	8.4266	84.42%
Value Line New Arithmetic PR	0.9334	0.8272	95.17%	9.3427	96.67%	9.5587	93.45%
Wilshire 5000 Equal Weight PR	0.9687	0.8693	94.82%	9.8181	94.45%	9.3997	88.86%
Wilshire 5000 Total Market PR	0.2925	0.8174	95.25%	9.2322	91.85%	8.6749	82.54%

a. P-value estimated through the bootstrap of 10,000 scenarios.

TABLE III. PANEL B: EFFICIENCY LEVELS OF EMU STOCK INDICES

Index	Kandel & Stambaugh	GRS		Wald		GMM	
		GRS	p-value	WT	p-value[a]	J_1	p-value[a]
EURO STOXX 50 GR	-0.1769	1.0359	93.54%	11.7001	84.45%	11.9253	46.52%
EURO STOXX 50 NR	-0.2880	1.2282	92.23%	13.8718	76.12%	14.0423	33.94%
EURO STOXX 50 PR	-0.6786	2.0008	88.40%	22.5977	51.76%	22.6221	8.74%
EURO STOXX 50 EW NR	0.1666	0.7281	96.00%	8.2236	90.71%	8.6829	75.04%
EURO STOXX 50 EW PR	-0.3364	1.2819	91.89%	14.4778	74.88%	14.4842	29.67%
EURO STOXX GR	-0.0009	0.7247	96.03%	8.1850	92.74%	8.4055	71.82%
EURO STOXX NR	-0.1026	0.7887	95.49%	8.9075	92.03%	9.0460	70.40%
EURO STOXX PR	-0.4623	1.1754	92.57%	13.2753	74.03%	13.2399	36.01%
FTSEurofirst 80 TR	-0.1171	0.9100	94.50%	10.2780	88.80%	10.5630	54.28%
FTSEurofirst 300 Eurozone PR	-0.5457	1.3676	91.38%	15.4463	69.31%	15.4468	30.79%
MSCI EMU GR	-0.0426	0.7537	95.78%	8.5126	92.23%	8.6949	71.07%
MSCI EMU NR	-0.1526	0.8392	95.07%	9.4775	90.22%	9.5809	65.26%
MSCI EMU PR	-0.5261	1.2915	91.83%	14.5866	71.97%	14.5407	33.50%
S&P Euro PR	-0.5345	1.3751	91.34%	15.5306	68.40%	15.5419	27.71%
S&P Euro TR	-0.0580	0.7794	95.56%	8.8030	91.91%	9.0043	64.94%

a. P-value estimated through the bootstrap of 10,000 scenarios.

TABLE IV. PANEL A: NET OUTRANKING FLOWS OF THE EFFICIENCY MEASURES OF US STOCK INDICES

	Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	$\varphi^+(i)$	$\Phi(i)$
1	DJ Composite Average PR	0.00	0.25	1.00	0.50	0.50	0.75	0.75	0.00	0.25	0.75	0.25	0.00	0.75	0.25	0.00	0.75	0.75	0.75	0.25	0.25	0.75	0.25	0.46	-0,07
2	DJ Composite Average TR	0.75	0.00	1.00	0.50	0.50	0.75	0.75	0.25	0.25	0.75	0.50	0.50	0.75	0.50	0.25	0.75	0.75	0.75	0.50	0.50	0.75	0.50	0.60	0,19
3	DJ Industrial Average PR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-1,00
4	DJ Industrial Average TR	0.50	0.50	1.00	0.00	0.50	0.75	0.75	0.00	0.25	0.75	0.25	0.00	0.75	0.25	0.00	0.75	0.75	0.50	0.25	0.25	0.75	0.25	0.46	-0,07
5	MSCI USA GR	0.50	0.50	1.00	0.50	0.00	0.75	1.00	0.00	0.25	1.00	0.00	0.00	1.00	0.00	0.00	1.00	1.00	0.75	0.25	0.25	1.00	0.25	0.52	0,05
6	MSCI USA NR	0.25	0.25	1.00	0.25	0.25	0.00	1.00	0.00	0.25	1.00	0.00	0.00	0.75	0.00	0.00	1.00	1.00	0.00	0.25	0.25	0.75	0.00	0.39	-0,21
7	MSCI USA PR	0.25	0.25	1.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	1.00	0.00	0.00	0.25	0.00	0.00	0.15	-0,69
8	Russell 1000 CEW TR	1.00	0.75	1.00	1.00	1.00	1.00	1.00	0.00	0.50	1.00	1.00	0.75	1.00	1.00	0.75	1.00	1.00	1.00	0.25	0.75	1.00	1.00	0.89	0,79
9	Russell 1000 SEW TR	0.75	0.75	1.00	0.75	0.75	0.75	1.00	0.50	0.00	1.00	0.75	0.50	1.00	0.75	0.50	0.75	1.00	0.75	0.75	1.00	1.00	0.75	0.80	0,60
10	Russell 1000 PR	0.25	0.25	1.00	0.25	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.25	1.00	0.00	0.00	0.25	0.00	0.00	0.21	-0,57
11	Russell 1000 TR	0.75	0.50	1.00	0.75	1.00	1.00	1.00	0.00	0.25	1.00	0.00	0.00	1.00	0.00	0.00	1.00	1.00	1.00	0.25	0.25	1.00	0.50	0.63	0,26
12	Russell 3000 CEW TR	1.00	0.50	1.00	1.00	1.00	1.00	1.00	0.25	0.50	1.00	1.00	0.00	1.00	1.00	0.75	1.00	1.00	1.00	0.50	0.75	1.00	1.00	0.87	0,74
13	Russell 3000 PR	0.25	0.25	1.00	0.25	0.00	0.25	0.75	0.00	0.00	0.75	0.00	0.00	0.00	0.00	0.00	0.25	1.00	0.00	0.00	0.25	0.25	0.00	0.25	-0,50
14	Russell 3000 TR	0.75	0.50	1.00	0.75	1.00	1.00	1.00	0.00	0.25	1.00	1.00	0.00	1.00	0.00	0.00	1.00	1.00	1.00	0.25	0.25	1.00	0.50	0.68	0,36
15	S&P 500 Equal Weighted TR	1.00	0.75	1.00	1.00	1.00	1.00	1.00	0.25	0.50	1.00	1.00	0.25	1.00	1.00	0.00	1.00	1.00	1.00	0.25	0.75	1.00	1.00	0.85	0,69
16	S&P 500 NR	0.25	0.25	1.00	0.25	0.00	0.00	1.00	0.00	0.25	0.75	0.00	0.00	0.75	0.00	0.00	0.00	1.00	0.00	0.25	0.25	0.50	0.00	0.31	-0,38
17	S&P 500 PR	0.25	0.25	1.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.10	-0,81
18	S&P 500 TR	0.25	0.25	1.00	0.50	0.25	1.00	1.00	0.00	0.25	1.00	0.00	0.00	1.00	0.00	0.00	1.00	1.00	0.00	0.25	0.25	1.00	0.00	0.48	-0,05
19	Value Line New Arithmetic PR	0.75	0.50	1.00	0.75	0.75	0.75	1.00	0.75	0.25	1.00	0.75	0.50	1.00	0.75	0.75	0.75	1.00	0.75	0.00	0.75	0.75	0.75	0.76	0,52
20	Wilshire 5000 Equal Weight PR	0.75	0.50	1.00	0.75	0.75	0.75	0.75	0.25	0.00	0.75	0.75	0.25	0.75	0.75	0.25	0.75	0.75	0.75	0.25	0.00	0.75	0.50	0.61	0,21
21	Wilshire 5000 Total Market PR	0.25	0.25	1.00	0.25	0.00	0.25	1.00	0.00	0.00	1.00	0.00	0.00	0.75	0.00	0.00	0.50	1.00	0.00	0.25	0.25	0.00	0.00	0.32	-0,36
22	Wilshire 5000 Total Market TR	0.75	0.50	1.00	0.75	0.75	1.00	1.00	0.00	0.25	1.00	0.50	0.00	1.00	0.50	0.00	1.00	1.00	1.00	0.25	0.50	1.00	0.00	0.65	0,31
Negative outranking flow $\phi^-(i)$		0,54	0.40	1.00	0.54	0.48	0.61	0.85	0.11	0.20	0.79	0.37	0.13	0.75	0.32	0.15	0.69	0.90	0.52	0.24	0.39	0.68	0.35		

TABLE IV. PANEL B: OUTRANKING FLOWS OF THE EFFICIENCY MEASURES OF EMU STOCK INDICES

	Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	$\varphi^+(i)$	$\Phi(i)$
1	EURO STOXX 50 GR	0.00	1.00	1.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	1.00	1.00	0.00	0.50	0,00
2	EURO STOXX 50 NR	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.50	0.00	1.00	0.00	0.00	1.00	1.00	0.00	0.39	-0,21
3	EURO STOXX 50 PR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-1,00
4	EURO STOXX 50 EW NR	1.00	1.00	1.00	0.00	1.00	0.50	0.75	1.00	1.00	1.00	0.75	1.00	1.00	1.00	0.75	0.91	0,82
5	EURO STOXX 50 EW PR	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.50	0.00	0.75	0.00	0.00	0.75	1.00	0.00	0.29	-0,43
6	EURO STOXX GR	1.00	1.00	1.00	0.50	1.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.96	0,93
7	EURO STOXX NR	1.00	1.00	1.00	0.25	1.00	0.00	0.00	1.00	1.00	1.00	0.00	1.00	1.00	1.00	0.50	0.77	0,54
8	EURO STOXX PR	0.00	0.50	1.00	0.00	0.50	0.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	1.00	0.00	0.36	-0,29
9	FTSEurofirst 80 TR	1.00	1.00	1.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	0.25	1.00	1.00	0.00	0.59	0,18
10	FTSEurofirst 300 Eurozone PR	0.00	0.00	1.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.75	0.00	0.14	-0,71
11	MSCI EMU GR	1.00	1.00	1.00	0.25	1.00	0.00	1.00	1.00	1.00	1.00	0.00	1.00	1.00	1.00	1.00	0.88	0,75
12	MSCI EMU NR	1.00	1.00	1.00	0.00	1.00	0.00	0.00	1.00	0.75	1.00	0.00	0.00	1.00	1.00	0.25	0.64	0,29
13	MSCI EMU PR	0.00	0.00	1.00	0.00	0.25	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.23	-0,54
14	S&P Euro PR	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.09	-0,82
15	S&P Euro TR	1.00	1.00	1.00	0.25	1.00	0.00	0.50	1.00	1.00	1.00	0.00	0.75	1.00	1.00	0.00	0.75	0,50
Negative outranking flow $\phi^-(i)$		0,54	0.50	0.61	1.00	0.09	0.71	0.04	0.23	0.64	0.41	0.86	0.13	0.36	0.77	0.91	0.25	

The Russell 3000 CEW TR, despite its high diversification, is less efficient than the Russell 1000 CEW TR: the latter is the second index in the ranking. This evaluation is significant, because it highlights how the potential impact of the size effect is not relevant in this case.

The Standard & Poor's 500 Equal Weight TR is the third most efficient index. The value of this result is particularly high: in fact, the stocks of S&P 500 are the components of the beta-sorted portfolios used in this analysis. As a consequence, tests carried out on the indices of the S&P 500 family do not show any potential bias deriving from the use of components which are not perfectly coincident with those of the index itself or due to the presence of different procedures of income reinvestment or of different tax rates. The ranking depends only on the weighting schemes, holding all else constant.

The fifth index for level of efficiency is the Value Line New Arithmetic PR. It is the equal weighted index of the about 1,700 stocks that are covered by the "Value Line Investment Survey". In this case, high efficiency is reached despite the fact that it is a price index. This result can be justified by the large diversification of the index and by the possible presence of selection bias caused by the active selection of the companies included: a fact that denies the character of objectivity typical of a good benchmark. Moreover, the returns of this index are calculated as the arithmetic mean of the daily returns of its components. This method implies a daily rebalancing, which would make the replication of this index too expensive and highly impractical. On the contrary, the other equal weight indices in this sample are subject to quarterly rebalancing. This implies a larger departure from perfect equal weighting, but, at the same time, a sharp reduction in transaction costs for an indexed investor.

This efficiency ranking offers some reflections about the treatment of income in the computing of indices. Within each "family" of indices calculated by the same index provider, the same ranking repeats itself. In the first place, there are gross total return indices; at a second level, given the reduction of reinvested income caused by taxes, there are net total return indices; finally, price indices are the least efficient. The float weight indices S&P 500 TR, NR and PR not only show a lower level of efficiency than their equal weight version, but also a clear distinction based upon the reinvestment of income, placing themselves respectively 12th, 17th and 21st. This same ranking is shown by MSCI indices. Also the Wilshire 5000 indices show a higher efficiency for equal weighting. The equal weight index of the "Wilshire 5000" series, even though it is just a price index, shows almost the same efficiency of the float weight gross total return version.

The Dow Jones indices have been analyzed due to their widespread use by financial media, even though their price weighting is devoid of any practical advantage today. The level of net outranking flow shown by the Dow Jones Industrial Average PR, equal to −1, is also significant. As a consequence, price weighted asset allocation can be regarded as the least efficient among the methodologies of indexed investing analyzed in this study.

As reported in table 4, panel B, the indices representative of the most relevant companies quoted on the markets of the European Monetary Union show an higher degree of efficiency when they are total return, especially when income, e.g. dividends, is reinvested gross of taxes. This outcome, on the other hand, does not repeat itself mechanically: the choice of an index can lead to optimal risk/reward profiles also independently from this first rule.

The Euro Stoxx NR, a net total return index, shows a degree of relative efficiency that is higher than that of three gross total return indices. This is an apparently counterintuitive outcome, but is justified by the higher degree of diversification of this index if compared to the other ones: the Euro Stoxx is composed by about 300 stocks.

The MSCI EMU NR, another net total return index, is its direct competitor as a benchmark for Eurozone stock markets. It is fifth in the ranking of relative efficiency and is composed of about 260 stocks, selected among the largest companies for free-float value. On the contrary, the Euro Stoxx NR represents a wider diversification, because it is composed of large, mid and small cap companies.

The indices of the Euro Stoxx 50 series offer an interesting example of how much construction techniques influence the level of efficiency. In fact, the stocks of the Euro Stoxx 50 are the components of the beta-sorted portfolios used in this analysis. As a consequence, tests carried out on the Euro Stoxx 50 family of indices do not show any potential bias deriving from the use of components not perfectly coincident with those of the index itself or due to the presence of different procedures of income reinvestment. The ranking, instead, depends only on the weighting schemes, holding all else constant.

On the other hand, weighting schemes play the key role: the Euro Stoxx 50 EW NR is, in fact, the second index for efficiency, despite its narrow sample of components and the impact of taxation on reinvested income. What makes it different from the other indices of this sample is, in fact, its construction technique: equal weighting (each component has the same weight of the other ones on every recalibration date). The Euro Stoxx 50 EW NR is, moreover, the most efficient index that can be realistically tracked by a passive investor. In fact it is calculated net of taxes on income, unlike the first index in the raking, the Euro Stoxx GR (gross total return).

To summarize the findings, then, we can conclude that equal weighting, holding all else constant, is the most efficient index calculation technique. Both in the US and EMU stock markets equal weight indices have surpassed, in terms of risk-adjusted return in the MPT and CAPM framework, any other kind of index.

V. Conclusions

The results of the empirical analysis provide a strong indication that equal weighted portfolios offer a superior risk-adjusted return if compared to traditional capitalization and price weighted ones. The causes of this phenomenon may be traced to a statistical interpretation, that can be divided into three coexisting theories.

The first approach, proposed by Treynor [23], underlines how the presence of "noise" in the prices of securities causes an excess weighting of overpriced stocks and, conversely, an underweighting of underpriced ones in cap-weighted indices, which are consequently subject to underperformance when prices tend to revert to their fair value.

The second theory, within the strictly statistical framework, that may explain the superior efficiency of equal weighted portfolios has been formulated by DeMiguel, Garlappi and Uppal [24], who have simulated the returns of portfolios constructed following several different techniques and have found that equal weighting provides the best out-of-sample risk-adjusted performance, even if compared to portfolios optimized "à la Markowitz". This result has been explained with the problem of estimation error, i.e. the investors' inability to measure the moments of returns distribution, which is so severe to make equal weighting the most efficient technique, since it ignores statistical measures in the portfolio construction process. Windcliff and Boyle [25], moreover, had already noticed this phenomenon, even though they had not measured it, explicitly linking the 1/N heuristic to the minimization of estimation error. The outcome of the empirical analysis is in accordance with these theoretical explanations.

The third theory traces the extra-performance of naïve portfolios to the so-called "diversification return", i.e. an incremental return earned by a rebalanced portfolio of assets [26] [27]. This return has been attributed to the underweighting, if compared to capitalization weighted indices, of assets that have earned a higher performance in the past and, conversely, to the overweighting of low-performance assets. Implicitly, the diversification return is traced to a mean-reversion of asset prices, but empirical analyses have excluded that this is the only source of extra-performance. As noted by Cuthbertson et al. [29], the primary factor behind diversification return is the reduction of volatility typical of equal weighted portfolios, due to their diversification, higher than that of capitalization weighted portfolios. This volatility contraction increases efficiency both by reducing risk and by increasing compounded returns, following the mathematical rule that geometric mean equals arithmetic mean when variance is zero.

The present study does not rule out the possibility that indexed investors who, for example, are able to construct a perfect proxy of the "CAPM market portfolio", or non-indexed investors, who follow active strategies such as stock picking, might reach higher levels of efficiency, but, according to this empirical analysis and the theoretical explanations put forward by scientific literature, the "1/N rule" can be regarded as the most efficient one for an indexed investor. As a consequence, for these subjects the 1/N heuristic should not be regarded as a "naïve" investment strategy but, rather, as a rational and simple method aimed at achieving superior efficiency.

Appendix A

The $\boldsymbol{\eta} = \mathbf{I}_N \otimes [1,\ 0]$: matrix composed of only 1s and 0s;

$\boldsymbol{\Omega}_T = [\mathbf{D}'_T\, \mathbf{S}_T^{-1} \mathbf{D}_T]^{-1}$: covariance matrix of the regression parameters;

$$\mathbf{D}_T = -\frac{1}{T}\sum_{t=1}^{T}[\mathbf{I}_N \otimes \mathbf{Z}_t \mathbf{Z}_t\,']$$: symmetrical matrix made of square submatrices, aligned along the main diagonal, containing the descriptive statistics of *P*;

$$\mathbf{S}_T = \frac{1}{T}\sum_{t=1}^{T}[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}'_t \otimes \mathbf{Z}_t \mathbf{Z}_t\,']$$: spectral density matrix;

$\mathbf{Z} = [\mathbf{1},\ \mathbf{r}_P\,']'$: matrix $2 \times T$ with only 1s on its first row and the excess returns of *P* on its second row.

References

[1] H. Markowitz, "Portfolio Selection," The Journal of Finance, vol. 7, pp. 77-91, 1952.

[2] W. Sharpe, "Capital Asset Prices: A Theory of Market Equilibrium," The Journal of Finance, vol. 19, pp. 425-442, 1964.

[3] J. Lintner, "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," Review of Economics and Statistics, vol. 47, pp. 13-37, 1965.

[4] J. Mossin, "Equilibrium in a Capital Asset Market." Econometrica, vol. 34, pp. 768-783, 1966.

[5] K. Fisher, and M. Statman, "The Mean-Variance-Optimization Puzzle: Security Portfolios and Food Portfolios," Financial Analysts Journal, vol. 53, pp. 41-50, 1997.

[6] K. Fisher, and M. Statman, "Investment Advice from Mutual Fund Companies. Closer to the Talmud than to Markowitz," The Journal of Portfolio Management, vol. 24, pp. 9-25, 1997.

[7] S. Benartzi, and R.H. Thaler, "Naive Diversification Strategies in Defined Contribution Saving Plans," The American Economic Review, vol. 91, pp. 79-98, 2001.

[8] D. Read, and G. Lowenstein, "Diversification Bias: Explaining the Discrepancy in Variety Seeking between Combined and Separated Choices," Journal of Experimental Psychology: Applied, vol. 1, pp. 34-49, 1995.

[9] I. Simonson, "The Effect of Purchase Quantity and Timing on Variety-Seeking Behavior," Journal of Marketing Research, vol. 27, pp. 150-162, 1990.

[10] G. Loomes, and R. Sudgen, "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty," The Economic Journal, vol. 92, pp. 805-824, 1982.

[11] J. Zweig, "How the Big Brains Invest at TIAA-CREF," Money, vol. 27, pp. 114-118, 1998.

[12] F. Black, M. Jensen, and M. Scholes, "The Capital Asset Pricing Model: Some Empirical Tests," in Studies in the Theory of Capital Markets, M. Jensen, Ed. New York: Praeger, 1972, pp. 79-121.

[13] R. Roll, "A Critique to the Asset Pricing Theory's Tests. Part I: On Past and Potential Testability of the Theory," Journal of Financial Economics, vol. 4, pp. 129-176, 1977.

[14] M. Gibbons, S. Ross, and J. Shanken, "A Test for the Efficiency of a Given Portfolio," Econometrica, vol. 57, pp. 1121-1152, 1989.

[15] J. Shanken, "Statistical Methods in Tests of Portfolio Efficiency: A Synthesis," in Handbook of Statistics, Vol. 14: Statistical Methods in Finance, G.S. Maddala, and C.R. Rao, Eds. Amsterdam: Elsevier Science, 1996, pp. 693-711.

[16] J. Campbell, A. Lo, and A.C. MacKinlay, The Econometrics of Financial Markets, Princeton: Princeton University Press, 1997.

[17] E. Sentana, "The Econometrics of Mean-variance Efficiency Tests: a Survey," Econometrics Journal, vol. 12, pp. C65-C101, 2009.

[18] A.C. MacKinlay, and M. Richardson, "Using Generalized Method of Moments to Test Mean-Variance Efficiency," The Journal of Finance, vol. 46, pp. 511-527, 1991.

[19] P. Chou, and G. Zhou, "Using Bootstrap to Test Portfolio Efficiency," Annals of Economics and Finance, vol. 7, pp. 217-249, 2006.

[20] S. Kandel, and R.F. Stambaugh, "Portfolio Inefficiency and the Cross-section of Expected Returns," The Journal of Finance, vol. 50, pp. 157-184, 1995.

[21] R.O. Michaud, Efficient Asset Management, Boston: Harvard Business School Press, 1998.

[22] J.-P. Brans, L'ingénierie de la décision. Elaboration d'instruments d'aide à la décision. Méthode PROMETHEE, Québec: Université Laval, 1982.

[23] J. Treynor, "Why Market-Valuation-Indifferent Indexing Works," Financial Analysts Journal, vol. 61, pp. 65-69, 2005.

[24] V. DeMiguel, L. Garlappi, and R. Uppal, "Optimal Versus Naïve Diversification: How Inefficient Is the 1/N Portfolio Strategy?" The Review of Financial Studies, vol. 22, pp. 1915-1953, 2009.

[25] H. Windcliff, and P. Boyle, "The 1/n Pension Investment Puzzle," North American Actuarial Journal, vol. 8, pp. 32-45, 2004.

[26] C. Erb, and C. Harvey, "The Strategic and Tactical Value of Commodity Futures," Financial Analysts Journal, vol. 62, pp. 69-97, 2006.

[27] G. Gorton, and K.G. Rouwenhorst, "Facts and Fantasies about Commodity Futures," Financial Analysts Journal, vol. 62, pp. 47-68, 2006.

[28] K. Cuthbertson, S. Hayley, N. Motson, and D. Nitzsche, "Diversification Return, Rebalancing Return and Volatility Pumping," Cass Business School, 2015.