

La valutazione dell'efficienza degli indici: una verifica empirica sul mercato azionario statunitense

GUIDO ABATE*

Abstract

The study of the efficiency of stock indices has increased its importance in the asset management industry due to the widespread use of benchmarks and indexed investment products.

This paper evaluates the efficiency level of the most commonly used US stock indices. The empirical study employs four measures of efficiency: the GRS, a small-sample multivariate test based upon the CAPM; the large-sample Wald test, implemented through a bootstrap simulation; the Generalized Method of Moments test, which has been applied in a non-Gaussian framework through a block bootstrap simulation; and Kandel and Stambaugh's relative efficiency measure.

The empirical results provide strong evidence of the superior efficiency of equally weighted indices. These findings are interpreted according to the scientific literature available, analyzing the different theoretical and empirical causes that have been proposed.

ADEIMF Keywords: Asset Allocation, Behavioural Finance, Performance Evaluation.

JEL Codes: G11, G14, G15.

* Dipartimento di Economia Aziendale, Università degli Studi di Brescia, C.da S. Chiara, 50 – 25122 Brescia (BS). E-mail: abate@eco.unibs.it.

Indice

1. Introduzione	1
2. La misura dell'efficienza: analisi della letteratura	1
2.1. <i>Le prime prove empiriche e il test GRS</i>	<i>1</i>
2.2. <i>Il test GRS: interpretazione geometrico-finanziaria.....</i>	<i>4</i>
2.3. <i>Il test GMM</i>	<i>5</i>
2.4. <i>La misura di efficienza relativa di Kandel e Stambaugh.....</i>	<i>7</i>
3. L'analisi empirica	10
3.1. <i>La composizione del campione</i>	<i>10</i>
3.2. <i>La costruzione dei portafogli beta-sorted.....</i>	<i>11</i>
3.3. <i>La metodologia applicativa dell'analisi empirica</i>	<i>12</i>
3.4. <i>I risultati dell'analisi empirica</i>	<i>14</i>
4. Conclusioni	18
Bibliografia	20

1. Introduzione

Gli indici del mercato azionario hanno subito, nel corso degli anni, profonde trasformazioni, non solo in termini di tecniche di calcolo, ma anche e soprattutto di finalità a essi attribuite. Come rilevato in Basile (2002), da strumenti meramente informativi, concepiti per rappresentare in un unico valore l'andamento di un intero mercato o di un singolo settore economico, essi sono divenuti metri di valutazione, ossia *benchmark*, dell'operatività degli *asset manager* da parte sia degli studiosi che della platea degli investitori. Da ultimo, sono stati trasformati in oggetti essi stessi di investimento, tramite l'utilizzo di strumenti finanziari quali, a titolo d'esempio, i fondi comuni d'investimento indicizzati e gli *exchange traded fund*.¹

Lo scopo che si propone questo studio è quindi quello di analizzare quali tra i principali indici del mercato azionario statunitense risultino più efficienti, in modo da poter fornire agli investitori e agli *asset manager* indicazioni sulle caratteristiche che dovrebbero essere rispettate dai *benchmark* selezionati per i loro investimenti.

L'efficienza assume un elevato rilievo applicativo sia nella sua accezione fornita dalla Modern Portfolio Theory (MPT) sia secondo quella formulata dal Capital Asset Pricing Model (CAPM). Markowitz (1952) considera come efficienti tutti quei portafogli che si collocano sulla frontiera a varianza minima: sono quindi molteplici le combinazioni di rischio-rendimento ottimali e ogni investitore, sulla base della sua avversione al rischio, può allocare le proprie risorse finanziarie in quella che massimizza la sua funzione di utilità. Nel CAPM, sviluppato da Sharpe (1964), Lintner (1965) e Mossin (1966), al contrario, un portafoglio è efficiente non solo in quanto posto sulla frontiera di Markowitz, ma anche perché i profili di rischio-rendimento di tutti i titoli sono esplicabili tramite una relazione lineare con tale portafoglio, definito "di mercato". Diverse propensioni al rischio possono essere soddisfatte, quindi, tramite investimenti in tale portafoglio caratterizzati da differenti gradi di *leverage*.

Appare quindi utile il ricorso a *test* di efficienza concepiti secondo i paradigmi di entrambi i modelli, anche alla luce delle diverse finalità che possono essere attribuite ai due concetti di efficienza. Mentre infatti la MPT è una teoria normativa, volta a indirizzare gli investitori verso allocazioni razionali e ottimali, il CAPM è una teoria positiva, che postula l'effettivo verificarsi delle proprie tesi nei processi di formazione dei prezzi dei titoli quotati nei mercati finanziari.

I dettagli delle tecniche di valutazione dell'efficienza impiegate in questo studio sono esposti nella sezione 2. Sono stati utilizzati i *test* GRS e Wald, che rientrano pienamente nell'orizzonte concettuale del CAPM, il *test* GMM, che considera rendimenti non gaussiani e autocorrelati, e la misura di Kandel e Stambaugh, impiegata invece nell'ambito della MPT.

La sezione 3, strutturata come il prosieguo applicativo della precedente, è dedicata interamente alla esposizione delle prove empiriche effettuate, tramite la programmazione di appositi *script* di Matlab, su di un campione complessivo di 22 indici azionari e di dieci portafogli *beta-sorted*.

Nelle conclusioni il risultato empirico viene analizzato sulla scorta delle diverse teorie finanziarie, delle quali sono esplorate le interazioni ed è evidenziata la complementarità.

2. La misura dell'efficienza: analisi della letteratura

2.1. Le prime prove empiriche e il test GRS

L'articolo del 1972 di Black, Jensen e Scholes,² antecedente alla critica di Roll³ e anzi spunto di riflessione per quest'ultimo autore, ha, seppure in modo implicito, esposto un primo esempio di *test*

¹ Oltre che, ovviamente, come oggetto di speculazione, arbitraggio e copertura tramite il ricorso a molteplici strumenti derivati.

² Black et al. (1972).

di efficienza. Si tratta di una procedura “implicita” perché lo scopo dichiarato dagli autori è quello di appurare la validità empirica del CAPM e, per ottenere questo risultato, è stato necessario costruire una *proxy* dell’indice di mercato. Tale prova, un *t-test* univariato, consiste nell’accertare la significatività statistica dell’intercetta α_i nell’equazione:

$$E(R_i) - R_f = \alpha_i + \beta_i [E(R_p) - R_f]$$

Affinché si possa ritenere empiricamente fondato il CAPM, è necessario che le intercette di tutti i titoli *i*-esimi siano nulle.

Alla luce della critica di Roll, secondo la quale è impossibile verificare la validità del CAPM data la non osservabilità del portafoglio di mercato, è possibile invertire la logica dei *test* del CAPM. In questo caso, il modello è considerato come verificato per ipotesi e l’oggetto dell’analisi sono gli indici di borsa via via utilizzati come *proxy* del *market portfolio*.

Gibbons, Ross e Shanken hanno quindi proposto un *test* multivariato, nel quale l’ipotesi nulla oggetto di valutazione è $H_0: \hat{\alpha} = \mathbf{0}$, con $\hat{\alpha}$ vettore $N \times 1$ delle intercette delle regressioni sull’extrarendimento⁴ di P degli extrarendimenti della serie storica di durata T degli N titoli componenti l’indice P :⁵

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i r_{p,t} + \hat{\varepsilon}_{i,t}$$

I residui $\hat{\varepsilon}$ (matrice $N \times T$) si distribuiscono secondo una normale con media zero e matrice di varianza-covarianza $\hat{\Sigma}$ di dimensioni $N \times N$, ipotizzata diagonale: ne consegue che i residui devono essere scorrelati tra di loro. L’ipotesi di normalità, per quanto imposta dagli autori, non sarebbe strettamente necessaria ai fini della valutazione del risultato del *test*, ma studi successivi⁶ hanno evidenziato una sua sensibilità alla presenza di eteroschedasticità condizionale⁷ degli $\hat{\varepsilon}$.

Allo scopo di valutare la significatività delle intercette è possibile considerare l’impiego di un *test* di Wald nella seguente formulazione:

$$WT = T \left[1 + \frac{\hat{\mu}_P^2}{\hat{\sigma}_P^2} \right]^{-1} \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha}$$

$1 \times N$ $N \times N$ $N \times 1$

Indicando con $\hat{\mu}_P$ e $\hat{\sigma}_P^2$ rispettivamente la media e la varianza degli extra-rendimenti del portafoglio P di cui si desidera valutare l’efficienza.

Il problema di questo approccio consiste nell’indeterminatezza della distribuzione della statistica WT nel caso in cui il campione sia finito. Per ovviare a tale seria limitazione, è però possibile ricorrere al teorema di Muirhead (1982), grazie al quale è possibile dimostrare che se H_0 è vera allora:

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

In altri termini, è possibile effettuare una trasformazione lineare del *test* di Wald che si distribuisca come una F con gradi di libertà N al numeratore e $T - N - 1$ al denominatore. La distribuzione F risulterebbe però non centrata nel caso in cui l’ipotesi nulla non fosse accolta, dato che il suo parametro di non centralità si annulla per $\hat{\alpha} = \mathbf{0}$. Si tratta comunque di una distribuzione nota che permette il calcolo dei quantili corrispondenti ai valori della statistica in oggetto.⁸

³ Roll (1977).

⁴ Con extrarendimento $r_{i,t}$ di un titolo *i*-esimo si intende la differenza tra il suo rendimento $R_{i,t}$ nel periodo *t*-esimo e il tasso *risk-free* R_f .

⁵ Gibbons et al. (1989).

⁶ Per una rassegna di tali studi, si rimanda a Shanken (1996).

⁷ In altri termini, la volatilità degli $\varepsilon_{i,t}$ dato $r_{p,t}$ varia dal tempo *t* al tempo *t*+1 e mostra una dipendenza da $r_{p,t}$.

⁸ Si noti che gli autori, nel loro articolo originale, hanno dimostrato la derivazione del loro *test* partendo dal *test* T^2 di Hotelling (ovvero una generalizzazione multivariata del *t-test* utilizzato in Black et al., 1972); si è preferito ricorrere

Apportando le modifiche indicate al *test* di Wald si può quindi definire il *test* di Gibbons, Ross e Shanken:

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

$1 \times N$ $N \times N$ $N \times 1$

La formula del *test* GRS contiene alcuni fattori di agevole interpretazione economica. Si può infatti notare che il rapporto tra il quadrato della media degli extrarendimenti di P e la sua varianza altro non è che il quadrato dell'indice di Sharpe, indicato con SR_P .⁹

Meno intuitivo è il valore rappresentato dalla forma quadratica $\hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha}$: tale prodotto matriciale è infatti la sommatoria dei quadrati degli alfa divisi per le varianze dei residui (sempre e soltanto nell'ipotesi che questi ultimi siano indipendenti). Si può ora ricordare che una misura di *risk-adjusted* performance marginale è l'*appraisal ratio*, definito come il rapporto tra l'intercetta della regressione dei rendimenti di un titolo rispetto al *benchmark* e la volatilità dei residui. In conclusione:

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

Il *test* GRS può quindi essere riscritto nella seguente notazione, non più strettamente econometrica bensì basata sugli strumenti valutativi tipici dell'*asset management*:

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

Sulla base della sua distribuzione F , sappiamo che la probabilità che il portafoglio P sia efficiente cresce all'approssimarsi a zero del valore della statistica GRS. Quindi, data la formulazione precedente, possiamo affermare che il grado di efficienza di P è:

- direttamente proporzionale al quadrato del suo Sharpe *ratio*;
- inversamente proporzionale alla somma dei quadrati degli *appraisal ratio* dei componenti.

La considerazione sull'indice di Sharpe appare di per sé evidente: secondo il CAPM, il portafoglio di tangenza di una frontiera efficiente giace sulla Capital Market Line, ovvero la *capital allocation line* con il coefficiente angolare più elevato (nel piano cartesiano media-deviazione *standard*). Poiché tale coefficiente altro non è che l'indice di Sharpe, è graficamente dimostrato che maggiore è il suo valore, maggiore è il livello di efficienza dell'attività finanziaria cui si riferisce.

Anche la significatività dell'*appraisal ratio* è legata alla formulazione stessa del CAPM: poiché la presenza di intercette alfa non è contemplata dalla sua equazione fondamentale, l'osservazione di AR_i elevati sarebbe contraria all'efficienza dell'indice sottoposto ad analisi. D'altro canto, l'*appraisal ratio* non è calcolato solamente sulla base dell'alfa, ma anche della deviazione *standard* dei residui: al crescere di quest'ultima, il rapporto tende a zero, evidenziando quindi una maggiore efficienza. Ciò è dovuto al fatto che ogni misura perde di significatività all'aumentare della sua volatilità e quindi valori anche elevati delle intercette potrebbero essere controbilanciati dalla loro spiccata aleatorietà.

E' d'altronde necessario evidenziare che l'*appraisal ratio* è introdotto nella formulazione del *test* GRS in forma quadratica: questo implica che anche intercette significativamente negative sarebbero indicatori di inefficienza dell'indice P . Tale conclusione è apparentemente controintuitiva, e

all'esposizione di Campbell et al. (1997) poiché nel presente studio il *test* di Wald sarà anch'esso oggetto di una prova empirica, basata sull'impiego della tecnica di simulazione *bootstrap*.

⁹ Si ricorda che lo Sharpe *ratio* rappresenta il rapporto tra la media degli extrarendimenti di un generico titolo i , al numeratore, e la loro deviazione *standard* al denominatore. Lo scopo di tale indicatore è quello di indicare sinteticamente quale livello di extrarendimento sia offerto dal titolo i per ciascuna unità di rischio che esso comporta per l'investitore.

richiede quindi un approfondimento. Sulla scorta del CAPM, tutti i rendimenti medi attesi dei titoli sono linearmente dipendenti dal rendimento del portafoglio di mercato M , il quale a sua volta è efficiente e si compone di tutte le attività investibili, ciascuna con coefficiente di ponderazione positivo. Ciò porta a rigettare l'efficienza di un indice il quale, se utilizzato come sostituto di M , generi intercette di qualsiasi entità, anche inferiore a zero.

Le analisi empiriche condotte sul *test* GRS¹⁰ hanno mostrato come la sua potenza¹¹ sia sensibile alle dimensioni del campione. In particolare, al crescere di T la potenza è sempre crescente, mentre declina fortemente al crescere di N : si ritiene quindi che sia buona norma ricondurne la dimensione entro un massimo di 10 elementi¹² e, in ogni caso, entro un rapporto di circa uno a tre rispetto a T .¹³

2.2. Il test GRS: interpretazione geometrico-finanziaria

Della statistica GRS è inoltre possibile derivare un'interpretazione in termini di geometria cartesiana applicata al piano media-deviazione *standard*. Allo scopo è necessario iniziare dalla costruzione dei portafogli giacenti sulla frontiera efficiente tramite la nota tecnica che consiste nella minimizzazione della varianza di un generico portafoglio k dato uno specifico livello di extra-rendimento atteso μ_k^* .

La procedura seguita da Gibbons, Ross e Shanken non segue la medesima logica di Markowitz, bensì l'ipotesi esposta da Lintner¹⁴, secondo la quale è possibile considerare la facoltà di vendere allo scoperto i titoli. Di conseguenza, il problema non deve più sottostare ai vincoli relativi ai pesi dei componenti dei portafogli di frontiera e può essere risolto analiticamente come una ottimizzazione quadratica. All'interno di quest'ottica teorica, ricorrendo al moltiplicatore di Lagrange e con alcuni passaggi algebrici si può dimostrare che il quadrato dello Sharpe *ratio* del portafoglio di tangenza è:

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

E' inoltre dimostrabile che:¹⁵

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

In altri termini, si espone nuovamente la constatazione che, in presenza del vero portafoglio di mercato, le intercette delle regressioni basate sull'equazione del CAPM sono pari a zero e, al contempo, che tanto maggiore è la distanza tra lo Sharpe *ratio* del portafoglio P rispetto a quello del portafoglio di mercato, tanto maggiore è la sua inefficienza.

Si noti, inoltre, che entrambe le equivalenze qui esposte sono esatte solo e solamente nel caso in cui siano verificate le ipotesi in termini di normalità dei rendimenti e, soprattutto, di indipendenza dei residui. In caso contrario, come solitamente avviene quando si opera con dati reali, le diverse formulazioni non portano a risultati coincidenti.

Sostituendo nella formula originaria del *test* GRS i valori equivalenti qui esposti, si ottiene:

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

L'ultimo fattore può essere anche visto in questa diversa forma:

¹⁰ Gibbons et al. (1989); Campbell et al. (1997); Sentana (2009).

¹¹ La potenza di un *test* è la probabilità di non incorrere in errori del secondo tipo: è quindi la probabilità di rigettare correttamente l'ipotesi H_0 come falsa quando infatti essa non è verificata.

¹² Campbell et al. (1997).

¹³ Gibbons et al. (1989).

¹⁴ Lintner (1965). Si noti che Sharpe segue invece esplicitamente l'impostazione *long-only* di Markowitz (Sharpe, 1964).

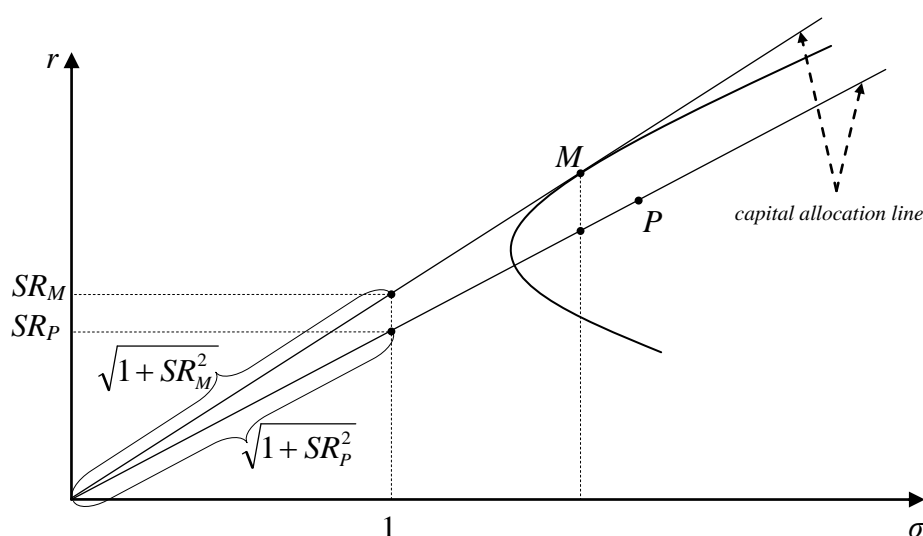
¹⁵ Gibbons et al. (1989), Grinold (1992).

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

Questa diversa notazione della formula è alla base della sua interpretazione geometrica (figura 1): il numeratore e il denominatore altro non sono che la lunghezza, secondo il teorema di Pitagora, dell'ipotenusa di due triangoli rettangoli. Tali triangoli sono composti da un cateto orizzontale identico per entrambi formato dall'asse delle ascisse fino al punto (1, 0) e da due cateti verticali rappresentati dai segmenti paralleli all'asse delle ordinate aventi lunghezza pari agli Sharpe ratio di M e P .

Quando la pendenza della *capital allocation line* di P si approssima a quella di M , il rapporto delle ipotenuse tende ad uno e quindi il valore di GRS si avvicina a zero, corrispondente nella distribuzione F alla massima probabilità che l'efficienza di P sia accettabile statisticamente.

Figura 1: Rappresentazione geometrica del test GRS



Fonte: Gibbons et al. (1989).

2.3. Il test GMM

Il test GRS rientra nella categoria dei test cosiddetti *small sample*, ovvero basati su un campione finito di dati. Tali metodi di valutazione dell'efficienza sono particolarmente utili nell'ipotesi di normalità dei rendimenti, in quanto in tal caso i loro valori assumono una distribuzione nota, ma possono risultare inadeguati se tale restrizione viene superata. In questo caso è invece necessario ricorrere a test per i quali la distribuzione dei valori della statistica non è nota e, di conseguenza, è necessario ricorrere a tecniche di ricampionamento per definire la funzione di densità del test. Per questa ragione, si opera in un'ottica *large sample*.

Nel caso di eteroschedasticità è possibile ricorrere al test GMM, proposto da MacKinlay e Richardson¹⁶ sulla base delle precedenti ricerche econometriche condotte da Hansen sul *generalized method of moments*.¹⁷

Il test GMM può essere effettuato utilizzando diverse notazioni, in particolare esso può assumere la forma di un test di Wald, ed è definito come:¹⁸

¹⁶ MacKinlay, Richardson (1991).

¹⁷ Hansen (1982).

$$J_1 = T \hat{\boldsymbol{\alpha}}' \left[\boldsymbol{\eta} \underbrace{[\mathbf{D}'_T \mathbf{S}_T^{-1} \mathbf{D}_T]^{-1}}_{\mathbf{\Omega}_T} \boldsymbol{\eta}' \right] \hat{\boldsymbol{\alpha}}$$

$\begin{matrix} 1 \times N & N \times 2N & 2N \times 2N & 2N \times 2N & 2N \times 2N & 2N \times N & N \times 1 \\ & & \mathbf{\Omega}_T & & & & \end{matrix}$

con:

$\boldsymbol{\eta} = \mathbf{I}_N \otimes [1, 0]$: matrice composta solamente da numeri 1 e 0;

$\mathbf{\Omega}_T = [\mathbf{D}'_T \mathbf{S}_T^{-1} \mathbf{D}_T]^{-1}$: matrice di varianza-covarianza dei parametri della regressione, calcolata tramite il prodotto di:

$\mathbf{D}_T = -\frac{1}{T} \sum_{t=1}^T [\mathbf{I}_N \otimes \mathbf{Z}_t \mathbf{Z}_t']$: matrice simmetrica composta da sottomatrici quadrate identiche,

allineate lungo diagonale principale, contenenti le statistiche descrittive del portafoglio P ;

$\mathbf{S}_T = \frac{1}{T} \sum_{t=1}^T [\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \otimes \mathbf{Z}_t \mathbf{Z}_t']$: matrice di densità spettrale;

$\mathbf{Z} = [\mathbf{1}, \mathbf{r}_p']$: matrice $2 \times T$ avente nella prima riga solo numeri 1 e nella seconda riga la serie storica degli extrarendimenti del portafoglio P .

In altri termini, si effettua un *test* simile a quello di Wald, ma nel quale le varianze-covarianze dei residui dipendono anche dai rendimenti dei componenti dell'indice P , come evidenziato in particolare dalla costruzione della matrice di densità spettrale \mathbf{S}_T . Formalmente, si può quindi affermare che il *test* GMM opera in una ipotesi di eteroschedasticità dei residui.

La statistica J_1 si distribuisce asintoticamente come una chi-quadro con N gradi di libertà: dovendo quindi utilizzare un campione di durata potenzialmente illimitata è necessario ricorrere a tecniche di ricampionamento. Tale soluzione può però portare a risultati subottimali in una situazione di autocorrelazione dei residui se il ricampionamento è effettuato tramite le comuni metodologie. Esse infatti prevedono che si definisca una funzione di densità dalla quale poi effettuare estrazioni casuali.

A sfavore di tale procedura si possono però evidenziare alcune controindicazioni, *in primis* la perdita di informazione riguardo ai rendimenti. Tali rendimenti, infatti, sono correlati sia tra di loro (correlazione *cross-section*) sia con quelli che li precedono e seguono (autocorrelazione o correlazione *time series*): se fossero ricampionati per ciascun periodo temporale t , potrebbero conservare solo la prima delle due forme di dipendenza. Il campionamento *cross-section* è sempre possibile per una distribuzione multivariata N -dimensionale, mentre la scelta di rendimenti isolati temporalmente cancella completamente qualsiasi autocorrelazione *time series* presente nel campione di dati originari. Al fine di superare queste limitazioni, è possibile far ricorso ad una tecnica euristica quale il *block bootstrap*.

Il *block bootstrap* consiste nell'estrazione congiunta di blocchi di residui dei rendimenti, ciascuno di durata b predefinita.¹⁹ E' proprio la durata di tali blocchi a permettere di rappresentare il fenomeno dell'autocorrelazione, ma solamente all'interno di ciascuno di essi: si noti che nel momento in cui il ricampionamento passa al blocco successivo, questo può essere estratto dalla serie storica in un qualsiasi punto non consecutivo al precedente ed essere quindi privo di correlazione con questo. Di conseguenza la scelta di b risulta essere sottoposta a due contrastanti problematiche:

- al diminuire di b si associa una minore considerazione per l'autocorrelazione;
- al crescere di b corrisponde un minor numero di permutazioni possibili sulla base della serie storica disponibile, necessariamente limitata.

¹⁸ Chou, Zhou (2006).

¹⁹ Shao, Tu (1995).

Al fine di applicare il *block bootstrap* al test GMM di efficienza dell'indice P è necessario seguire la seguente procedura:²⁰

1) effettuare la regressione degli extrarendimenti degli N titoli sull'extrarendimento dell'indice P ;

2) calcolare i residui per ciascuno degli N titoli in ogni periodo t sulla base della seguente regressione *unrestricted*: $\hat{\varepsilon}_{it} = r_{it} - \hat{\beta}_i r_{Pt} - \hat{\alpha}_i$;²¹

3) stimare i coefficienti di regressione β_i nel caso *restricted*: $r_i = \tilde{\beta}_i r_P$;²²

4) estrarre un numero l di blocchi *cross-section* di residui, ciascuno di durata b , dalla matrice $\hat{\varepsilon}$ (dimensioni $N \times T$), con $T = l \times b$;²³

5) simulare M serie storiche dell'extrarendimento di P tramite il *block bootstrap*, ottenendo la matrice \mathbf{r}_P^* di dimensioni $M \times T$;²⁴

6) simulare N serie di rendimenti, ciascuna di durata T , secondo la seguente equazione esemplificativa relativa al solo titolo i al tempo t : $r_{it}^* = \tilde{\beta}_i r_{Pt}^* + \varepsilon_{ik}^*$, con $1 < k < T$;

7) ripetere il punto 6 un numero M molto elevato di volte (circa 10.000).

In questo modo si ottiene una matrice tridimensionale di extrarendimenti simulati, avente dimensioni $N \times T \times M$.

Con quest'ultimo passaggio si conclude l'implementazione del *block bootstrap* e quindi i dati così simulati possono essere utilizzati per calcolare M valori della statistica *test* J_1^* . Tali valori sono utilizzati per ricostruire la distribuzione empirica di J_1 tramite la seguente procedura:

1) regredire in ognuno degli M scenari la serie storica simulata degli N componenti sugli extrarendimenti simulati dell'indice P , al fine di stimare il vettore $N \times 1$ delle intercette $\hat{\alpha}_i^*$ e la matrice $N \times T$ dei residui $\hat{\varepsilon}^*$;

2) calcolare la statistica J_1^* sulla base dei dati stimati al punto 1 per ciascuno scenario;

3) confrontare il valore delle M statistiche J_1^* con la statistica J_1 calcolata con i dati storici e stimarne quindi il *p-value*;

4) nel caso in cui tale *p-value* sia superiore ad uno meno l'intervallo di confidenza prescelto, l'indice P è efficiente.

2.4. La misura di efficienza relativa di Kandel e Stambaugh

I *test* sin qui esposti operano tutti all'interno della cornice teorica del CAPM, seppure con alcune distinzioni rispetto a questo modello. Il metodo di analisi dell'efficienza di un indice proposto da Kandel e Stambaugh trascende invece il CAPM, evitando di ricorrere alle sue stringenti ipotesi e, in genere, a concetti quali la dipendenza lineare dei rendimenti dei titoli rispetto al portafoglio di mercato.²⁵ Al contrario, esso riparte dai fondamenti stessi della MPT.

Kandel e Stambaugh privilegiano la valutazione dell'efficienza relativa, misurata non rispetto ad un modello di *pricing* degli strumenti finanziari, bensì rispetto alla frontiera efficiente. Si tratta

²⁰ Chou, Zhou (2006).

²¹ La regressione è definita *unrestricted* (non vincolata) in quanto effettuata senza porre alcuna limitazione al valore dell'intercetta α_i .

²² In questo caso la regressione, *restricted* (vincolata), segue la teoria del CAPM, la quale postula che in equilibrio i rendimenti dei titoli dipendano solamente dalla loro sensibilità al portafoglio di mercato e quindi non mostrino intercette α_i diverse da zero.

²³ Si evidenzia che l'estrazione è compiuta con sostituzione (*replacement*): un blocco di dati, o una sua parte, può essere campionato nuovamente in seguito.

²⁴ L'asterisco a fianco del rendimento è qui utilizzato convenzionalmente per indicare che si tratta di un valore ottenuto in seguito alla simulazione *block bootstrap*.

²⁵ Kandel, Stambaugh (1995).

quindi di effettuare un confronto tra l'extrarendimento di P rispetto al portafoglio a varianza minima g e l'extrarendimento di x , portafoglio efficiente avente la medesima volatilità di P , sempre rispetto a g .

In termini formali, si definisce ψ_P , ovvero la misura di efficienza relativa del portafoglio P , come:

$$\psi_P = \frac{\mu_P - \mu_g}{\mu_x - \mu_g}$$

Il range di ψ_P è compreso tra $-\infty$ e $+1$.²⁶ A una perfetta efficienza è associato il valore pari a $+1$, che si verifica nel caso in cui si assista ad una coincidenza tra P e x : il portafoglio P giace quindi sulla frontiera. Nel caso in cui, invece, ψ_P assuma il valore pari a -1 essa indica che P è posto sulla parte inferiore della frontiera efficiente, ovvero sul segmento di tale curva che rappresenta i portafogli a rendimento minimo dato un certo livello di volatilità. L'asimmetria della distribuzione di ψ_P è dovuta al fatto che la frontiera efficiente è costruita considerando anche la presenza del portafoglio P : nel caso in cui questo fosse perfettamente efficiente potrebbe al massimo essere collocato su tale curva, limitando quindi a $+1$ i valori che possono essere assunti da ψ_P . Al contrario, non esiste alcun vincolo teorico che imponga un livello massimo di inefficienza raggiunto da parte di un indice.

Se invece si ragiona non in termini di rendimenti medi, bensì di varianze, la misura ψ_P deve essere valutata al quadrato:²⁷

$$\psi_P^2 = \frac{\sigma_y^2 - \sigma_g^2}{\sigma_P^2 - \sigma_g^2}$$

Questo cambio di variabili può essere giustificabile nel caso in cui P presenti una volatilità superiore a quella del portafoglio efficiente con rendimento massimo²⁸ e, al contempo, un rendimento tale da non posizionarlo sulla frontiera. In questa situazione, infatti, la notazione originaria dell'indice ψ_P non potrebbe essere utilizzata, in quanto nessuna combinazione lineare efficiente dei titoli disponibili avrebbe la medesima deviazione *standard* di P e quindi sarebbe impossibile calcolare μ_x .

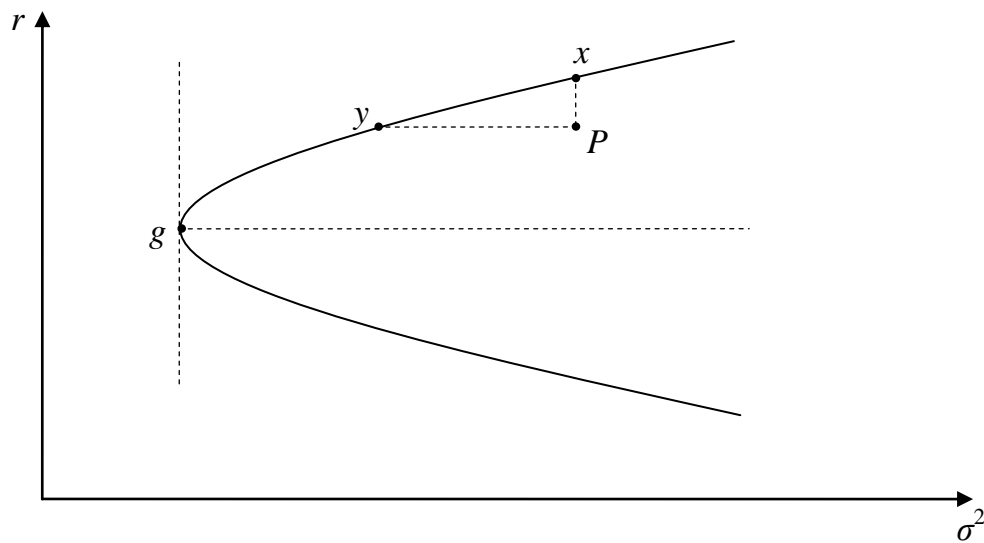
Nell'impiego del quadrato di ψ_P il confronto avviene con il portafoglio y , ovvero il punto sulla frontiera efficiente avente il medesimo rendimento di P e, salvo nel caso in cui quest'ultimo sia perfettamente efficiente, contraddistinto da una volatilità inferiore. Questa notazione è però altamente discutibile, in quanto assegna il medesimo livello di efficienza relativa ai portafogli posti simmetricamente rispetto alla retta parallela all'asse delle ascisse passante per il portafoglio a varianza minima g . Quindi, in presenza di varianze identiche ma extrarendimenti differenti, due portafogli sarebbero classificati sullo stesso livello di efficienza: un risultato chiaramente inaccettabile.

²⁶ Sarebbe compreso tra -1 e $+1$ nel caso in cui la frontiera fosse costruita in assenza del vincolo di non negatività dei pesi dei componenti.

²⁷ Tale formulazione discende direttamente dall'equazione della parabola. Si rammenta infatti che la frontiera efficiente *unconstrained* assume la conformazione di un'iperbole nel piano media-deviazione *standard* e di una parabola, con asse di simmetria parallelo all'asse delle ascisse σ^2 , nel caso in cui il rischio sia misurato tramite la varianza (figura 2).

²⁸ Si ricorda che, in assenza di vendite allo scoperto, la frontiera efficiente è limitata tra il portafoglio a varianza minima g e il portafoglio composto al 100% dal titolo i -esimo avente rendimento massimo.

Figura 2: Misura di efficienza di Kandel e Stambaugh



Fonte: Kandel, Stambaugh (1995).

Questo approccio, per quanto semplice, offre alcuni vantaggi rispetto a quelli esposti in precedenza e fondati sui presupposti teorici del CAPM: tramite la misura di Kandel e Stambaugh è sempre possibile costruire classifiche di efficienza relativa degli indici, anche se la validità empirica del modello di Sharpe, Lintner e Mossin non fosse verificata.

La finalità delle classifiche stilte tramite la misura di Kandel e Stambaugh, quindi, non è quella di determinare quale sia il portafoglio di mercato, da utilizzare poi, ad esempio, per l'analisi del valore atteso in equilibrio dei titoli oppure per la stima di indicatori di *performance* relativa quale l'*appraisal ratio*. Al contrario, questa metodologia d'analisi risulta particolarmente utile nel campo dell'*asset management*, in quanto consente di stabilire, a titolo d'esempio, quale sia l'indice più efficiente all'interno di un dato mercato, allo scopo di impiegarlo come *benchmark* per investimenti passivi.

Nell'impiego della misura di Kandel e Stambaugh si è ritenuto opportuno ricorrere alla tecnica euristica del ricampionamento (*resampling*) per la costruzione della frontiera efficiente, al fine di limitare i ben noti problemi di *error maximization* tipici del modello di Markowitz.²⁹ In particolare, l'impiego della tecnica del *resampling* consente di ottenere una stima dell'indicatore ψ_P nettamente più robusta rispetto a quella calcolata con il metodo proposto dei due autori tre anni prima che Michaud esponesse il proprio procedimento.

Ciò nonostante l'applicazione del ricampionamento al modello di Kandel e Stambaugh non è del tutto esente da difetti. La frontiera ricampionata è costituita infatti da un insieme discreto di portafogli, non da una serie continua. Nella formula della misura di efficienza ψ_P si ipotizza che sia possibile costruire un qualsiasi portafoglio efficiente, vincolato a un dato livello di deviazione *standard*. Questo è perfettamente possibile allorquando si effettua una singola ottimizzazione media-varianza, ma in presenza di una simulazione Monte Carlo può accadere che in almeno uno degli M scenari nessun portafoglio efficiente abbia una volatilità pari a quella dell'indice oggetto di studio P .

Per superare questo problema di ordine eminentemente pratico si è quindi deciso di procedere con la costruzione dell'intera frontiera sulla base dei pesi medi dei portafogli efficienti *resampled* per ciascuno degli M scenari simulati. In seguito, la stima dell'extrarendimento del portafoglio x ,

²⁹ La procedura del ricampionamento è stata esposta per da Richard O. Michaud nella prima edizione, del 1998, di Michaud, Michaud (2008).

ovvero di quel portafoglio efficiente avente la stessa deviazione *standard* di P , è stata effettuata per interpolazione. Il metodo prescelto è il cosiddetto *cubic spline*, che permette di superare sia il problema di linearità a tratti (con conseguente concavità nulla) della comune interpolazione lineare, in disaccordo con la conformazione concava della frontiera efficiente, sia il fenomeno di Runge³⁰ tipico delle interpolazioni polinomiali di grado elevato.

3. L'analisi empirica

3.1. La composizione del campione

Lo scopo della presente analisi consiste nella valutazione dell'efficienza dei principali indici del mercato azionario utilizzati nell'ambito dell'*asset management*, al fine di individuare i migliori *benchmark* in termini di efficienza. Nello specifico, è stata presa in considerazione l'*asset class* azionaria USA, data l'ampia diversificazione tra le tecniche costruttive degli indici che ne sono maggiormente rappresentativi (tabella 1).

La selezione è stata effettuata sulla base di criteri quali-quantitativi. Gli indici prescelti sono infatti tra i più utilizzati come *benchmark* nella prassi operativa, inoltre devono rispettare le seguenti caratteristiche:

- avere un *track record* di rendimenti mensili, disponibili sul *database* Morningstar Direct, risalente almeno a gennaio 2003;
- essere composti, completamente o in parte, da titoli azionari appartenenti ai portafogli *beta-sorted*, descritti in dettaglio nel paragrafo 3.2.

Il tasso *risk free* è stato considerato pari all'indice Citigroup Treasury Bill 1 Month USD, calcolato come media mensile dei rendimenti dei T-Bill con scadenza all'emissione pari a quattro settimane.

Come evidenziato nel paragrafo riguardante il *test* GRS, le analisi statistiche qui esposte sono influenzate dalla dimensione del campione utilizzato. Inoltre, e indipendentemente dal *test* impiegato, deve essere sottolineato che la matrice di varianza e covarianza dei residui $\hat{\Sigma}$, nel caso in cui il numero di *asset* N superi il numero di intervalli temporali disponibili T , sarebbe singolare e quindi non invertibile: ciò renderebbe inutilizzabile il *test* di Wald e i suoi derivati. Con riguardo invece al *test* di Kandel e Stambaugh, l'impatto dimensionale è meno importante da un punto di vista strettamente statistico, ma assume rilevanza in un'ottica di *asset management*. Al crescere di N , infatti, diviene più probabile includere nel campione titoli con *performance* estreme. Da ciò deriverebbe una maggiore probabilità di tracciare una frontiera efficiente composta da portafogli costituiti da titoli caratterizzati da profili di rischio e rendimento difficilmente replicabili *out of sample*.

³⁰ Tale fenomeno consiste nell'oscillazione della linea interpolante tra i punti lontani dal centro della serie e nell'incremento dell'errore di interpolazione al crescere del grado del polinomio. In altri termini, polinomi di grado elevato (o, comunque, superiore al secondo) mostrano derivate seconde di segno alterno tra un punto e l'altro della serie da interpolare e derivate prime via via di valore più elevato al crescere del loro grado.

Tabella 1: Indici azionari statunitensi

Indice	Ponderazione	Proventi reinvestiti
DJ Composite Average PR	Prezzo	No
DJ Composite Average TR	Prezzo	Lordi
DJ Industrial Average PR	Prezzo	No
DJ Industrial Average TR	Prezzo	Lordi
MSCI USA GR	Flottante	Lordi
MSCI USA NR	Flottante	Netti
MSCI USA PR	Flottante	No
Russell 1000 CEW TR	Equiponderato	Lordi
Russell 1000 SEW TR	Equiponderato per settori	Lordi
Russell 1000 PR	Flottante	No
Russell 1000 TR	Flottante	Lordi
Russell 3000 CEW TR	Equiponderato	Lordi
Russell 3000 PR	Flottante	No
Russell 3000 TR	Flottante	Lordi
S&P 500 Equal Weighted TR	Equiponderato	Lordi
S&P 500 NR	Flottante	Netti
S&P 500 PR	Flottante	No
S&P 500 TR	Flottante	Lordi
Value Line New Arithmetic PR	Equiponderato	No
Wilshire 5000 Equal Weight PR	Equiponderato	No
Wilshire 5000 Total Market PR	Flottante	No
Wilshire 5000 Total Market TR	Flottante	Lordi

Fonte: manuali metodologici degli indici.

3.2. La costruzione dei portafogli *beta-sorted*

Date queste premesse, si è posto il problema di ridurre il rapporto N/T agendo sul numeratore. A tal fine, la letteratura scientifica propone un metodo di aggregazione degli *asset* ampiamente utilizzato: il ricorso ai portafogli *beta-sorted*.³¹

Il primo passo per la loro costruzione consiste nella stima del vettore $\hat{\beta}_i$ dei coefficienti angolari delle rette della regressione OLS dei rendimenti degli N titoli sul rendimento del portafoglio di cui essi sono componenti. In seguito, questi titoli sono ordinati dal minore al maggiore a seconda del livello del loro beta e suddivisi in un numero arbitrario di quantili (indicato con Q). Nel caso specifico, si è deciso di proseguire con il metodo esposto da Gibbons, Ross e Shanken e quindi di utilizzare dieci quantili.³²

Ogni q -esimo portafoglio *beta-sorted* è composto in parti uguali dai titoli del relativo quantile: si tratta quindi di un portafoglio equiponderato, il cui rendimento medio è pari alla media dei rendimenti. Da questo deriva che $\hat{\beta}_q$ è pari al rischio sistematico medio delle azioni di q e che $\hat{\alpha}_q$ è l'intercetta media.

Un rischio cui è possibile incorrere con questa aggregazione è quello di introdurre una distorsione nella selezione dei componenti di ogni portafoglio *beta-sorted*. Ciò è dovuto alla

³¹ Black et al. (1972).

³² Si noti che dieci è anche dimensione massima ritenuta idonea per i *test* qui utilizzati (Campbell et al., 1997).

presenza di errori di stima, tali per cui i beta di valore più elevato contengono al loro interno una componente spuria positiva, mentre l'opposto si verifica per i beta di valore minore. A causa di tale *bias* i titoli (e quindi i portafogli *beta-sorted*) con una forte esposizione al rischio sistematico mostrerebbero intercette inferiori al reale, mentre avverrebbe l'opposto in presenza di beta di valore limitato. Questa distorsione può essere evitata separando temporalmente la costruzione dei portafogli e il loro utilizzo all'interno dei *test*, rendendo indipendenti questi ultimi dalla precedente selezione dei titoli in base al loro beta.

Per effettuare l'analisi di efficienza è stato necessario selezionare le serie storiche dei componenti di un indice rappresentativo del mercato statunitense sul *database* Datastream. Tale indice è stato individuato nello Standard & Poor's 500 TR, formato dalle prime cinquecento società, per capitalizzazione, quotate presso il New York Stock Exchange e calcolato considerando il reinvestimento dei proventi lordi (dividendi, assegnazione di azioni a titolo gratuito, ecc.).

Dato che l'indice è soggetto ad una revisione trimestrale, con riguardo sia ai componenti sia ai loro pesi, sarebbe stato errato utilizzare i soli titoli azionari presenti nell'ultima data disponibile. Al contrario, è stata seguita la seguente procedura:

- è stato scaricato da Datastream l'elenco dei componenti per ogni trimestre a partire da dicembre 2002 fino a settembre 2010;
- di ogni elenco sono stati scaricati 63 mesi di rendimenti *total return*, calcolati da Datastream, di cui i primi 60 mesi (*in sample*) utilizzati per la stima dei beta e 3 (*out of sample*) per la costruzione dei portafogli *beta-sorted*;
- per ogni finestra *rolling* di 63 mesi, sono stati regrediti sull'indice i primi 60 rendimenti di quei componenti aventi almeno 24 mesi *in sample* e almeno due mesi *out of sample*;
- i titoli sono stati riordinati in base al beta e aggregati in dieci portafogli *beta-sorted*;
- il rendimento mensile R_q di ciascun portafoglio *beta-sorted* è la media aritmetica semplice dei rendimenti dei suoi componenti nei mesi *out of sample*, coprendo il periodo gennaio 2003-dicembre 2010.

Da questo procedimento deriva che la composizione di ciascun portafoglio è variabile nel tempo, modificandosi ogni tre mesi: di fatto, i *beta-sorted portfolio* possono essere considerati alla stregua di fondi comuni, dotati di una loro identità autonoma ma la cui composizione interna è sempre variabile.

Relativamente alla non-stazionarietà dei beta, la ricalibrazione dei portafogli comporta un importante vantaggio: essa infatti consente di riallocare i titoli in portafogli diversi a seconda della variazione della loro esposizione al fattore di rischio sistematico, seppure non istantaneamente bensì con un *lag* temporale massimo di tre mesi.³³ Quindi, a fronte di una variabilità nella composizione dei portafogli *beta-sorted*, si mantiene sostanzialmente costante il profilo di rischio di ciascuno di loro, in quanto sono i titoli ad essere trasferiti nel caso in cui il loro beta subisse delle variazioni e non è invece il portafoglio nel suo complesso a subire tale mutamento.

3.3. La metodologia applicativa dell'analisi empirica

Le serie *out of sample* dei portafogli sono state utilizzate come *panel* di componenti di tutti gli indici della *asset class* USA, indipendentemente dal fatto che tali *beta-sorted portfolio* siano o meno costituiti dai medesimi titoli inclusi in ciascuno degli indici analizzati. Questa scelta, oltre ad essere dettata dalla mancanza di serie storiche dei *component* di ogni indice, è fondata anche su di una *ratio* teorica: al fine di individuare il miglior *benchmark* rappresentativo di un mercato, è utile

³³ A titolo di paragone, si rammenta che Black et al. (1972) ricalibrano i loro portafogli *beta-sorted* ogni 12 mesi, quindi con minore precisione. Va comunque notato che due periodi *in sample* consecutivi si sovrappongono per 57 mesi, quindi la variabilità dei beta è alquanto ridotta, essendo dovuta solamente a *shock* intervenuti nei tre rendimenti mensili non comuni ad entrambe le finestre temporali.

che tutti gli indici azionari di quello specifico mercato siano confrontati con lo stesso campione di titoli.

Data l'ipotesi di normalità degli extrarendimenti insita nei *test* effettuati, con l'eccezione del GMM, prima di tutto è stato verificato l'eventuale scostamento delle serie storiche qui utilizzate rispetto alla distribuzione gaussiana tramite l'utilizzo del noto *test* di Jarque-Bera (tabella 2). Questa misura statistica, che si distribuisce asintoticamente come una chi-quadro con due gradi di libertà, soffre però di serie distorsioni nel caso in cui il campione sia di dimensioni limitate. Per questa ragione l'analisi è stata effettuata utilizzando il programma Matlab, che stima il *p-value* del *test* non tramite la distribuzione asintotica, bensì ricorrendo ad una tabella di valori critici computata tramite simulazioni Monte Carlo, tale da correggere il *bias* che si manifesta in particolare per le serie storiche di durata inferiore ai 2.000 periodi, come in questo caso (96 periodi).

Tabella 2 Statistiche descrittive degli extrarendimenti degli indici

Indice	Media	Deviazione Standard	Asimmetria	Curtosi	JB stat	p-value (1)	Normalità
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
Wilshire 5000 Total Market TR	0,56%	4,53%	-0,86	5,04	28,48	0,14%	No

(1) Arrotondato a 0,10% da Matlab se tendente a zero.

Fonte: elaborazione dell'autore su rendimenti mensili, gennaio 2003 - dicembre 2010.

Alla luce di questa misurazione, si ritiene che risulterà particolarmente significativo il *test* GMM, in quanto estraneo all'ipotesi di normalità.

Il *test* GRS è stato utilizzato nella notazione originaria, anziché nella sua scomposizione in *Sharpe ratio* e *appraisal ratio*, in quanto quest'ultima risulta troppo distorta in presenza di correlazioni tra i residui.

La seconda delle prove condotte è stata il *test* di Wald, realizzato attraverso la tecnica del *bootstrap* al fine di costruire una distribuzione empirica che consenta di superare le problematiche

relative a campioni di dimensione finita. Sono stati simulati 10.000 scenari per ogni gruppo di dieci portafogli *beta-sorted* per ciascun indice sottoposto a valutazione tramite il ricorso ad una procedura identica a quella esposta nel paragrafo 2.3, eccetto che per la durata b dei blocchi, che in questo caso è posta pari ad uno.³⁴ In ognuno degli scenari è stato calcolato il valore assunto dalla statistica WT^* e confrontato con WT , ottenuto con le serie storiche originarie. Il p -value è risultato quindi pari al numero di WT^* superiori a WT , diviso per 10.000.

A differenza degli altri *test*, quello proposto da Kandel e Stambaugh non richiede che il campione si mantenga entro determinati vincoli di numerosità, sia con riguardo ai componenti, sia alle durate. Ciò nonostante si è deciso di impiegare nella costruzione della frontiera efficiente *resampled* non i singoli componenti degli indici, bensì, anche in questo caso, i dieci portafogli *beta-sorted*. La scelta è stata dettata sia da ragioni di coerenza con gli altri indicatori di efficienza qui utilizzati sia per motivi statistici. Il raggruppamento in portafogli, infatti, consente di limitare l'impatto di rendimenti *outlier*, riducendo ulteriormente l'impatto dell'*error maximization*, già parzialmente compensato grazie all'uso del *resampling*. In dettaglio, quest'ultima procedura è stata applicata compiendo i seguenti *step*:

- in 10.000 scenari, ciascuno composto da 96 periodi, sono stati simulati congiuntamente gli extrarendimenti mensili dei dieci portafogli *beta-sorted* e di tutti gli indici appartenenti al medesimo mercato azionario;
- in ogni scenario e per ciascun indice p -esimo sono state calcolate le frontiere efficienti composte dai dieci portafogli *beta-sorted* e da un singolo indice alla volta;
- per ogni indice è stata calcolata la frontiera *resampled* e, tramite un'interpolazione *cubic spline*, è stato individuato il corrispondente portafoglio efficiente x ;
- infine per ogni indice è stato calcolato il valore di ψ_p .

Nel *test* GMM si è fatto ricorso alla distribuzione empirica stimata tramite *block bootstrap* di 10.000 scenari, utilizzando blocchi di lunghezza unitaria pari a sei periodi. Questa durata è stata definita sulla base dell'autocorrelazione degli extrarendimenti degli indici azionari selezionati: nella grande maggioranza dei casi (19 su 22 indici), infatti, l'autocorrelazione è significativa fino al quarto *lag* temporale. L'uso di blocchi di sei periodi è un compromesso che consente di catturare, all'interno di ciascun blocco, l'autocorrelazione:

- del primo ordine in cinque periodi su sei;
- del quarto ordine in due periodi su sei.³⁵

3.4. I risultati dell'analisi empirica

La valutazione dell'efficienza degli indici del mercato azionario tramite *test* molto diversi, sia come impostazione teorica sia come costruzione dei campioni (serie storiche contro serie ricampionate), ha portato a risultati sostanzialmente concordi e utili sia per finalità di ordine pratico, quale l'individuazione dei *benchmark* ottimali, sia di tipo teorico, distinguendo nettamente tra loro le diverse tecniche di costruzione degli indici.

La tabella 3 riporta i valori assunti dalle statistiche dei diversi *test* e, ove disponibile, il livello percentuale del p -value, ossia della probabilità che non si possa rigettare l'ipotesi di efficienza di ciascun indice.

³⁴ Si tratta, in altri termini, di un *bootstrap* tradizionale, non di un *block bootstrap*. Il *test* di Wald, infatti, presume che i residui siano indipendentemente e identicamente distribuiti: la presenza di correlazione seriale e quindi la necessità di ricorrere a blocchi di durata maggiore di uno sono escluse *a priori* dal modello.

³⁵ In Chou, Zhou (2006) i periodi utilizzati sono pari a tre, ma la logica sottostante tale scelta non è argomentata dagli autori.

Tabella 3: Livello di efficienza degli indici

Indice	Kandel e	GRS		Wald		GMM	
	Stambaugh	GRS	p-value	WT	p-value (1)	J1	p-value (1)
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%
Wilshire 5000 Total Market TR	0,6333	0,7949	95,43%	8,9778	95,04%	8,3249	86,23%

(1) P-value stimato tramite la simulazione di 10.000 scenari.

Fonte: elaborazione dell'autore su rendimenti mensili, gennaio 2003 - dicembre 2010.

Si osserva, innanzitutto, che nella finestra temporale considerata, caratterizzata da una crescita generalizzata dei corsi azionari seguita da un periodo di forte turbolenza, tutti gli indici risultano essere efficienti. Inoltre, la grande maggioranza degli indici mostra livelli di *p-value* superiori al 90% in almeno uno degli indicatori. Accanto a questo giudizio complessivo risulta però maggiormente utile analizzare i *ranking* conseguiti utilizzando i risultati dei *test*.

Poiché i *test* utilizzati sono stati quattro, non è possibile avere la certezza *a priori* di esser sempre in grado di pervenire ad un giudizio univoco. Al fine, quindi, di costruire una classifica unitaria e non più suddivisa in *test* distinti è possibile seguire un approccio di analisi multicriterio, tipico della teoria delle decisioni, quale il PROMETHEE.³⁶ In tale tecnica si ricorre al flusso netto di dominanza $\phi(i) = \phi^+(i) - \phi^-(i)$, ove con $\phi^+(i)$ e $\phi^-(i)$ si indicano rispettivamente la somma dei gradi di preferenza π^{37} di *i* rispetto a ciascuno degli altri indici e la somma dei gradi di preferenza di ognuno degli altri indici rispetto a *i*, definiti come:

$$\phi^+(i) = \frac{1}{I-1} \sum_{\forall j \neq i} \pi(i, j)$$

$$\phi^-(i) = \frac{1}{I-1} \sum_{\forall j \neq i} \pi(j, i)$$

con *I*: numero complessivo di indici oggetto di valutazione.

³⁶ Preference Ranking Organization Method for Enrichment of Evaluations (Brans, 1982).

³⁷ Il grado di preferenza multicriterio per ciascuna coppia di indici *i, j* è pari alla media semplice delle funzioni di preferenza per ognuna delle quattro misure di efficienza, che assumono valore 1 se il *ranking* di *i* è maggiore di quello di *j* secondo una data misura di efficienza, e valore 0 in caso contrario.

Tabella 4: Flussi di dominanza delle misure di efficienza degli indici statunitensi

Indice	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	$\phi^+(i)$	$\phi(i)$
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	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
Flusso negativo di dominanza $\phi^-(i)$	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		

Fonte: elaborazione dell'autore.

Cinque dei sei indici di borsa costruiti tramite la tecnica dell'equiponderazione si collocano nelle prime cinque posizioni nel *ranking* di efficienza.³⁹

All'interno di questo sottocampione è possibile osservare, tra gli indici Russell, la presenza di due metodi di ponderazione distinti:

- *constituent equal weight* (CEW);
- *sector equal weight* (SEW).

Il primo segue il comune concetto di equiponderazione, che prevede un peso identico per ciascun componente a ogni data di ribilanciamento. Ad esempio, mentre nel Russell 1000 (indice *float weighted*) le prime 100 società incidono per il 58% della capitalizzazione complessiva, nella versione equiponderata, per definizione, tale percentuale si riduce al 10%.

Al contrario, la tecnica dell'equiponderazione settoriale è stata sviluppata internamente dalla società Russell e attualmente non è impiegata da altri *index provider*. Essa è stata implementata allo scopo di evitare una rappresentazione distorta dei settori economici (*sector bias*). Di conseguenza, il peso di ciascun titolo viene stabilito tramite due passaggi:

- il mercato azionario è suddiviso in nove settori economici, tutti inclusi nell'indice con lo stesso peso;
- all'interno di ciascun settore le singole azioni sono equiponderate.

Tramite questo procedimento Russell supera la problematica legata all'eccessiva concentrazione in quei settori, quali il *Financial services* e il *Consumer discretionary*, che sono rappresentati sui mercati azionari statunitensi da parte di un numero di società particolarmente elevato.

Il Russell 1000 SEW TR, quarto nel *ranking*, nonostante sia stato costruito seguendo questa tecnica innovativa, appare meno efficiente del più semplice Russell 1000 CEW TR (primo in classifica), ma deve essere sottolineato che questo risultato è influenzato dal giudizio derivante dal *test GRS* che, in presenza di deviazioni dalla gaussiana, non è del tutto affidabile.

Il Russell 3000 CEW TR, nonostante l'elevata diversificazione, è inferiore, in termini di efficienza, rispetto al Russell 1000 CEW TR: quest'ultimo è infatti il secondo indice nella classifica. Questa valutazione è significativa, in quanto evidenzia come il potenziale impatto dell'anomalia nota come *size effect* non sia stato rilevante in questo caso.

Lo Standard & Poor's 500 Equal Weight TR risulta essere il terzo indice più efficiente. Il valore di questo risultato è particolarmente elevato: sono infatti i titoli appartenenti all'indice S&P 500 i componenti che sono stati utilizzati per la costruzione dei portafogli *beta-sorted*. Ne consegue che le analisi condotte sugli indici della famiglia S&P 500 non presentano alcun potenziale *bias* derivante dall'utilizzo di componenti non perfettamente coincidenti con quelli dell'indice o dalla presenza di difformi procedure di reinvestimento dei proventi, così come da diverse aliquote impositive applicate sugli stessi. La graduatoria invece dipende solamente dalle diverse metodologie di calcolo, essendo identico ogni altro aspetto.

Il quinto indice per efficienza è il Value Line New Arithmetic PR. Si tratta dell'indice equiponderato dei circa 1.700 titoli che sono oggetto di analisi nel *Value Line Investment Survey*, un *report* periodico, pubblicato da oltre ottant'anni, dedicato all'analisi fondamentale del mercato azionario statunitense.⁴⁰ In questo caso l'elevata efficienza è raggiunta nonostante si tratti di un

³⁹ Il solo Wilshire 5000 Equal Weight PR è superato da indici *value weighted*, nei quali il peso attribuito a ciascun componente è proporzionale alla sua capitalizzazione (*capitalization* o *cap weight*) o al suo flottante (*float weight*).

⁴⁰ Sin dal 30 giugno 1961 la società Value Line calcola il valore anche del Value Line Geometric Index. Nonostante sia, storicamente, il primo indice equiponderato oggetto di ampia diffusione, esso è stato ormai completamente soppiantato dalla sua versione Arithmetic, introdotta nel 1988, a causa delle diverse modalità di calcolo e delle finalità ad esse sottese. Il rendimento giornaliero del primo indice, infatti, è una media geometrica dei rendimenti dei suoi componenti, non aritmetica come invece avviene nell'altro indice equiponderato. La scelta di questa statistica, solitamente impiegata in calcoli *time series* e non, come in questo caso, *cross section*, si spiega con la decisione di concepire tale indice come un indicatore rappresentativo del rendimento mediano dei suoi componenti, dato che media

indice di prezzo. D'altro canto essa è giustificabile dall'ampia diversificazione e dalla potenziale presenza di *selection bias* derivante dal metodo di selezione dei titoli da includere, fatto quest'ultimo che priva l'indice delle caratteristiche di oggettività tipiche di un buon *benchmark*. Inoltre, le variazioni di questo paniere sono calcolate come la media aritmetica dei rendimenti dei suoi componenti: questo metodo di calcolo implica un ribilanciamento quotidiano, rendendo il New Arithmetic un indice estremamente oneroso da replicare. Al contrario, gli altri esempi di equiponderazione presenti nel campione sono soggetti a ribilanciamento trimestrale. Ciò comporta un maggiore disallineamento dei coefficienti di ponderazione tra una data e l'altra di revisione, ma al contempo un netto contenimento dei costi di transazione in capo a un investitore indicizzato.

Il prosieguo della classifica offre alcuni spunti di riflessione riguardo al trattamento dei proventi. All'interno di ogni "famiglia" di indici dello stesso *provider* si ripete la medesima classifica di efficienza, posizionando al primo posto i panieri *gross total return*. La riduzione dell'ammontare reinvestibile derivante dall'applicazione delle aliquote delle imposte sui redditi di capitale comporta un grado di efficienza inferiore in capo agli indici *net total return*, mentre gli indici di prezzo si posizionano per ultimi.

Ad esempio, i noti indici *float weight* S&P 500 TR, NR e PR (rispettivamente *gross total return*, *net total return* e *price*) mostrano un livello di efficienza nettamente inferiore rispetto alla versione equiponderata, collocandosi rispettivamente al 12°, 17° e al 21° posto.

Come nel caso dell'S&P 500, anche gli indici Wilshire 5000 sono caratterizzati da una maggiore efficienza, a parità di reinvestimento dei proventi, in capo alla versione equiponderata. Questa, nonostante sia rappresentata da un indice di prezzo, è in grado di collocarsi in prossimità del Wilshire 5000 Total Market Index TR (al lordo delle imposte). Ulteriore evidenza empirica a favore dell'*equal weighting*.

Gli indici della serie Averages, calcolati dalla Dow Jones Indexes, sono stati oggetto di analisi meramente per la loro importanza tradizionale nella divulgazione finanziaria, essendo essi costruiti tramite la ponderazione per il prezzo, priva di fondamenti teorici e, almeno oggi, di limitatissima utilità nella semplificazione del loro calcolo. Particolarmente significativo il valore assunto dal flusso netto di dominanza del Dow Jones Industrial Average PR: -1 . Il più antico indice del mercato azionario statunitense, tuttora ampiamente citato dai *mass media*, è quindi il meno efficiente in assoluto, essendo dominato da tutti gli altri indici oggetto di questo studio.

Gli MSCI USA presentano, indipendentemente dal trattamento dei proventi, un grado di efficienza in linea con gli indici S&P 500 e inferiore solamente rispetto a quello dei più diversificati Russell e Wilshire.

4. Conclusioni

Alla luce delle analisi effettuate, è possibile concludere che la scelta di indici del mercato azionario, sia allo scopo di replica passiva sia allo scopo di *performance evaluation* di un portafoglio attivo, richiede un'approfondita analisi delle loro caratteristiche costruttive.

Le tecniche di valutazione impiegate in questo studio hanno consentito di giungere a risultati sostanzialmente univoci, seppure sia necessario evidenziare la limitata capacità discriminante dimostrata dal *test* GRS in condizioni di accentuata non-normalità. Si ritiene quindi che sia opportuno il ricorso a tecniche di simulazione, come quelle impiegate nei *test* di Wald e GMM e nell'applicazione del *resampling* alla misura di Kandel e Stambaugh, date le conformazioni delle distribuzioni dei rendimenti dei titoli azionari e le problematiche derivanti dalla stima dei loro parametri sulla base di serie storiche di durata limitata.

geometrica e mediana coincidono in presenza di log-normalità. Va infatti ricordato che prima dell'utilizzo di programmi informatici per la statistica la stima della mediana era molto più gravosa rispetto al calcolo della media geometrica.

Tra le caratteristiche costruttive degli indici, il trattamento dei proventi è risultato cruciale per l'incremento dell'efficienza, sconsigliando quindi l'impiego sia degli indici di prezzo sia di quelli *gross total return*, i primi in quanto dominati dalle tecniche che prevedono il reinvestimento degli utili, i secondi solamente perché fondati su ipotesi di calcolo irrealistiche.

Ancor più rilevante appare il ruolo svolto dal metodo di ponderazione. Se già la teoria finanziaria escludeva la razionalità del *price weighting*, i risultati di questa analisi hanno attribuito una netta superiorità alle tecniche di equiponderazione rispetto al *cap weighting*. Queste evidenze empiriche, per quanto in netto contrasto con il Capital Asset Pricing Model, sono corroborate da un'ampia letteratura.

Come notato da Markowitz (2005), una volta che siano poste in discussione alcune delle ipotesi più irrealistiche del CAPM, anche nella sua versione più generale concepita da Fischer Black, il portafoglio ottimo non è più necessariamente ponderato per la capitalizzazione dei suoi componenti. Treynor (2005) ha proposto la Noisy Market Hypothesis, secondo la quale, in presenza di errori di *pricing* anche limitati, i portafogli *cap weighted* affidano necessariamente un peso più elevato ai titoli sopravvalutati.

Oltre alle critiche poste al CAPM, il comportamento non strettamente razionale degli investitori può contribuire a generare una richiesta più elevata per portafogli equiponderati. Secondo la finanza comportamentale possono essere principalmente tre le ragioni a favore di un investimento *equally weighted*. Fisher e Statman (1997a, 1997b) hanno valutato le prassi d'investimento dei piani pensionistici statunitensi, evidenziando comportamenti che possono essere ricondotti alla Regret Theory. In condizioni di forte incertezza, l'investitore potrebbe trovare troppo gravoso il peso di una scelta non ottimale, preferendo quindi un'allocazione meccanica secondo la regola $1/N$.

L'euristica della diversificazione, di cui l'equiponderazione è un caso particolare, è stata portata come esempio da parte di Benartzi e Thaler (2001). Tale euristica consiste nella tendenza degli investitori a diversificare in modo *naive* quando si trovano in situazioni che richiedano l'esecuzione di una scelta multipla simultanea. Al contrario, allorché le decisioni possono avvenire sequenzialmente, essi sono propensi alla massimizzazione della loro utilità selezionando di volta in volta ciò che ritengono migliore, potendo agire anche sulla base dell'esperienza accumulata in precedenza. E' evidente come il processo di *asset allocation*, specialmente se concepito come propedeutico a un investimento uniperiodale, rientri nella categoria delle decisioni multiple simultanee.

L'ultimo apporto comportamentale a favore dell'equiponderazione è stato fornito da Windcliff e Boyle (2004), che per primi hanno collegato l'euristica della diversificazione al problema dell'*estimation error*, valutandola come una forma di approccio bayesiano semplificato caratterizzato dall'impiego di *prior* identici.

Questa interpretazione può essere vista come il *trait d'union* con il fondamentale studio di DeMiguel, Garlappi e Uppal (2009) riguardante l'impatto dell'*estimation error* sulla *performance* dei portafogli ottimizzati. Gli autori, sulla base di un'estesa analisi empirica, hanno riscontrato una netta dominanza, *ex post*, dei portafogli equiponderati rispetto alle altre forme di allocazione.

Il loro contributo ha una portata unificante riguardo sia alle ipotesi di piena razionalità degli operatori economici sia rispetto all'opposta concezione tipica della finanza comportamentale. Esso, infatti, consente di ricondurre a un paradigma di razionalità le scelte di *asset allocation* apparentemente subottimali compiute da molti investitori, una volta che sia tenuta nella dovuta considerazione l'incertezza sui parametri delle distribuzioni dei rendimenti.

Bibliografia

- Basile, Ignazio (a cura di) (2002) *Benchmark e performance dei portafogli azionari e obbligazionari*. Roma: Bancaria Editrice.
- Benartzi, Shlomo, Richard H. Thaler (2001) Naive Diversification Strategies in Defined Contribution Saving Plans. *The American Economic Review*, 91 (1), 79-98.
- Black, Fischer, Michael C. Jensen, Myron Scholes (1972) The Capital Asset Pricing Model: Some Empirical Tests. In Michael C. Jensen (a cura di) *Studies in the Theory of Capital Markets*, New York, NY: Praeger, 79-121.
- Brans, Jean-Pierre (1982) *L'ingénierie de la décision. Elaboration d'instruments d'aide à la décision. Méthode PROMETHEE*, Québec: Université Laval.
- Campbell, John Y., Andrew W. Lo, A. Craig MacKinlay (1997) *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.
- Chou, Pin-Huang, Guofu Zhou (2006) Using Bootstrap to Test Portfolio Efficiency. *Annals of Economics and Finance*, 7 (2), 217-249.
- DeMiguel, Victor, Lorenzo Garlappi, Raman Uppal (2009) Optimal Versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy? *The Review of Financial Studies*, 22 (5), 1915-1953.
- Fisher, Kenneth L., Meir Statman (1997a) The Mean-Variance-Optimization Puzzle: Security Portfolios and Food Portfolios. *Financial Analysts Journal*, 53 (4), 41-50.
- Fisher, Kenneth L., Meir Statman (1997b) Investment Advice from Mutual Fund Companies. Closer to the Talmud than to Markowitz. *The Journal of Portfolio Management*, 24 (1), 9-25.
- Gibbons, Michael R., Stephen A. Ross, Jay Shanken (1989) A Test for the Efficiency of a Given Portfolio. *Econometrica*, 57 (5), 1121-1152.
- Grinold, Richard C. (1992) Are Benchmark Portfolios Efficient? *The Journal of Portfolio Management*, 19 (1), 34-40.
- Hansen, Lars Peter (1982) Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50 (4), 1029-1054.
- Kandel, Shmuel, Robert F. Stambaugh (1995) Portfolio Inefficiency and the Cross-section of Expected Returns. *The Journal of Finance*, 50 (1), 157-184.
- Lintner, John (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), 13-37.
- MacKinlay, A. Craig, Matthew P. Richardson (1991) Using Generalized Method of Moments to Test Mean-Variance Efficiency. *The Journal of Finance*, 46 (2), 511-527.
- Markowitz, Harry M. (1952) Portfolio Selection. *The Journal of Finance*, 7 (1), 77-91.
- Markowitz, Harry M. (2005) Market Efficiency: A Theoretical Distinction and So What? *Financial Analysts Journal*, 61 (5), 17-30.
- Michaud, Richard O., Robert O. Michaud (2008) *Efficient Asset Management*. 2nd ed. Oxford: Oxford University Press.
- Mossin, Jan (1966) Equilibrium in a Capital Asset Market. *Econometrica*, 34 (4), 768-783.
- Muirhead, Robb J. (1982) *Aspects of Multivariate Statistical Theory*. New York, NY: John Wiley & Sons.

- Roll, Richard (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), 129-176.
- Sentana, Enrique (2009) The Econometrics of Mean-variance Efficiency Tests: a Survey. *Econometrics Journal*, 12 (3), C65-C101.
- Shanken, Jay (1996) Statistical Methods in Tests of Portfolio Efficiency: A Synthesis. In *Handbook of Statistics, Vol. 14: Statistical Methods in Finance*, a cura di G. S. Maddala, C. R. Rao. Amsterdam: Elsevier Science.
- Shao, Jun, Dongsheng Tu (1995) *The Jackknife and Bootstrap*. New York, NY: Springer Verlag.
- Sharpe, William F. (1964) Capital Asset Prices: A Theory of Market Equilibrium. *The Journal of Finance*, 19 (3), 425-442.
- Treynor, Jack (2005) Why Market-Valuation-Indifferent Indexing Works. *Financial Analysts Journal*, 61 (5), 65-69.
- Windcliff, Heath, Phelim P. Boyle (2004) The $1/n$ Pension Investment Puzzle. *North American Actuarial Journal*, 8 (3), 32-45.