“European indexed investments: how to identify the most efficient index”

AUTHORS Guido Abate https://orcid.org/0000-0002-7406-0650

ARTICLE INFO Guido Abate (2013). European indexed investments: how to identify the most efficient index. *Investment Management and Financial Innovations*, *10*(1-1)

RELEASED ON Friday, 19 April 2013

JOURNAL "Investment Management and Financial Innovations"

FOUNDER LLC “Consulting Publishing Company “Business Perspectives”

Guido Abate (Italy)

European indexed investments: how to identify the most efficient index

Abstract

This article provides a guideline for the evaluation of index efficiency, with a focus on the euro area stock markets. Efficiency plays a key role in defining the level of risk-adjusted performance reached by an investment, and it becomes gradually more important as the share of indexed investments increases within the portfolio.

The theoretical assumptions of the Modern Portfolio Theory (MPT) and of the Capital Asset Pricing Model (CAPM) provide the background for the statistical tests employed in the empirical analysis, carried out on the most important indices representative of the euro area stock markets. Indices have been selected according to quali-quantitative criteria, in order to include most of the construction techniques available on the market. The tests comprise robust techniques based upon simulation processes, such as the bootstrap, block-bootstrap and resampling.

The results of the analysis show the relevant influence of equal weighting schemes upon the level of efficiency, an outcome that is apparently in contrast with the CAPM and, to a lesser extent, the MPT, but that can be justified both by statistical and behavioral models.

Keywords: indexed investing, portfolio efficiency, efficiency tests, index construction.
JEL Classification: G11, G23.

Introduction

Indexed investments and efficiency. Indexed investments are commonplace in European portfolios, but the asset allocation process often overlooks the effects of the characteristics of the underlying indices. Once the investor has defined his asset allocation, in fact, the choice of the indexed financial instruments should take into account not only their structure, costs and liquidity, but also the indices tracked by the instruments themselves. The aim of this study is to evaluate the efficiency level of the most important indices representative of the Eurozone stock market, taking into account the influence of their construction techniques. In order to reach this scope, the empirical analysis presented in this article employs four different efficiency tests and measures, which have been applied on a sample of Eurozone indices. The joint use of these tests provides a new insight on the relative efficiency of each construction method, highlighting which characteristics should be favored by indexed investors in order to aim at higher risk-adjusted performances.

The concept of efficiency plays a key role in investments choice, but it should be underlined that financial theory has developed two distinct approaches to its definition. One of them can be described as "normative", i.e. a theory that describes a norm of behavior that investors should follow in portfolio selection, and the other one as "positive", i.e. an hypothesis about investors' aggregate behavior in real-life investments (Fabozzi, Gupta and Markowitz, 2002).

Markowitz (1952, 1959) has been the first author to propose a normative model: the Modern Portfolio Theory (MPT), which postulates that investors should diversify their investments and, at the same time, maximize their utility function, which is directly proportional to the expected return of the portfolio and inversely proportional to its volatility. Portfolios with the least variance, given a certain expected return, are defined "efficient portfolios" and lie on the "efficient frontier". MPT states that every investor should choose an asset allocation equal to that of an efficient portfolio, coincident with the tangency point between his utility function and the efficient frontier on the mean-standard deviation plane.

Following this prescriptive approach, Sharpe (1964), Lintner (1965) and Mossin (1966) developed the positive theory of the Capital Asset Pricing Model (CAPM), which instead assumes to be a realistic description of investors' behavior. In this framework, then, investors should not simply follow the CAPM because it is a rational choice, but are also assumed to apply it in all their investments, at least at the aggregate level.

According to the CAPM, if its underlying hypotheses are met, all the investors hold the same portfolio of every investible risky asset, 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 follows that the market portfolio must be capitalization weighted. The investment weights in each asset must be, therefore, strictly positive and proportional to the ratio of the asset's capitalization to the total capitalization of the universe of investible assets.

It can be concluded, then, that the concept of optimal, i.e. the most efficient, portfolio can vary 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. In order to take into account this discrepancy, this study employs different tests, based upon the MPT and the CAPM frameworks.

1. Measures of portfolio efficiency

Following Black, Jensen and Scholes (1972), Roll (1979) has set the theoretical background for CAPM-based efficiency tests: once the empirical validity of the CAPM is assumed as verified, the only testable hypothesis is the efficiency of a proxy of the market portfolio.

Given these premises, Gibbons, Ross and Shanken (1989) have developed a multivariate test, under the null hypothesis $H_0 : \hat{\alpha} = 0$, with $\hat{\alpha}$ being an 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 ($N \times T$ matrix $\hat{\varepsilon}$) are distributed as a Normal with mean zero and diagonal covariance matrix $\hat{\Sigma}$ (dimensions $N \times N$), since the residuals are uncorrelated by hypothesis. The normality hypothesis, imposed by the authors, would not be strictly necessary in order to evaluate the test statistic, but Shanken (1996) has underlined its sensibility to conditional heteroskedasticity[1] of $\hat{\varepsilon}$.

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

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

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

However, the Wald test suffers from a practical shortcoming: only its asymptotical distribution is known. In order to overcome this problem, Gibbons, Ross and Shanken have applied the following correction to the 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 employ 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 noncentral when H_0 cannot be accepted, because its noncentrality parameter is zero when $\hat{\alpha} = 0$. The Gibbons, Ross and Shanken (GRS) test is then:

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

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

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

Thus the GRS test can be reformulated using measures typical of performance evaluation in the asset management industry:

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

According to the F distribution, the probability that 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 P's components.

The observation about the Sharpe ratio is in accordance with the CAPM, because the Sharpe ratio of P is the slope of the capital allocation line passing through P and the higher is the slope, the higher is the degree of efficiency of an asset. The significance of the appraisal ratio is also linked to CAPM theory. Given that the presence of intercepts, be they positive or negative, is not envisaged by this model, significant appraisal ratios would be in contrast with the notion of efficiency of P. Recalling

[1] In other words, residual volatility is time-varying and shows dependence on the excess returns of P.

that AR_i is the ratio between the intercept α_i and the standard deviation of residuals in the CAPM regression, a value of AR_i near zero implies that either the intercept is small or that it is statistically not significant due to its volatility.

Empirical analyses carried out on the GRS test by Gibbons, Ross and Shanken (1989), Campbell, Lo and MacKinlay (1997) and Sentana (2009) 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 (1997) suggest to keep N not larger than 10.

Under the assumption that residuals are i.i.d., Gibbons, Ross and Shanken (1989) show that:

$$SR_M^2 = \hat{\mu}'\hat{V}^{-1}\hat{\mu} \tag{7}$$

and

$$\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha} = \sum_{i=1}^{N} AR_i^2 = SR_M^2 - SR_P^2. \tag{8}$$

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

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

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

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

This new formulation shows that the GRS-stat is proportional to the ratio of the lengths of the hypotenuses of two right-angled triangles (see Figure 1). The ratio converges to 1, and thus GRS to zero, as the Sharpe ratio of P approaches the Sharpe ratio of the market portfolio. Given that the GRS test is a small-sample adjustment of the Wald test, the same interpretation can also be applied to the latter.

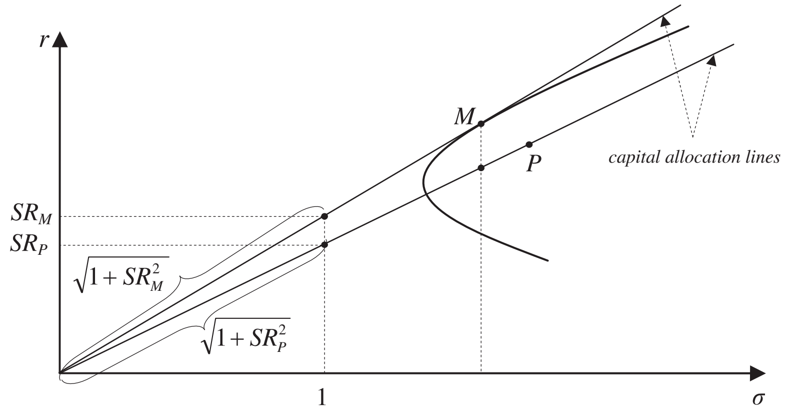

Fig. 1. Graphical representation of GRS and Wald tests

The GRS test is based upon a finite sample of data, which is an advantage, given that no empirical analysis can be carried out on samples of infinite length, but suffers from its assumption of normality of returns. In order to model the presence of the heteroskedasticity of residuals, it is possible to apply the Generalized Method of Moments test (MacKinlay and Richardson, 1991). The most common notation used for this test is the Wald-like one, as reported by Chou and Zhou (2006):

$$J_1 = T\hat{\alpha}'[\eta\,[D_T' S_T^{-1} D_T]^{-1}\,\eta']\hat{\alpha} \tag{11}$$

$\eta = I_N \otimes [1, 0]$ is the matrix composed only of 1s and 0s; $\Omega_T = [D_T' S_T^{-1} D_T]^{-1}\eta']^{-1}$ is the covariance matrix of the regression parameters; $D_T = -\frac{1}{T}\sum_{t=1}^{T}[I_N \otimes Z_t Z_t']$ is the symmetrical matrix made of square submatrices, lined along the main diagonal, containing the descriptive statistics of P; $S_T = \frac{1}{T}\sum_{t=1}^{T}[\varepsilon_t\,\varepsilon_t' \otimes Z_t Z_t']$ is the spectral density matrix; $Z = [1, r_P']$ is the matrix $2\times T$ with only 1s on its first row and the excess returns of P on its second row.

The GMM test, in other words, is a Wald-like test, in which covariances of residuals are correlated with the returns of the components of index P.

A shortcoming of this test is that the distribution of J_1 is known only asymptotically, and thus it is necessary to utilize sampling techniques. This solution, on the one hand, can lead to sub-optimal results in case of serial correlation of residuals. Returns, in fact, show cross-section correlation and time series correlation and thus, if resampling is to be applied in each time t of length equal to one, only the former type of correlation can be simulated. In order to overcome these limitations, it is necessary to employ heuristic techniques such as the block bootstrap.

The 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 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 importance 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 can be applied only in the CAPM framework, the measure of relative efficiency by Kandeland Stambaugh (1996) follows an approach closer to the MPT: it evaluates relative efficiency with respects to the efficient frontier. It compares the excess return of *P* to that of *x*, 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:

$$\psi_P = \frac{\mu_P - \mu_g}{\mu_x - \mu_g}. \quad (12)$$

The highest 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.

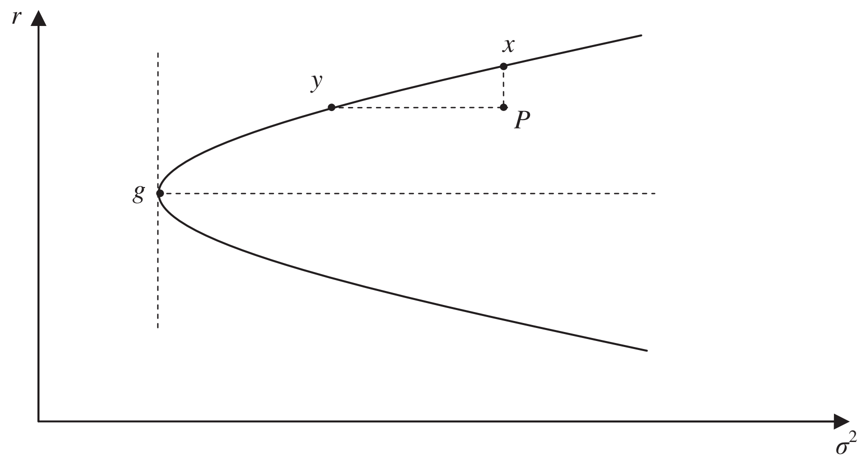

Fig. 2. Kandel and Stambaugh's efficiency measure

This measure of efficiency was conceived two years before the introduction of the technique of resampling by Michaud (1998). This simulation method attempts to overcome the problems of error maximization typical of the usual construction of the efficient frontier, which causes excessive 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.

2. Empirical analysis

2.1. Composition of the sample. The focus on European investments of this empirical study has led to the selection of indices representative of the euro area stock markets. The choice of indices composed of securities denominated in the same currency is very important, for the scope of this analysis, because it allows to avoid the influence of the movements of exchange rates on the returns of indices and their components.

Thanks to the diversification of construction techniques shown by indices representative of the euro area stock market, it is possible to evaluate the influence of each one on the level of efficiency.

The list of indices and their characteristics is reported in Table 1. The indices chosen are among the benchmarks most utilized by practitioners, and they have to comply with the following criteria:

- 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 section.

The chosen proxy of the risk-free rate is the return of 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.

Table 1. Eurozone stock indices

Index	Weighting	Reinvestment of income
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

2.2. Beta-sorted portfolios. The tests employed in this analysis are subject to potential biases depending upon sample size, 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 is N, in fact, the higher is the probability of including assets with extreme in-sample performances that could not repeat themselves also out-of-sample.

Given these premises, it has been necessary to solve the problem of reducing the N/T ratio by limiting the numerator. Black, Jensen and Scholes (1972) provide an aggregation method that has become standard in scientific literature: 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 excess returns of the N assets on the excess 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 this study, following Gibbons, Ross and Shanken (1989) and Campbell, Lo and MacKinlay (1997), Q has been set as being equal to ten.

Each beta-sorted portfolio is an equal weight average of the assets within a quantile, thus the slope of the portfolio is equal to the systematic risk of the assets of the portfolio and its intercept is equal to the average intercept.

In order to implement the analysis it has been necessary to select the time series of the returns of the components of an index that can be regarded as representative of the euro area stock market. Such an index has been identified in the Euro Stoxx 50, a free-float weighted average of the 50 supersector leaders from the Euro Stoxx index.

Given that the index is subject to quarterly revisions, both with regard to its components and to their weights, the following procedure has been implemented:

- 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, of which the first 60 months (in-sample) have 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 returns of the components that have at least 24 months in-sample and two out-of-sample have been regressed on the index;
- stocks have been ordered according to their beta and aggregated into ten beta-sorted portfolios;
- the monthly return of each beta-sorted portfolio is the arithmetic average of the returns of its components in the out-of-sample months, covering the period January 2003-December 2010.

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 betas, if it occurs, even though it includes a temporal lag of three months at worst[1]. Thus, with regard to the variable composition of the beta-sorted portfolios, their risk profile is kept constant, because stocks are transferred to other portfolios when their beta migrates to another quantile.

2.3. Methodology of the empirical analysis. The out of sample returns of beta-sorted portfolios have been used as a panel of components for all the indices of the "Eurozone stock market" asset class, 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 composition of every index, is founded also on theoretical bases: in order to identify the most efficient construction

[1] Black, Jensen and Scholes (1972) reconstruct their beta-sorted portfolios yearly, thus with a lower precision. It should be noted, however, that two in-sample periods are overlapping for 57 months and thus the variability of betas is somewhat limited, given that it can be ascribed only to the shocks that happened in the three non-overlapping months.

methods for a market benchmark, it is useful to compare all the stock indices to the same sample of assets.

Possible deviations of returns from the Gaussian distribution have been analyzed through the Jarque-Bera test (Table 2)[1].

Table 2. Descriptive statistics of monthly excess returns

Index	Mean	Standard deviation	Asymmetry	Kurtosis	JB stat	p-value (1)	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

Note: (1) Rounded to 0,10% by Matlab if tending to zero.

Given the outcome of the JB tests, it can be inferred that the GMM test will be very significant, because it is the only one which rejects the hypothesis of normality. The GRS test has been used in its original notation and not in its decomposition into the Sharpe ratio and appraisal ratio, because this latter is too biased in presence of correlated residuals. The Wald test has been implemented through a bootstrap simulation, in order to model an empirical distribution and thus overcome the problems linked to limited samples. For each index and the ten beta-sorted portfolios, 10,000 scenarios have been simulated through a bootstrap simulation (with replacement).

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 also 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. The grouping of stocks into portfolios, in fact, limits the impact of outlier returns, further reducing error maximization. In detail, the procedure 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, 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 $J1$ 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 the excess returns in the indices: in their large majority (14 out of 15), in fact, autocorrelation is statistically significant up to the fourth lag. The use of blocks with a length of six periods is a compromise that allows to capture, within each block, an autocorrelation of:

- first order, in five periods out of six;
- fourth order, in two periods out of six.

2.4. Results of the empirical analysis. Despite the use of tests with different theoretical bases and implementation techniques, the evaluation of efficiency has reached results that are substantially in agreement with each other and are useful to identify optimal construction techniques. Table 3 shows the results of the tests and, where available, their percentage of p-value, i.e. of the probability that the hypothesis of efficiency cannot be rejected.

[1] This statistical measure 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.

Table 3. Efficiency levels

Index	Kandel & Stambaugh	GRS		Wald		GMM	
		GRS	p-value	WT	p-value (1)	J1	p-value (1)
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%

Note: (1) P-value estimated through the bootstrap of 10,000 scenarios.

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 time of strong turbulence. Along with this overall judgment, we can 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) by Brans (1982). 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 (negative) outranking flow which expresses how much an alternative is outranking (outranked by) all the others. The calculation of these flows requires, as a first step, to define the dummy variable of the following preference function:

$$P_k(i, j) = \begin{cases} 0 \text{ if } d_k \leq 0 \\ 1 \text{ if } d_k > 0, \end{cases} \tag{13}$$

where i and j are two stock market indices and d_k is the distance between the rankings of i and j according to the efficiency measure k. Then the aggregate preference index $\pi(i, j)$, i.e. an average (weight $w_k = 1/4$) of the preference functions of the pair of indices i and j for each of the four efficiency measures, is calculated: $\pi(i, j) = \sum_{k=1}^{4} P_k(i, j) w_k$.

The last step is the summation of the preference indices $\pi(i, j)$ across the pairs of the sample, where I is the total number of stock market indices present in the empirical analysis, in order to determine the outranking flows:

$$\begin{aligned} \varphi^{+}(i) &= \frac{1}{I-1} \sum\nolimits_{\forall j \neq i} \pi(i, j), \\ \varphi^{-}(i) &= \frac{1}{I-1} \sum\nolimits_{\forall j \neq i} \pi(j, i). \end{aligned} \tag{14}$$

As reported in Table 4, the indices representative of the most relevant companies quoted on the markets of the European Monetary Union show a higher degree of efficiency when they are total return, especially when income, i.e. dividends, is reinvested gross of taxes. This outcome, on the other hand, does not repeat itself mechanically: the choice of an index, then, 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 the factors in index construction that affect positively the level of efficiency are equal weighting, diversification and reinvestment of income. The latter, in accordance with the reality of financial markets, can be tracked by investors only when they take into account the tax rates on income.

Table 4. Net outranking flows of the efficiency measures of Eurozone stock indices

	Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	$\varphi^+(i)$	$\varphi(i)$	Rank
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*	8th
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*	9th
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*	15th
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*	2nd
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*	11th
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*	1st
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*	4th
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*	10th
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*	7th
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*	13th
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*	3rd
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*	6th
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*	12th
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*	14th
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*	5th
Negative outranking flow $\varphi^-(i)$		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			

Conclusions

The results of the empirical analysis suggest that equal weighting offers a superior risk-adjusted return if compared to traditional cap-weighting. The causes of this phenomenon can be traced both to behavioral and statistical reasons. If we accept the former interpretation, we may conclude that the influence of investors who follow the 1/*N* heuristic is such that they are able to shape the structure of the market and thus to turn equal weighting into the most efficient construction method for indices.

The statistical interpretation can, however, be divided into two coexisting theories. Treynor (2005) has underlined how the presence of "noise" in prices causes an excess weighting of overpriced stocks (and, conversely, an underweighting of underpriced ones) in cap-weighted indices. These indices, then, are 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 weighting has been formulated by DeMiguel, Garlappi and Uppal (2009). These authors 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. 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 that equal weighting, which ignores statistical measures for portfolio construction, is the most efficient technique[1]. Windcliff and Boyle (2004), moreover, had already noticed, even though they had not measured, this phenomenon, explicitly linking the 1/*N* heuristic to the minimization of estimation error.

The results of the empirical analysis are thus in accordance with these theoretical explanations that,

[1] It can be proven that, when return moments are not known, the tangency portfolio is equally weighted.

indeed, can be regarded as two faces of the same coin. Either the 1/*N* heuristic is "irrational", but is followed by so many investors that it influences the markets, thus becoming a rational investment rule, or estimation error is so relevant that no other asset allocation is more efficient than equal weighting, and thus the 1/*N* heuristic is the most rational approach. In both cases, equal weighting plays a key role in indices efficiency, a counterintuitive outcome that should make investors aware of the importance of measuring the efficiency of construction techniques when they are going to select an indexed investment.

References

1. Black, F., M. Jensen, M. Scholes (1972). The Capital Asset Pricing Model: Some Empirical Tests, M. Jensen (editor), *Studies in the Theory of Capital Markets*, New York: Praeger, pp. 79-121.
2. Brans, J.-P. (1982). L'ingénierie de la décision. Elaboration d'instruments d'aide à la décision. Méthode PROMETHEE, Colloque d'Aide à la Décision, Québec: Université Laval, pp. 183-213.
3. Campbell, J., A. Lo, A.C. MacKinlay (1997). *The Econometrics of Financial Markets*, Princeton: Princeton University Press, 611 pp.
4. Chou, P., G. Zhou (2006). Using Bootstrap to Test Portfolio Efficiency, *Annals of Economics and Finance*, 7 (2), pp. 217-249.
5. DeMiguel, V., L. Garlappi, R. Uppal (2009). Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies*, 22 (5), pp. 1915-1953.
6. Fabozzi, F., F. Gupta, H. Markowitz (2002). The Legacy of Modern Portfolio Theory, *The Journal of Investing*, 11 (3), pp. 7-22.
7. Gibbons, M., S. Ross, J. Shanken (1989). A Test for the Efficiency of a Given Portfolio, *Econometrica*, 57 (5), pp. 1121-1152.
8. Kandel, S., R. Stambaugh (1995). Portfolio Inefficiency and the Cross-section of Expected Returns, *The Journal of Finance*, 50 (1), pp. 157-184.
9. Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *Review of Economics and Statistics*, 47 (1), pp. 13-37.
10. MacKinlay, A.C., M. Richardson (1991). Using Generalized Method of Moments to Test Mean-Variance Efficiency, *The Journal of Finance*, 46 (2), pp. 511-527.
11. Markowitz, H. (1952). Portfolio Selection, *The Journal of Finance*, 7 (1), pp. 77-91.
12. Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*, New York: Wiley, 344 pp.
13. Michaud, R. (1998). *Efficient Asset Management*, Boston: Harvard Business School Press, 152 pp.
14. Mossin, J. (1966). Equilibrium in a Capital Asset Market, *Econometrica*, 34 (4), pp. 768-783.
15. Roll, R. (1977). A Critique to the Asset Pricing Theory's Tests. Part I: On Past and Potential Testability of the Theory, *Journal of Financial Economics*, 4 (2), pp. 129-176.
16. Sentana, E. (2009). The Econometrics of Mean-variance Efficiency Tests: a Survey, *Econometrics Journal*, 12 (3), pp. C65-C101.
17. Shanken, J. (1996). Statistical Methods in Tests of Portfolio Efficiency: A Synthesis, In G. S. Maddala, C.R. Rao (editors) Handbook of Statistics, Vol. 14, *Statistical Methods in Finance*, Amsterdam: Elsevier Science, pp. 693-711.
18. Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium, *The Journal of Finance*, 19 (3), pp. 425-442.
19. Treynor, J. (2005). Why Market-Valuation-Indifferent Indexing Works, *Financial Analysts Journal*, 61 (5), pp. 65-69.
20. Windcliff, H., P. Boyle (2004). The 1/n Pension Investment Puzzle, *North American Actuarial Journal*, 8 (3), pp. 32-45.