



Munich Personal RePEc Archive

The asset allocation of hedge funds during the financial crisis: an empirical investigation

Piluso, Fabio and Amerise, Ilaria Lucrezia

10 January 2011

Online at <https://mpra.ub.uni-muenchen.de/28178/>

MPRA Paper No. 28178, posted 19 Jan 2011 20:42 UTC

L'asset allocation dei fondi hedge durante la crisi finanziaria: un'analisi empirica

PILUSO F.^a – AMERISE I.L.^b

^a Dipartimento di Scienze Aziendali, Facoltà di Economia - Università della Calabria, Via P. Bucci, Rende, Cosenza, 0984492848 – fabio.piluso@unical.it

^b Dipartimento di Economia e Statistica, Facoltà di Economia - Università della Calabria, Via P. Bucci, Rende, Cosenza, 0984492402 – iamerise@economia.unical.it

Abstract

La crisi finanziaria innescata a partire dal 2007 ha individuato nei fondi hedge i suoi principali responsabili. Eppure i dati di mercato evidenziano che se in molti casi tali fondi hanno prodotto enormi perdite, in tanti altri hanno realizzato enormi profitti.

Obiettivo del presente paper è quello di analizzare il mercato londinese dei fondi hedge allo scopo di mappare gli stessi in base alle strategie utilizzate dai diversi gestori e arrivare così a individuare le tecniche che maggiormente si avvicinano alla frontiera efficiente di rischio-rendimento durante periodi di panico finanziario o di forte volatilità dei mercati.

A tal fine, si è deciso di clusterizzare un campione di 788 fondi inglesi attraverso il metodo di Ward (1963). Noi crediamo che la forte instabilità dei mercati dovrebbe favorire i fondi che investono in beni rifugio (commodity) e quelli che sfruttano l'alta volatilità dei titoli quotati (strategia long/short).

Key words: hedge fund, investment strategy, cluster analysis, financial crisis

JEL Classification: G23, G11, C10, G01

INTRODUZIONE

Gli hedge fund, o fondi speculativi, animano da tempo un acceso dibattito sulla loro gestione e sull'opportunità di eventuali regolamentazioni e restrizioni.

Accusati, da alcuni (i governi di Francia e Germania), di aver in buona parte causato, o quanto meno amplificato, l'attuale crisi finanziaria, gli hedge fund continuano però a riscuotere consensi nei Paesi anglosassoni innescando un confronto che diventa sempre più politico oltre che economico. In effetti, già nel corso del G8 del 2007, i ministri finanziari hanno discusso e approvato il "Rapporto sull'Industria degli Hedge Funds" presentato dal Governatore della Banca d'Italia, Mario Draghi, in qualità di Presidente del Financial Stability Forum¹.

In tale documento, tuttavia, non viene fatto alcun esplicito riferimento ad un eventuale "codice di condotta" richiesto dalla presidenza tedesca, ma fronteggiato da Stati Uniti e Gran Bretagna, i due Paesi dove opera la maggioranza dei fondi, per i quali il codice di condotta poteva significare solamente regolamentazione stringente e precisa dei fondi hedge.

Il dibattito e il confronto sul tema è stato ultimamente ripreso in seguito all'emanazione da parte degli Stati Uniti del "Private Fund Transparency Act of 2009" che andrà a sostituire l'"Investment Advisers Act of 1940" costringendo così i gestori di fondi hedge a registrarsi presso uno speciale elenco stilato dalla Sec al fine di proteggere gli investitori e il mercato da nuovi possibili rischi sistemici. Ancora una volta, però, nessun accenno è stato fatto a una possibile regolamentazione stringente che, nelle valutazioni di Stati Uniti, Regno Unito, Svezia e (più recentemente) della stessa BCE, potrebbe indurre i grandi gestori a emigrare verso mercati "più liberi" danneggiando così il sistema finanziario europeo, già di per sé in crisi di fiducia e di liquidità.

Data la complessità, oltre che la già accennata attualità del tema, i fondi hedge sono stati oggetto di un'ampia letteratura in ambito internazionale nel corso degli ultimi anni. Diversi sono stati gli obiettivi che gli accademici e i ricercatori si sono posti quale nucleo centrale della propria analisi, ma al momento non risulta uno studio approfondito sul confronto dei fondi hedge in relazione alle strategie di investimento utilizzate in un arco temporale definito e su una piazza finanziaria circoscritta. Obiettivo del presente lavoro è perciò quello di analizzare il rapporto rischio rendimento ottenuto dai fondi hedge sulla piazza finanziaria di Londra in relazione alle diverse strategie di investimento seguite dai gestori in un periodo cruciale per l'operatività di tali intermediari (considerata la crisi finanziaria in atto).

¹ Il documento esaminava i potenziali rischi connessi alle attività dei fondi e formulava "cinque raccomandazioni" rivolte ad autorità, investitori e gestori per favorire il mantenimento della stabilità dei mercati finanziari e per migliorare la trasparenza sul mercato. Il documento suggeriva un approccio "indiretto" alla vigilanza su tali operatori realizzato attraverso l'attenta collaborazione di banche e intermediari che operano con gli hedge funds e la definizione di linee guida per migliorare le pratiche del settore.

Ci si attende che l'alta volatilità dei mercati finanziari durante tale periodo abbia dovuto favorire quei gestori che guadagnano in tali condizioni di mercato (strategia *long/short*) nonché quei gestori che hanno puntato sul trading di beni rifugio (strategia *cta/managed futures*).

Ai fini della nostra analisi si è ritenuto utile applicare la tecnica della cluster analysis in luogo delle tradizionali tecniche di ottimizzazione media-varianza. Ciò anche in relazione ai risultati di un recente studio prodotto da The Bank of New York Mellon e Oxford Metrica² che dimostra la superiorità dei risultati raggiunti attraverso l'utilizzo della cluster analysis rispetto alle tecniche tradizionali di massimizzazione nel comparto degli hedge fund.

La presente analisi si basa pertanto sulla tecnica della clusterizzazione dei fondi in base alle strategie e si pone l'obiettivo di individuare la tipologia di fondo che ottimizza il rapporto rischio - rendimento. All'interno del cluster "ottimale", si individua altresì il fondo che, nello specifico, realizza il punteggio più alto in relazione alle variabili considerate nella presente analisi.

L'applicazione della cluster analysis è stata articolata in due fasi:

- in primo luogo viene effettuata la scelta delle variabili di classificazione delle unità osservate. Tale scelta, inevitabilmente, rispecchia le ipotesi alla base della ricerca. Sebbene si tratti di una operazione che implica un grado molto alto di soggettività la scelta delle variabili è coerente con il tipo di relazioni alla cui identificazione è orientata la ricerca;
- una volta identificate le variabili sulle quali operare, il secondo step ha riguardato la scelta di una misura della dissomiglianza esistente fra le unità statistiche³. Poiché le variabili in input hanno carattere quantitativo possono essere utilizzati vari tipi di indici di distanza.

La struttura del lavoro presenta la seguente sequenza logico-operativa. Nel secondo paragrafo viene presentata una selezionata letteratura sul tema dei fondi hedge e i principali risultati raggiunti. Nel terzo paragrafo viene brevemente presentato l'andamento a livello globale del settore nel corso degli anni 2007 e 2008. Nel quarto si analizza la metodologia utilizzata per l'analisi. Nel quinto si espongono i risultati ottenuti. Nel sesto, infine, vengono commentati i risultati ottenuti.

² Cfr. Knight R., Whitaker J. (2008).

³ In termini generali i caratteri statistici possono essere espressi in quattro distinte scale di misura: nominali, ordinali, per intervalli e per rapporti. I caratteri qualitativi possono essere misurati solo con riferimento alle prime due, mentre le variabili ammettono scale di qualunque tipo.

ANALISI DELLA LETTERATURA

E' bene da subito precisare come sul tema dei fondi hedge esista un'ampia letteratura internazionale tesa a verificare diversi aspetti di un comparto abbastanza opaco e, per certi versi, poco trasparente.

Una prima linea di ricerca ha affrontato il tema del confronto tra le performance realizzate dai fondi hedge rispetto ai fondi tradizionali.

A tal riguardo, Ackermann et al. (1999) e Liang (1999) hanno evidenziato come gli hedge fund costantemente sovra-performino i fondi comuni. Ackermann e Ravenscraft (1998) hanno sottolineato come le minori performance dei fondi comuni siano dovute alle limitazioni di investimento imposte dalla normativa. Agarwal e Naik (2000), dal canto loro, hanno dimostrato l'esistenza di una certa persistenza nelle maggiori performance fatte registrare dai fondi hedge rispetto ai fondi comuni. Questo aspetto legato alla persistenza delle performance è particolarmente importante nel caso degli hedge fund, come evidenziato da Brown et al. (1999), Goetzmann, Hiraki, Otsuki e Shiraishi (2001), Brown, Goetzmann e Park (1997) e Liang (1999).

Una seconda linea di ricerca ha riguardato l'analisi della relazione rischio-rendimento di un portafoglio che include fondi speculativi.

In tal senso, Fung e Hsieh (1997), Schneeweis e Spurgin (1997) e Cowell (2003) hanno dimostrato che gli hedge fund possono sensibilmente migliorare il profilo rischio/rendimento del portafoglio degli investitori grazie al basso livello di correlazione con i fondi comuni e le altre asset class tradizionali. L'esistenza di un basso livello di correlazione è dimostrato anche da Liang (1999) e Agarwal e Naik (1999). Amin e Kat (2002) arrivano a concludere che affinché si abbia un impatto significativo (in termini di rischio – rendimento) sul portafoglio degli investitori è necessario allocare ai fondi hedge una percentuale decisamente superiore al range 1%-5% che tipicamente viene scelta per questo tipo di asset class. Tuttavia, il miglioramento del profilo rischio-rendimento non è a costo nullo, poiché comporta una *skewness*⁴ inferiore e una maggiore *kurtosis*⁵ [Amin e Kat, 2002]. Agarwal e Naik (2001), Fung e Hsieh (2001), Lo (2001), Brooks e Kat (2002) e Favre e Galeano (2002), però, hanno messo in evidenza come gli elementi di rischio e volatilità che contraddistinguono gli hedge fund sono molto più complessi rispetto a quelli che contraddistinguono i titoli azionari e obbligazionari.

⁴ La skewness è una misura della asimmetria di una distribuzione normale. Se il suo valore è pari a zero, ciò implica che la distribuzione è perfettamente simmetrica e la media coincide con la mediana. Se il suo valore è diverso da zero, allora la distribuzione potrà essere concentrata maggiormente in uno dei due lati della stessa risultando quindi asimmetrica.

⁵ La kurtosis è un parametro che descrive lo spessore delle code di una funzione di densità della probabilità di una variabile casuale, ovvero il grado di appiattimento di una distribuzione. Se la kurtosis supera il valore di 3, la stessa è definita leptocurtica (maggiore allungamento della distribuzione), mentre per valori inferiori a 3 si definisce platicurtica (maggiore appiattimento della distribuzione).

Una terza linea di analisi ha valutato la capacità dei gestori di fondi hedge di ottenere extra-rendimenti rispetto ai fondi tradizionali.

All'uopo, Brown et al. (1999) e Agarwal e Naik (2000) hanno utilizzato una serie di dati sugli hedge fund americani senza tuttavia aver trovato alcuna evidenza sul fatto che le capacità dei gestori possano influenzare i rendimenti ottenuti dai fondi hedge.

Una quarta linea di indagine ha affrontato il tema dell'individuazione dei modelli di ottimizzazione del portafoglio dei fondi speculativi.

A tale tema la letteratura riserva modelli di ottimizzazione tradizionali media-varianza [Markowitz, 1952; Amin e Kat, 2003; Hoevenaars, Molenaar, Schotman e Steenkamp, 2005] e modelli di ottimizzazione più avanzati [Cvitanić, Lazrak, Martellini e Zapatero, 2002; Favre e Galeano, 2002; Davies, Kat e Lu, 2004; Morton, E. Popova e I. Popova, 2006]. I primi ipotizzano che i rendimenti delle attività finanziarie considerate siano distribuiti secondo una funzione normale; i secondi invece considerano che i rendimenti di un titolo possono non avere una distribuzione normale e quindi l'analisi è estesa anche ai momenti di terzo e quarto grado, rappresentati rispettivamente dall'asimmetria (*Skewness*) e dalla curtosi (*Kurtosis*) delle suddette distribuzioni. Fung e Hsieh (2000 e 2001), Amin e Kat (2003), Ghiringhelli (2004), Lhabitant (2004) hanno dal canto loro messo in evidenza come i fondi hedge presentino in generale una distribuzione dei rendimenti non normale perché risulta asimmetrica e leptocurtica. Più precisamente, Ghiringhelli (2004) e Amin e Kat (2003) hanno appurato che i fondi hedge presentano una distribuzione dei rendimenti leptocurtica e asimmetrica negativa. Alle stesse conclusioni sono giunti anche Brooks e Kat (2002) e Lamm, (2003).

I modelli di ottimizzazione tradizionali che la letteratura riserva alla definizione di asset allocation efficiente (ovvero i modelli di Amin e Kat del 2003 e di Hoevenaars, Molenaar, Schotman e Steenkamp del 2005) si rifanno al modello su cui poggia la moderna teoria del portafoglio: il modello media - varianza di Markowitz (1952). Per quanto riguarda i modelli di ottimizzazione innovativi - dall'analisi della letteratura - sono risultati di particolare interesse il modello di Davies, Kat e Lu (2004) e quello di Favre e Galeano (2002) poiché entrambi questi modelli hanno mostrato di adattarsi alle peculiarità che caratterizzano gli hedge fund.⁶

Una quinta linea di analisi ha riguardato la consistenza dei fondi hedge in relazione alla strategia di investimento.

⁶ Cfr. Marco Giorgino, Diego Viviani, *Hedge funds* e fondi comuni non armonizzati: come migliorare l'efficienza dei portafogli

Malkiel e Saha (2005)⁷ rilevano come il gruppo più ampio di hedge funds sia quello della categoria long/short, che generano rendimenti assumendo posizioni “long” su titoli sottovalutati, e si posizionano con posizioni *short* su titoli sopravvalutati; la maggior parte di questa tipologia di fondi, inoltre, cercava di essere *equity market neutral*, cioè non influenzata dall’andamento del mercato. Koh, Koh e Teo (2003) analizzano inoltre la correlazione tra i ritorni economici degli hedge fund e alcune loro caratteristiche - come la dimensione, il tipo di compenso dato ai gestori, l’età, le dimensioni delle società possedute e il periodo di rimborso - attraverso l’analisi un dataset dettagliato sui fondi asiatici. Gli stessi autori analizzano la relazione tra i ritorni e le caratteristiche dei fondi attraverso l’utilizzo del modello di Fama e MacBeth (1973) ed effettuano una regressione univariata e multivariata dei ritorni mensili sulle caratteristiche degli stock. Le regressioni vengono effettuate sia sui ritorni post-pagamento che su quelli pre-pagamento. In entrambe le regressioni le variabili indipendenti risultano fortemente correlate. Esistono infine delle banche dati di operatori indipendenti – quali Hedge Fund Research ed EuroHedge – che si occupano del peso delle strategie di investimento e della loro evoluzione nel corso degli anni.

IL MERCATO DEGLI HEDGE FUND NEL 2007 E NEL 2008

Dall’analisi dei rendimenti dei fondi hedge nel corso degli anni 2007 e 2008 si ottengono le seguenti informazioni. Nel 2007, nonostante l’elevata volatilità dei mercati finanziari, i gestori sono riusciti a sfruttare le fasi di ripresa e a proteggersi in gran parte dalle fasi di mercato in calo riuscendo a ottenere rendimenti medi superiori al 10%.

Tra gli andamenti più importanti vi è la performance decisamente migliorata della strategia *global macro*, che è riuscita a compensare la persistente debolezza delle strategie sensibili al credito.

A sei mesi dall’inizio del 2008 le strategie che hanno, invece, registrato le migliori performance sono proprio quelle che hanno scommesso sul ribasso dei mercati (*short seller* +2.69%) e quelle che hanno saputo sfruttare la volatilità di materie prime e valute. In questo secondo caso, i gestori *commodity trading advisor/managed futures* hanno ottenuto un rialzo del 9.60%, ovvero dieci volte quello ottenuto dai fondi *long/short equity* (+0.86%) nello stesso periodo di tempo. Buoni risultati sono inoltre stati offerti dai fondi *global macro* (+7.08%).

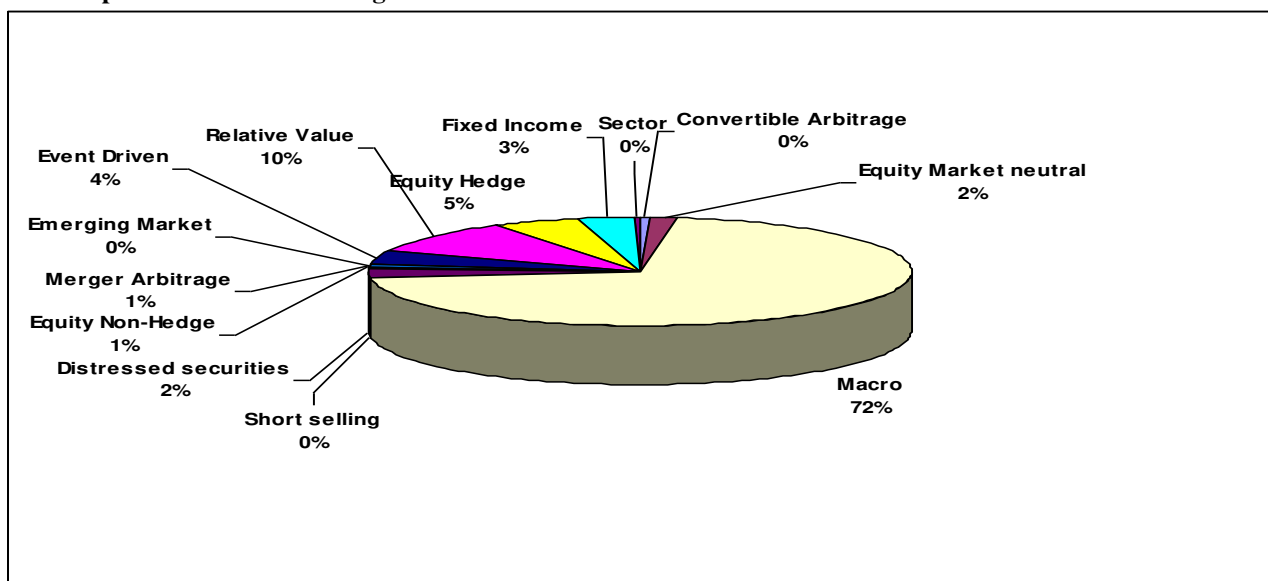
Tra le strategie “non vincenti” vanno collocate le *convertible arbitrage* (-5.22%) che hanno risentito del limitato numero di emissioni nel periodo e delle poche opportunità offerte dal mercato. Stessa sorte per chi ha puntato sul *reddito fisso* (-3.65%), così come i gestori che investono sui *mercati emergenti*, i quali registrano ancora una perdita di quasi due punti percentuali. In questo caso il

⁷ B. G. Malkiel e A. Saha, Hedge funds: Risk and return, Financial analysts journal, 2005.

mercato sta riassorbendo parte della “sbornia asiatica” che negli anni passati ha generato rendimenti positivi rispettivamente del 20.26% nel 2007, del 20.49% nel 2006 e del 17.39% nel 2005. Secondo le statistiche annuali fornite dall’Ufficio Studi Alternative Investment⁸ fra le strategie che hanno conseguito le migliori performance annualizzate registrate nel periodo gennaio 1994 - aprile 2008 troviamo le strategie *global macro* (+13.80%), *long/short equity* (+11.74%), *event driven* (+11.36%). Le altre categorie hanno ottenuto un rendimento compreso fra il 5% e il 10%; soltanto la strategia *dedicated short bias* ha riportato un rendimento negativo (-1.65%). Quest’ultima, infatti, ha presentato un Sharpe Ratio pari a -0.33. Tra le strategie che hanno mostrato una bassa volatilità e un elevato Sharpe Ratio si riscontrano la *multistrategy*, *l’event driven*, *l’equity market neutral* e la *convertible arbitrage*.

In particolare la strategia *market neutral* è quella che ha ottenuto il miglior risultato in termini di rischio-rendimento, registrando una performance pari a 9.88% e la più bassa percentuale di deviazione standard (2.80%). La strategia *emerging markets*, invece, ha registrato ottimi rendimenti ma una volatilità estremamente elevata (15.44%). Cambia, infine, il peso delle diverse strategie di investimento nel corso dell’ultimo ventennio. In effetti, se ad esempio, nel 1990 i fondi *macro* gestivano oltre il 70% delle attività, nel 2007 riducono il loro peso al 10% a vantaggio di altre categorie, quali *equity hedge* (27,68%) ed *event driven* (13,68%).

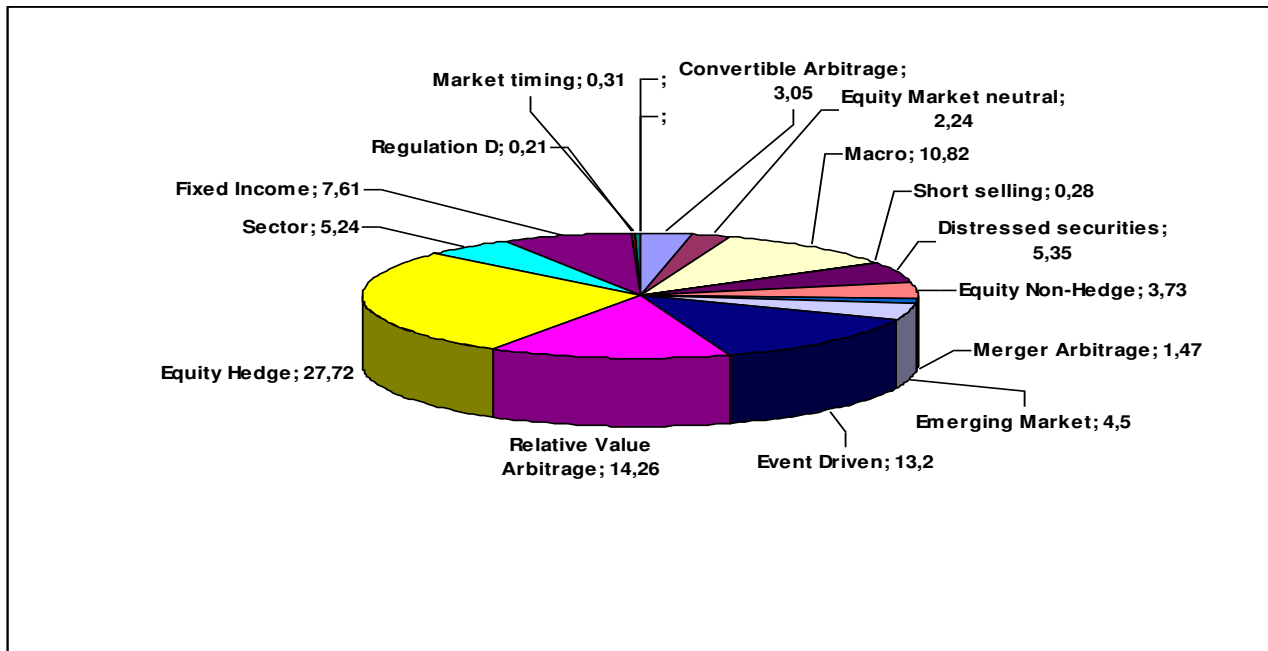
Tab. 1 Il peso delle diverse strategie a livello mondiale nel 1990



Fonte: Database Hedge Fund Research. Percentuali calcolate sul patrimonio in gestione

⁸ Cfr. www.studiobba.it/ufficiostudi/Documenti/Report%20Alternative.pdf

Tab. 2 Il peso delle diverse strategie a livello mondiale nel 2007



Fonte: Database Hedge Fund Research. Dati stimati aggiornati al terzo trimestre 2007. Percentuali calcolate sul patrimonio in gestione

Tra i Paesi europei, il Regno Unito si conferma leader in Europa con un peso del 77%, seguito da Francia (5,7%), Svizzera (3,5%), Svezia (2,9%) e Usa (2,8%). Tale dato evidenzia la differenza abissale tra il primo e il secondo paese europeo per peso di fondi hedge e giustifica l'idea della presente ricerca di analizzare esclusivamente il mercato del Regno Unito (eventualmente da considerare quale *proxy* del mercato europeo).

Tab. 3 I primi 5 Paesi dove sono localizzati i fondi single manager europei

	Patrimonio Gestito	Numero di fondi	Peso
UK	414,824	962	77,0%
Francia	30,850	114	5,7%
Svizzera	18,576	101	3,5%
Svezia	15,807	52	2,9%
Usa*	15,132	41	2,8%

Fonte: Database EuroHedge. Dati aggiornati al secondo trimestre 2007. Dati in miliardi di dollari.

* Società di gestione operative in Europa, ma localizzate in Usa

OBIETTIVI DELLA RICERCA E METODOLOGIA UTILIZZATA

L'analisi condotta sui fondi hedge presenti sulla piazza finanziaria di Londra utilizza le seguenti variabili indipendenti:

1. i rendimenti dei fondi negli anni 2007 e 2008;
2. la deviazione standard, rappresentativa della volatilità del fondo;
3. lo Sharpe ratio⁹;
4. il Sortino ratio¹⁰;
5. il maximum draw down¹¹;
6. la dimensione del fondo.

La mappatura dei fondi avviene attraverso l'utilizzo delle seguenti strategie di investimento, considerate variabili dipendenti: *Arbitrage, Bottom-Up, CTA / Managed Futures, Distressed Debt, Diversified Debt, Dual Approach, Event Driven, Fixed Income, Long / Short Equities, Macro, Multi-Strategy, Others, Relative Value, Top-Down*.

Come si può evincere dalle variabili scelte, l'analisi pone in evidenza la ricerca di una frontiera efficiente per ciò che attiene alle diverse strategie, ponendo a confronto il rapporto tra i rendimenti e i diversi indici di rischiosità (volatilità, Sharpe ratio, Sortino ratio, maximum draw down), oltre che la dimensione del fondo stesso. La tecnica di analisi che viene utilizzata è la *cluster analysis*.

Per analisi dei gruppi (cluster analysis)¹² si intende l'insieme delle procedure che permettono di individuare, entro un insieme di oggetti di natura qualsiasi, alcuni sottoinsiemi (gruppi, classi,

⁹ L'indice di Sharpe misura l'extra-rendimento di un portafoglio titoli (rispetto a un tasso risk free) in rapporto alla volatilità del portafoglio stesso espressa dalla deviazione standard. Si viene così a indicare il rendimento in termini percentuali rispetto a ogni unità di rischio del nostro investimento.

⁹ L'indice di Sortino è anch'esso un indicatore di rischio finanziario e si basa sulla probabilità che gli extra rendimenti (rispetto al tasso risk free) di un investimento si posizionano al di sopra di quello che l'investitore considera il "minimo accettabile". Un elevato indice di Sortino indica che la variabilità dei rendimenti non si concentra prevalentemente al di sotto del minimo ritenuto accettabile dall'investitore (basso down side risk), viceversa un basso indice di Sortino indica che la variabilità si concentra sotto il minimo accettabile (alto down side risk).

¹¹ Il maximum draw down rappresenta pure un indicatore di rischio e calcola la perdita percentuale massima di una strategia di investimento in un determinato periodo calcolata su base empirica. Permette perciò di comparare la rischiosità di diverse strategie.

¹² La cluster analysis consente di pervenire ai seguenti risultati:

- la generazione di ipotesi di ricerca: infatti per effettuare una analisi di raggruppamento non è necessario avere in mente alcun modello interpretativo;
- la riduzione dei dati in forma (anche grafica) tale da rendere facile la lettura delle informazioni rilevate e parsimoniosa la presentazione dei risultati;
- la ricerca tipologica per individuare gruppi di unità statistiche con caratteristiche distintive che facciano risaltare la fisionomia del sistema osservato;
- la costruzioni di sistemi di classificazione automatica (Jardine N., Sibson R., 1971);
- la ricerca di classi omogenee, dentro le quali si può supporre che i membri siano mutuamente surrogabili (Green P.E., Frank R.E., Robinson P.J., 1967).

Per un approfondimento si veda Fabris L., (1997).

cluster o grappoli) tendenzialmente omogenei al loro interno e sufficientemente distinti fra loro, sulla base delle informazioni disponibili¹³.

Le unità statistiche, in altri termini, sono ripartite in un certo numero di gruppi a seconda del loro livello di "somiglianza". Quest'ultima è valutata a partire dai valori che una serie di variabili prescelte assume in ciascuna unità.

A differenza di altre tecniche statistiche multivariate (ad esempio, l'analisi discriminante¹⁴), l'analisi dei gruppi non si fonda su alcuna assunzione "a priori" sulle tipologie fondamentali esistenti che possono caratterizzare il collettivo studiato.

Le tecniche di clustering, proprio per questo, assumono estrema utilità per la loro valenza esplorativa e di ricerca di strutture e legami logici latenti nei dati (approccio data driven). Esse, in ogni caso, restano incardinate su metodologie puramente empiriche di classificazione e, in quanto tali, rientrano in un approccio di ricerca induttivo.

L'analisi condotta nei paragrafi seguenti mira ad assegnare i fondi osservati a categorie non definite a priori, formando gruppi di osservazioni omogenei al loro interno (within) ed eterogenei tra loro (between).

Elemento cruciale per la comprensione del funzionamento della tecnica gerarchica applicata è costituito dalla rappresentazione grafica della struttura di raggruppamento tramite *diagramma ad albero* (o *dendrogramma*).

Nelle tecniche di clustering, il dendrogramma è impiegato per fornire una rappresentazione grafica del processo di raggruppamento delle unità statistiche di partenza (records), esprimendo graficamente sia la distanza cardinale tra i clusters (secondo una metrica definita dall'indice in ascissa) sia il livello gerarchico di aggregazione (attraverso valori interi positivi in ordinata)¹⁵.

Poiché le variabili da utilizzare debbono essere espresse nella stessa unità di misura si è proceduto a standardizzare le variabili assunte in input. Si sono calcolate media μ e varianza σ^2 delle distribuzioni di partenza e, quindi, si sono ottenuti i punti zeta di ciascuna variabile in base alla seguente espressione:

$$Z = \frac{X - \mu}{\sigma}$$

¹³ Roberto Biorcio (1993), *L'analisi dei gruppi*, Ed. Franco Angeli

¹⁴ Attraverso l'analisi discriminante è possibile la ripartizione di un insieme di individui in gruppi, predeterminati fin dall'inizio della ricerca in base alle diverse modalità assunte da uno o più caratteri della popolazione di partenza.

¹⁵ La lettura del grafico ad albero avviene in base alla distanza che intercorre tra l'origine e la linea orizzontale più vicina che connette le linee verticali corrispondenti a due elementi qualsiasi. Al fine di comprendere quale sia la relazione che intercorre tra i due elementi, è necessario perciò tracciare un percorso da uno dei due elementi all'altro, seguendo i rami del diagramma e percorrendo il cammino più breve. La distanza dall'origine alla linea orizzontale più esterna toccata dal percorso rappresenta una misura del grado di somiglianza tra i due elementi considerati.

La tecnica utilizzata considera una matrice di dissomiglianza che contiene le informazioni riguardanti il grado di dissomiglianza tra le diverse unità statistiche. La dissomiglianza tra unità coincide con la distanza tra le stesse. Allo scopo è stata utilizzata come metrica la distanza quadratica euclidea¹⁶.

Quella euclidea corrisponde al concetto geometrico di distanza nello spazio multidimensionale. Il quadrato della distanza euclidea è stato utilizzato al fine di pesare in modo progressivamente maggiore gli “oggetti” che stanno oltre una certa distanza.

Dopo aver adottato una misura di dissomiglianza, si è proceduto a scegliere un idoneo algoritmo di raggruppamento delle unità osservate. La distinzione che normalmente viene proposta è fra :

- metodi gerarchici, che conducono ad un insieme di gruppi ordinabili secondo livelli crescenti, con un numero di gruppi da n ad 1 (dove n è la numerosità della popolazione statistica di partenza);
- metodi non gerarchici, che forniscono un'unica partizione delle n unità in g gruppi, e dove g deve essere specificato a priori.

Allo scopo di ottenere una preliminare indicazione circa il numero di partizioni annidate nella popolazione statistica di riferimento è stata applicata dapprima una tecnica gerarchica di analisi dei gruppi.

I metodi gerarchici si prestano ad essere utilizzati in situazioni in cui si hanno n grappoli di una sola unità per giungere, attraverso successive fusioni dei grappoli meno distanti tra di loro, ad una situazione in cui si ha un solo grappolo che contiene tutte le n unità della popolazione di partenza.

Risultato finale dei metodi gerarchici non è, quindi, una singola partizione delle n unità, ma una serie di partizioni che possono essere rappresentate graficamente attraverso un "dendogramma" o "diagramma ad albero" nel quale sull'asse delle ascisse è riportato il livello di distanza, mentre sull'asse delle ordinate sono riportate le singole unità. Ogni ramo del diagramma (linea verticale) corrisponde ad un grappolo.

La linea di congiunzione (orizzontale) di due o più rami individua il livello di distanza al quale i grappoli si fondono.

I metodi gerarchici si distinguono per il modo in cui, dopo la p -esima fusione, vengono calcolate le distanze tra il nuovo grappolo ed i rimanenti.

¹⁶ Formalmente, sia X una matrice dati $n \times k$, X_i il vettore k -dimensionale della i -esima osservazione ed x_{ih} il suo elemento generico, la distanza euclidea è data dalla relazione:

$$d_{ij} = \left[\sum_{h=1}^k (x_{ih} - x_{jh})^2 \right]^{\frac{1}{2}}$$

Gli algoritmi gerarchici proposti in letteratura (metodo del legame singolo, metodo del legame completo, metodo del legame medio, metodo del centroide, metodo di Ward, solo per ricordarne alcuni) si differenziano unicamente per il diverso criterio che regola la valutazione delle distanze tra i gruppi ai fini delle aggregazioni in serie.

La tecnica scelta per l'analisi qui condotta è l'algoritmo di Ward (1963).

Descrizione del metodo di Ward

Il procedimento di raggruppamento parte dalla matrice di dissomiglianza tra elementi e procede iterativamente in due passi:

- raggruppando gli elementi più somiglianti;
- calcolando la matrice di dissomiglianza tra gruppi e/o elementi, utilizzando una misura di distanza come criterio per stabilire la distanza dei gruppi dai singoli elementi e/o dagli altri gruppi.

La condizione di arresto si verifica quando tutti gli elementi sono aggregati in un unico cluster.

Il metodo di Ward differisce dagli altri in quanto suggerisce di riunire, ad ogni tappa del processo, i due gruppi dalla cui fusione deriva il minimo incremento possibile della devianza data dalla relazione:

$$DEV_T = \sum_{s=1}^p \sum_{i=1}^n (x_{is} - \bar{x}_s)^2 = \sum_{i=1}^n \sum_{s=1}^p (x_{is} - \bar{x}_s)^2$$

dove \bar{x}_s è la media della variabile s con riferimento all'intero collettivo.

Data una partizione in g gruppi, tale devianza può essere scomposta in:

$$DEV_W = \sum_{k=1}^g \sum_{s=1}^p \sum_{i=1}^{n_k} (x_{is} - \bar{x}_{s,k})^2$$

che è la devianza entro i gruppi riferita alle p variabili con riferimento al gruppo k , dove $\bar{x}_{s,k}$ è la media della variabile s con riferimento al gruppo k ;

$$DEV_B = \sum_{s=1}^p \sum_{k=1}^g (\bar{x}_{s,k} - \bar{x}_s)^2 n_k$$

che è la devianza tra i gruppi.

Risulta che:

$$DEV_T = DEV_B + DEV_W$$

Nel passare da k+1 a k gruppi (aggregazione) DEV_w aumenta, mentre ovviamente DEV_B diminuisce. Ad ogni passo del metodo di Ward si aggregano tra loro quei gruppi per cui vi è il minor incremento di DEV_T.

La “best partition” è la suddivisione cui corrisponde la massimizzazione del rapporto

$$best = \max \left(\frac{tr(B)}{tr(W)} \right)$$

Dove

Tr(W) = traccia della matrice delle devianze e codevianze interne nei gruppi (within groups)

Tr(B) = traccia della matrice delle devianze e codevianze fra i gruppi (between groups)

Nel caso di una cluster gerarchica la scelta del numero di cluster può essere effettuata utilizzando in primo luogo la distanza di fusione.

Allo scopo è stato sezionato il dendrogramma ottenuto attraverso il metodo di Ward in corrispondenza di determinati livelli di dissomiglianza in modo da ottenere una partizione in gruppi disgiunti e omogenei dell'insieme di unità di partenza.

La distanza di fusione, in termini di distanza “rescaled”, può essere facilmente evinta dall'osservazione del dendrogramma: se nel passaggio da K gruppi a K+1 si registra un forte incremento della distanza di fusione si deve sezionare la popolazione in modo da ripartirla in K gruppi.

È opportuno sottolineare che tali metodi di scelta del numero di gruppi si basano comunque su una osservazione dei dati alla ricerca di una loro discontinuità, e questo può risultare una procedura azzardata e soggettiva¹⁷.

DATI E RISULTATI

Il dataset utilizzato per la scelta del collettivo statistico è stato tratto da EurekaHedge Hedge Fund, un database europeo fondato nel 2003. I fondi presenti nel database EurekaHedge Hedge Fund sono 1081; si è ritenuto tuttavia opportuno, ai fini della presente analisi, escludere le osservazioni con dati mancanti. Pertanto sono stati complessivamente osservati 788 fondi hedge presenti sulla piazza finanziaria di Londra suddivisi in base alle strategie adottate secondo lo schema riportato di seguito:

Variabile	Modalità	Frequenze	%
Investment Strategy	Arbitrage	28	3,553
	Bottom-Up	83	10,533
	CTA / Managed Futures	35	4,442
	Distressed Debt	7	0,888
	Diversified Debt	9	1,142
	Dual Approach	11	1,396
	Event Driven	30	3,807
	Fixed Income	46	5,838
	Long / Short Equities	394	50,000
	Macro	40	5,076
	Multi-Strategy	25	3,173
	Others	30	3,807
	Relative Value	32	4,061
Top-Down	18	2,284	

Dei citati fondi sono state osservate le seguenti variabili: i rendimenti dei fondi negli anni 2007 e 2008, la deviazione standard, lo Sharpe ratio, il Sortino ratio, il maximum draw down e la dimensione del fondo. Si è proceduto, in seguito, ad effettuare un'analisi interclasse, ovvero

¹⁷ Alcuni indici consentono una valutazione maggiormente oggettiva riguardo la scelta del numero di gruppi. Tra questi l'indice di Calinsky e Harabasz (1974), dato dall'espressione:

$$C_r = \frac{tr(B)}{tr(W)} \frac{n-r}{r-1}$$

L'indice viene rappresentato graficamente in funzione del numero di gruppi r . Se il grafico presenta un punto di massimo questo coincide con la numerosità della partizione ottimale. Se il grafico è sempre crescente si può concludere che le osservazioni non presentano una struttura favorevole alla suddivisione in gruppi. Se al contrario il grafico risulta sempre decrescente le osservazioni si prestano ad una classificazione di tipo gerarchico.

all'interno della stessa strategia si è individuato il fondo che, in base alle variabili sopra citate, ha realizzato il miglior rapporto rischio-rendimento.

Al fine di valutare la sostenibilità del metodo, si è ritenuto opportuno procedere con dei test preliminari; in primo luogo l'ACP (Cfr. sezione 1 in appendice), il cui scopo primario è la riduzione di un numero più o meno elevato di variabili (rappresentanti altrettante caratteristiche del fenomeno analizzato); in secondo luogo l'ANOVA (Cfr. sezione 2 in appendice), che permette di confrontare due o più gruppi di dati comparando la variabilità *interna* a questi gruppi con la variabilità *tra* i gruppi.

CONCLUSIONI

Dall'analisi suddetta si rileva una suddivisione delle strategie in 3 rispettivi gruppi (Cfr. grafico 1 in appendice): il cluster composto dalle strategie *cta/managed futures*, *long/short equities* e *macro* risulta il vincente tra i gruppi osservati, in coerenza anche con quanto rilevato dalla letteratura citata nei paragrafi precedenti.

Cluster 1

Fondo Hedge

CTA / Managed Futures	<i>Mulvaney Global Markets Fund Ltd</i>
Long / Short Equities	<i>CPIM Global Financials Fund LP Class AM-1</i>
Macro	<i>Danix Fund - Class M USD</i>

Cluster 2

Fondo Hedge

Arbitrage	<i>VCM Equity Market Neutral Segregated Portfolio</i>
Bottom-Up	<i>Polar Capital Funds Plc Japan Fund - GBP</i>
Distressed Debt	<i>Solent Distressed Credit Fund – EUR</i>
Diversified Debt	<i>Thames River High Income Fund (NOK)</i>
Dual Approach	<i>Parvest Japan Yen Bond</i>
Multi-Strategy	<i>GLC Diversified Fund</i>
Others	<i>CAL Global Macro Fund</i>
Relative Value	<i>Global Maritime Investments Limited - Class D1</i>
Top-Down	<i>Thames River Global Bond Fund (GBP) - EUR Dbn</i>

Cluster 3

Fondo Hedge

Event Driven	<i>Cumulus Energy Fund</i>
Fixed Income	<i>Stratton Street PPC Ltd Asian Bond Fund – USD</i>

In particolare, Malkiel e Saha (2005) avevano già rilevato come il gruppo più ampio di hedge fund fosse quello della categoria *long/short*, così come le statistiche annuali fornite dall'Ufficio Studi Alternative Investment avevano osservato che tra le singole strategie quelle che avevano conseguito le migliori performance annualizzate registrate nel periodo gennaio 1994 - aprile 2008 vi fossero la *global macro* (+13.80%) e la *long/short equity* (+11.74%); allo stesso modo dai dati di mercato risultava che già per i primi 6 mesi del 2008 i gestori *managed futures* avevano ottenuto un ottimo rialzo pari al 9.60% poiché avevano sfruttato la volatilità di materie prime e valute.

La strategia che nel periodo di riferimento osservato realizza la migliore performance è dunque la *CTA / managed futures* con il fondo Mulvaney Global Markets Fund Ltd che risulta il fondo con il punteggio più elevato in seno alla strategia; in secondo luogo rileviamo la strategia *long/short equities* con il fondo CPIM Global Financials Fund LP Class AM-1e, infine, la strategia *macro* con il fondo Danix Fund - Class M USD.

Il cluster composto, invece, dalle strategie *arbitrage, bottom-up, distressed debt, diversified debt, dual approach, multi-strategy, others, relative value, top-down* è il gruppo che presenta caratteristiche di sviluppo decisamente nella media e molto simili tra di loro in base alle variabili osservate.

Infine il cluster composto da *event driven* e *fixed income* - strategie che nell'ultimo ventennio hanno aumentato la percentuale di penetrazione nel mercato sfavorendo altre categorie come ad esempio la *macro* - raccoglie i fondi che hanno comunque avuto un ritorno positivo nel corso dell'ultimo anno ma si distinguono dagli altri per i valori estremamente negativi che riporta la variabile maximum draw down.

Passando all'analisi dei risultati, si evince chiaramente come la strategia *Cta/managed futures* sia risultata la vincente nel periodo preso a riferimento. Ciò è perfettamente coerente con le condizioni dei mercati finanziari nella piena bufera finanziaria e caratterizzati da enorme volatilità che ha permesso ai gestori di tale comparto di sfruttare a pieno la leva finanziaria (connessa ai futures) e puntare sull'enorme rialzo del prezzo delle principali commodities. L'alta volatilità dei trend dei mercati azionari ha contestualmente favorito la strategia *long/short equities*. Anche i fondi maggiormente capitalizzati - che hanno potuto utilizzare una strategia di tipo *Macro* - sono stati premiati dal mercato; in effetti occorre considerare che i gestori di tale comparto operano su tutte le classi di attività - titoli di Stato, azioni, materie prime, metalli preziosi, obbligazioni societarie - fanno uso di tutti gli strumenti di investimento - titoli, tassi di interesse, opzioni, futures, swap e altri strumenti derivati - e ricorrono alla vendita allo scoperto e alla leva finanziaria. Non sono previsti limiti geografici ai loro investimenti e pertanto la loro operatività si estende in ogni parte del mondo, dai mercati dei paesi sviluppati ai mercati emergenti.

Fixed income ed event driven sono infine risultate le strategie che hanno presentato statisticamente il più elevato livello di maximum draw down. Ciò è perfettamente coerente con le attuali condizioni di crisi finanziaria dove i titoli del reddito fisso (frutto prevalentemente di cartolarizzazioni e CDO) hanno realizzato grosse svalutazioni e dove le operazioni societarie legate a eventi di carattere straordinario sono state di numero minore e caratterizzate da un alto rischio.

Bibliografia

1. Ackermann C., McEnally R., Ravenscraft D.(1999), “The performance of hedge funds: risk, return, and incentives”, *Journal of Finance* 54 (3).
2. Ackermann C., Ravenscraft D.(1998), “The impact of regulatory restrictions on fund performance: A comparative study of hedge funds and mutual funds”, University of North Carolina Dissertation.
3. Agarwal V; Naik N.(2001), “Characterizing systematic risk of hedge funds with buy-and-hold and option-based strategies”, Working Paper, London Business School.
4. Agarwal V., Naik N.(2000), “Multi-period performance persistence analysis of hedge funds”, *Journal of Financial and Quantitative Analysis* 35 (3).
5. Agarwal V., Naik N.(2000), ”On taking the “alternative” route: the risks, rewards, and performance persistence of hedge funds”. *Journal of Alternative Investments* 2.
6. Agarwal V; Naik N.(1999), “On taking the 'Alternative' route: Risks, Rewards, Style and Performance Persistence of Hedge Funds”, IFA Working Paper.
7. Amin G. S., Kat H.(2003), *Stocks, Bonds and Hedge Funds: Not A Free Lunch!*, *Journal of Portfolio Management*.
8. Amin, G., Kat, H.(2003), “Hedge funds performance 1990–2000. Do the money machines really add value?”, *Journal of Financial and Quantitative Analysis* 38 (2).
9. Amin, G., Kat H.(2002), “Diversification and Yield Enhancement with Hedge Funds”, *Journal of Alternative Investments*.
10. Biorcio R. (1993), *L’analisi dei gruppi*, Franco Angeli
11. Burton G. Malkiel e Atanu Saha (2005), *Hedge funds: Risk and return*, *Financial analysts journal*.
12. Brooks C., Kat H.(2002), “The statistical properties of hedge fund index returns and their implications for investors”, *Journal of Alternative Investments* 5 (2).
13. Brown S.J., Goetzmann W.N., Ibbotson R.G.(1999), “Offshore hedge funds: survival and performance, 1989–95”, *The Journal of Business* 72.
14. Brown S.J., Goetzmann W.N., Park, J.(1997), “Conditions for survival: changing risk and the performance of hedge fund managers and CTAs”. Working Paper, Stern School of Business, New York University.

15. Brown S.J., Goetzmann W.N., (1997), "Mutual fund styles", *Journal of Financial Economics*, volume 43, pagg. 373-399.
16. Brown S.J., Goetzmann W.N., (2001), "Hedge funds with style", NBER Working Paper No. W8173, New York.
17. Calinski R.B., Harabasz J., (1974), A dendrite method for cluster analysis, *Communications in Statistics*, 3(1): 1-27.
18. Carhart M.M.(1997), "*On Persistence in Mutual Fund Performance*", *The Journal of Finance*, March.
19. Cowell, F.(2003), "Investment mandates for hedge funds", *Pensions: An Alternative Journal* 9 (2).
20. Cvitanić J., Lazrak A., Martellini L., Zapatero F.(2002), "Optimal Allocation to Hedge Funds : An Empirical Analysis", *Quantitative Finance*.
21. Davies R. J., Kat H. M., Lu S.(2004), "Fund of Hedge Funds Portfolio Selection: A Multiple-Objective Approach", AIRC Working Paper No. 18.
22. Fabris L, (1997), *Statistica multivariata. Analisi esplorativa dei dati*, Ed. McGraw-Hill Companies.
23. Fama E., MacBeth J. (1973), "Risk, Return and Equilibrium: Empirical Tests", *Journal of Political Economy*, Vol. 81, pp. 115-146
24. Favre L., Galeano J. A.(2002), Mean-Modified Value-at-Risk optimization with hedge funds, *Journal of Alternative Investments*.
25. Fung, W., Hsieh D.(2001), "The risk in hedge fund strategy: Theory and evidence from trend followers", *Review of Financial Studies* 14.
26. Fung, W., Hsieh D.(2000), "Performance Characteristics of Hedge Funds and Commodity Funds: Natural vs. Spurious Biases", *Journal of Financial and Quantitative Analysis* 35.
27. Fung, W., Hsieh, D.A.(1997), "Empirical characteristics of dynamic trading strategies: The case of hedge funds", *Review of Financial Studies* 10.
28. Ghiringhelli P.(2004), "Hedge Funds e gestione di portafogli mobiliari", EGEA.
29. Giorgino M., Viviani D. (2006), *Hedge funds e fondi comuni non armonizzati: come migliorare l'efficienza dei portafogli*, working paper, Politecnico di Milano.
30. Goetzmann W., Hiraki T., Otsuki T., Shiraishi N.(2001), The Japanese Open_End Fund Puzzle, *Journal of Business* 74.
31. Green P.E., Frank R.E., Robinson P.J. (1967), Cluster Analysis in text market selection, *Management science*.
32. Hoevenaars R., Molenaar R., Schotman P., Steenkamp T.(2005), "Strategic Asset Allocation with Liabilities: Beyond Stocks and Bonds", Department of Quantitative Economics Maastricht University.

33. Jardine N., Sibson R. (1971), *Mathematical taxonomy*, Ed. Wiley, Londra.
34. Koh F., Koh W.T.H., Teo M. (2003), *Asian Hedge Funds: Return Persistence, Style, and Fund Characteristics*, School of Business, Singapore Management University, working paper.
35. Lamm R.M. (2003), "Asymmetric returns and optimal hedge fund portfolios", *Journal of Alternative Investments*.
36. Lhabitant F.S. (2004), *Hedge Funds with Quantitative Insights*, John Wiley & Sons, Ltd.
37. Liang, B. (1999), "On the performance of hedge funds", *Financial Analysts Journal* 55 (4).
38. Lo A. (2001), "Risk management for hedge funds: introduction and overview". *Financial Analysts Journal* 57.
39. Malkiel B.G. e Saha A. (2005), *Hedge funds: Risk and return*, *Financial analysts journal*.
40. Markowitz H M. (1952), "Portfolio selection", *Journal of Finance* 7.
41. Morton D.P., Popova E., Popova I. (2006), "Efficient fund of hedge funds construction under downside risk measures", *Journal of Banking & Finance*, Vol. 30 (2).
42. Schneeweis, T., Spurgin, R. (1997), "Managed futures, hedge funds and mutual fund return estimation: A Multi-factor approach". CISDM Working Paper.

APPENDICE

Sezione 1

Analisi in Componenti Principali

Matrice di correlazione (Pearson(n)): l'indice di correlazione di Pearson tra due variabili aleatorie (di seguito esposto) è un coefficiente che esprime la linearità tra la loro covarianza e il prodotto delle rispettive deviazioni standard.

Variables	2008 Return	2007 Return	Ann. Std Dev	Sharpe Ratio	Max. Draw Down	Sortino Ratio	Fund Size (US\$m)
2008 Return	1	-0,618	0,781	0,339	-0,451	0,333	-0,166
2007 Return	-0,618	1	-0,601	-0,147	0,570	-0,138	0,137
Ann. Std Dev	0,781	-0,601	1	0,181	-0,498	0,179	-0,233
Sharpe Ratio	0,339	-0,147	0,181	1	-0,246	0,992	0,156
Max. Draw Down	-0,451	0,570	-0,498	-0,246	1	-0,206	0,261
Sortino Ratio	0,333	-0,138	0,179	0,992	-0,206	1	0,177
Fund Size (US\$m)	-0,166	0,137	-0,233	0,156	0,261	0,177	1

I valori in grassetto sono diversi da 0 al livello di significatività alfa=0,05

Test di sfericità di Bartlett: prova che la matrice di correlazione iniziale è una matrice identità. La matrice identità presenta il numero 1 sulla diagonale e 0 nelle posizioni extra- diagonali.

Chi-quadro corretto (Valore osservato)	63,988
Chi-quadro corretto (Valore critico)	32,671
GDL	21
p-value	< 0,0001
α	0,05

Interpretazione del testo:

H0: Non ci sono correlazioni significativamente diverse da 0 tra le variabili.

Ha: Almeno una delle correlazioni tra le variabili è significativamente diversa da 0.

Considerando che il p-value calcolato è inferiore al livello di significazione alfa=0,05, non è possibile accettare l'ipotesi nulla H0 ma bisogna accettare l'ipotesi alternativa Ha.

Il rischio di rifiutare l'ipotesi nulla H0 quando è vera è inferiore a 0,01%.

ACP avviene tramite una trasformazione lineare delle variabili che proietta quelle originarie in un nuovo sistema cartesiano nel quale la nuova variabile con la maggiore varianza viene proiettata sul primo asse, la variabile nuova (seconda per dimensione) della varianza sul secondo asse e così via. La riduzione della complessità avviene limitandosi ad analizzare le principali (per varianza) tra le nuove variabili. Diversamente da altre trasformazioni (lineari) di variabili, praticate nell'ambito della statistica, in questa tecnica sono gli stessi dati che determinano i vettori di trasformazione. Assumendo che a ciascuna delle variabili originarie venga sottratta la loro media la nuova variabile (X_i) ha media nulla,

$$\mathbf{w}_1 = \arg \max_{\|\mathbf{w}\|=1} E \left\{ (\mathbf{w}^T \mathbf{x})^2 \right\}$$

Con i primi (k-1) componenti, il k-esimo componente può essere trovato sottraendo i primi (k-1) componenti principali a X

$$\hat{\mathbf{x}}_{k-1} = \mathbf{x} - \sum_{i=1}^{k-1} \mathbf{w}_i \mathbf{w}_i^T \mathbf{x}$$

e sostituendo questo

$$\mathbf{w}_k = \arg \max_{\|\mathbf{w}\|=1} E \left\{ (\mathbf{w}^T \hat{\mathbf{x}}_{k-1})^2 \right\}.$$

Un metodo più semplice per calcolare la componente \mathbf{w}_i utilizza la matrice delle covarianze di \mathbf{x} . Trovando l'autovalore e l'autovettore della matrice di covarianza, si ottiene che l'autovettore con il maggiore autovalore corrisponde alla dimensione che ha la maggiore correlazione con l'insieme di dati. I dati originali sono infine proiettati nello spazio vettoriale ridotto.

Gli elementi dell'autovettore colonna corrispondente ad un autovalore esprimono il legame tra le variabili di partenza e la componente considerata attraverso dei pesi. Il numero di variabili latenti da considerare come componenti principali si fonda spesso su considerazioni soggettive (la grandezza relativa degli autovalori rispetto agli altri infatti, ci dice la percentuale di variazione causata da ogni variabile latente: da qua decideremo quali tenere).

Possiamo costruire quindi la **matrice dei fattori** (in pratica una matrice modale) che elenca per riga le variabili originarie, per colonna le variabili latenti: ogni valore ci dice quanto le seconde incidano sulle prime (con valori che vanno da 0 ad 1).

La **matrice del punteggio fattoriale**, invece, ha la stessa struttura della precedente ma, a differenza della stessa, ci dice quanto le singole variabili originarie abbiano pesato sulla determinazione della grandezza di quelle latenti.

Al fine di individuare il numero di componenti principali estraibili per ottenere la più efficace sintesi dei dati è stato utilizzato il criterio dell'autovalore maggiore di 1 (detto criterio di Kaiser).

Autovalori:

	F1	F2	F3	F4	F5	F6	F7
Autovalore	3,147	1,845	0,836	0,616	0,355	0,194	0,007
Volatilità (%)	44,953	26,354	11,942	8,798	5,073	2,776	0,103
% cumulata	44,953	71,307	83,249	92,047	97,120	99,897	100,000

Autovettori:

	F1	F2	F3	F4	F5	F6	F7
2008 Return	0,481	-0,109	0,185	0,435	0,065	0,727	-0,009
2007 Return	-0,422	0,239	-0,304	0,287	0,753	0,153	-0,011
Ann. Std Dev	0,453	-0,240	0,177	0,379	0,388	-0,641	0,020
Sharpe Ratio	0,329	0,576	-0,221	0,000	-0,065	-0,061	0,709
Max. Draw Down	-0,401	0,176	0,165	0,736	-0,464	-0,148	0,040
Sortino Ratio	0,321	0,586	-0,197	0,040	-0,091	-0,100	-0,703
Fund Size (US\$m)	-0,112	0,410	0,853	-0,204	0,226	0,020	0,008

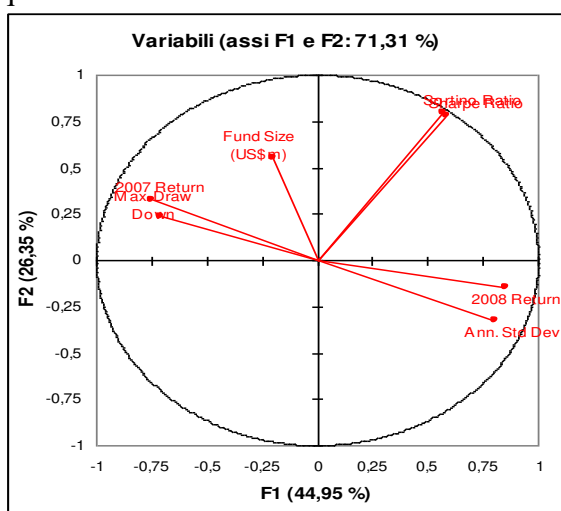
Coordinate delle variabili:

	F1	F2	F3
2008 Return	0,854	-0,147	0,169
2007 Return	-0,749	0,324	-0,278
Ann. Std Dev	0,804	-0,325	0,162
Sharpe Ratio	0,584	0,782	-0,202
Max. Draw Down	-0,711	0,239	0,151
Sortino Ratio	0,569	0,796	-0,180
Fund Size (US\$m)	-0,198	0,556	0,779

Rappresentazione grafica ACP

Se due variabili sono correlate positivamente, sono situate dalla stessa parte su di un asse.

Le variabili sono dei vettori e non dei punti. Non è la prossimità tra individuo e un insieme di punti rappresentante delle variabili che è importante ma l'allontanamento dell'individuo nella direzione di questo insieme di variabili.



Coordinate delle osservazioni:

Osservazioni	F1	F2	F3
Arbitrage	-0,729	0,779	-0,920
Bottom-Up	0,531	0,104	-0,001
CTA / Managed Futures	4,823	-1,670	1,168
Distressed Debt	-3,130	-3,692	0,615
Diversified Debt	-1,081	0,445	0,759
Dual Approach	-1,610	-0,688	-0,070
Event Driven	1,144	-0,093	-0,873
Fixed Income	-0,427	-0,182	-0,927
Long / Short Equities	0,577	0,310	-0,978
Macro	1,836	-0,412	-0,500
Multi-Strategy	-0,158	0,937	-0,071
Others	-0,295	0,509	-0,913
Relative Value	-0,830	1,841	0,582
Top-Down	-0,651	1,812	2,128

Sezione 2

ANOVA

L'ipotesi alla base dell'analisi della varianza è che, dati n gruppi, sia possibile scomporre la varianza in due componenti: varianza interna ai gruppi (anche detta Within) e varianza tra i gruppi (Between). La ragione che spinge a compiere tale distinzione è la convinzione, da parte del ricercatore, che determinati fenomeni trovino spiegazione in caratteristiche proprie del gruppo di appartenenza. Il confronto si basa sull'idea che se la variabilità interna ai gruppi è relativamente elevata rispetto alla variabilità tra i gruppi, allora probabilmente la differenza tra questi gruppi è soltanto il risultato della variabilità interna.

Questa analisi si basa sul confronto della varianza e usa variabili di test distribuite come la variabile casuale F di Snedecor.

Analisi della varianza (Variabile 2008 Return):

Fonte	GDL	Somma dei quadrati	Media dei quadrati	F	Pr > F
Modello	13	184597,930	14199,841	3,218	< 0,0001
Errore	774	3414914,562	4412,034		
Totale corretto	787	3599512,492			

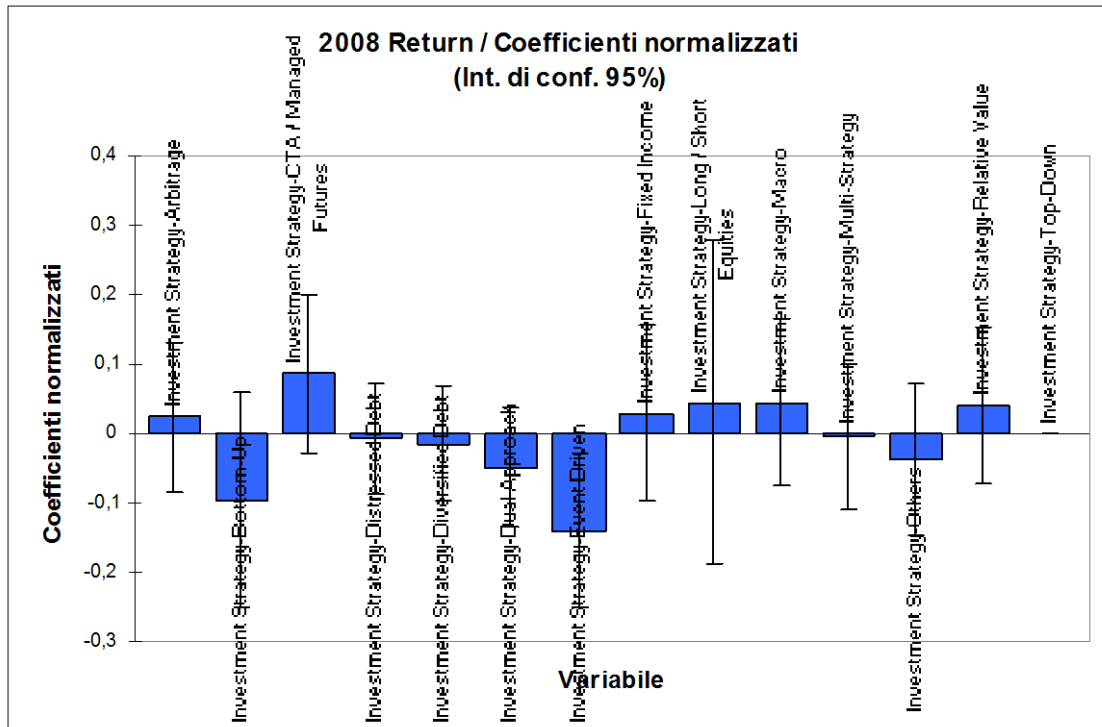
Calcolato contro il modello $Y=Media(Y)$

Parametri del modello (Variabile 2008 Return):

Fonte	Valore	Standard Deviation	T	Pr > t	Limite Inferiore (95%)	Limite Superiore (95%)
Intercetta	-20,413	15,656	-1,304	0,193	-51,147	10,320
Investment Strategy-Arbitrage	8,585	20,067	0,428	0,669	-30,808	47,977
Investment Strategy-Bottom-Up	-21,080	17,270	-1,221	0,223	-54,982	12,823
Investment Strategy-CTA / Managed Futures	28,217	19,266	1,465	0,143	-9,602	66,037
Investment Strategy-Distressed Debt	-5,364	29,587	-0,181	0,856	-63,445	52,717
Investment Strategy-Diversified Debt	-8,965	27,117	-0,331	0,741	-62,196	44,267
Investment Strategy-Dual Approach	-28,056	25,421	-1,104	0,270	-77,957	21,846
Investment Strategy-Event Driven	-49,174	19,804	-2,483	0,013	-88,049	-10,298
Investment Strategy-Fixed Income	8,484	18,467	0,459	0,646	-27,768	44,735
Investment Strategy-Long / Short Equities	5,962	16,010	0,372	0,710	-25,465	37,390
Investment Strategy-Macro	13,648	18,852	0,724	0,469	-23,360	50,656
Investment Strategy-Multi-Strategy	-1,607	20,533	-0,078	0,938	-41,913	38,700
Investment Strategy-Others	-13,526	19,804	-0,683	0,495	-52,401	25,350
Investment Strategy-Relative Value	14,018	19,570	0,716	0,474	-24,399	52,434
Investment Strategy-Top-Down	0,000	0,000				

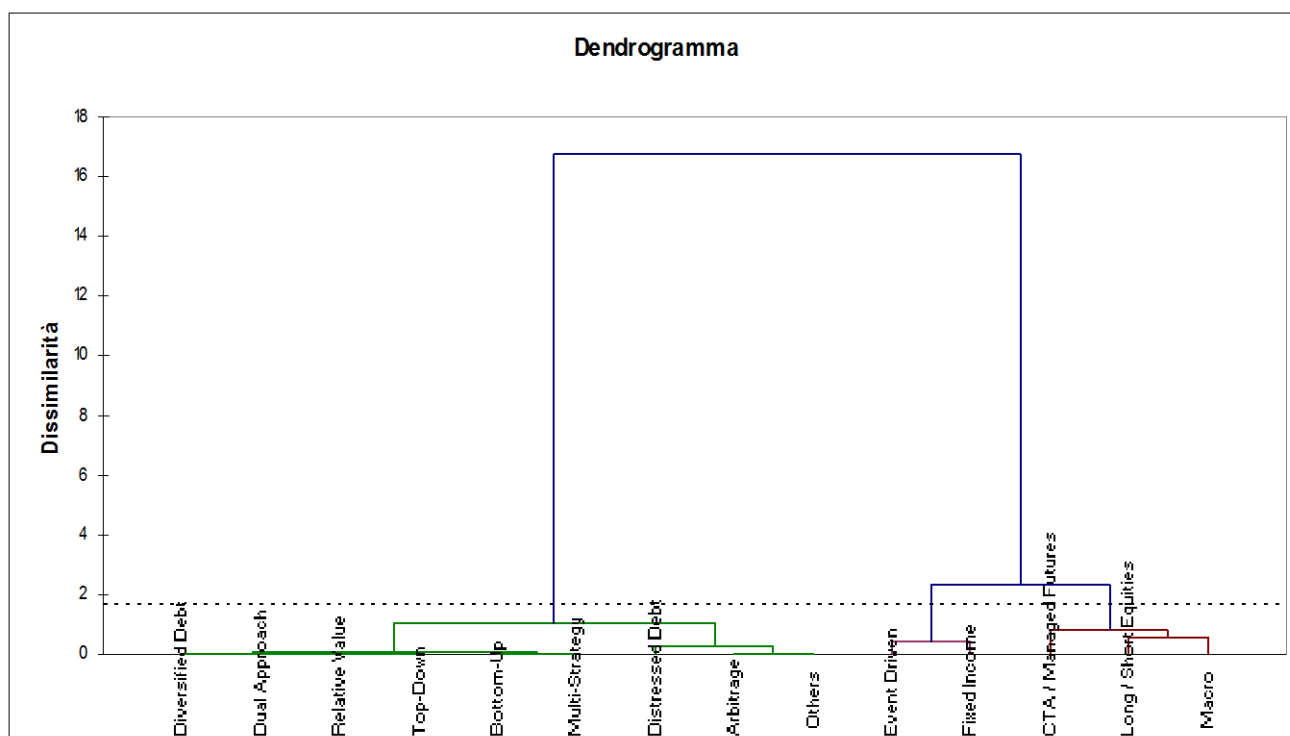
Equazione del modello (Variabile Return 2008):

2008 Return = -20,4132330246914+8,58473600088201*Investment Strategy-Arbitrage-21,0797070690167*Investment Strategy-Bottom-Up+28,217419532628*Investment Strategy-CTA / Managed Futures-5,36390983245138*Investment Strategy-Distressed Debt-8,9645447530863*Investment Strategy-Diversified Debt-28,0555169753085*Investment Strategy-Dual Approach-49,1735262345678*Investment Strategy-Event Driven+8,48359836285574*Investment Strategy-Fixed Income+5,96246138058542*Investment Strategy-Long / Short Equities+13,6483094135803*Investment Strategy-Macro-1,60688364197522*Investment Strategy-Multi-Strategy-13,5255354938271*Investment Strategy-Others+14,0175472608025*Investment Strategy-Relative Value



L'utilità della rappresentazione grafica dell'ANOVA attraverso il box-plot deriva dal fatto che una sola rappresentazione grafica si evidenziano gli aspetti principali di una distribuzione di frequenza. Il primo e terzo quartile costituiscono gli estremi della scatola del box plot la cui lunghezza è quindi rappresentata dalla differenza interquartile. La linea orizzontale all'interno della scatola indica la mediana, per cui a seconda di quanto questa sarà posizionata centralmente si individueranno distribuzioni più o meno simmetriche. Le due linee esterne fuori dalla scatola sono anche dette baffi e sono delimitati, rispettivamente, dai valori minimo e massimo della distribuzione. Un baffo eccessivamente lungo indica la presenza di valori anomali.

Grafico 1: Rappresentazione grafica della struttura di raggruppamento tramite diagramma ad albero (o dendrogramma)



La distanza di fusione può essere facilmente evinta dall'osservazione del dendrogramma: se nel passaggio da K gruppi a $K+1$ si registra un forte incremento della distanza di fusione si deve "tagliare" a K gruppi. Il taglio in questa analisi si verifica in corrispondenza di un livello di dissimilarità pari a circa 2. Ciò consente di individuare 3 cluster, 2 dei quali, come si nota dal grafico, hanno un livello di similarità elevato:

1. CTA / Managed Futures, Long / Short Equities, Macro
2. Event Driven, Fixed Income
3. Arbitrage, Bottom-Up, Distressed Debt, Diversified Debt, Dual Approach, Multi-Strategy, Others, Relative Value, Top-Down

Si riportano di seguito le elaborazioni dettagliate del campione osservato, suddivise per cluster e per oggetto.

Risultati per classe:

Classe	1	2	3
Oggetti	9	3	2
Somma dei pesi	9	3	2
Varianza intraclassa	96400,772	3765,575	695,577
Distanza minima al baricentro	26,240	34,653	18,649
Distanza media al baricentro	248,751	47,752	18,649
Distanza massima al baricentro	600,936	69,014	18,649
		CTA / Managed	
	Arbitrage	Futures	Event Driven
		Long / Short	
	Bottom-Up	Equities	Fixed Income
	Distressed		
	Debt	Macro	
	Diversified		
	Debt		
	Dual Approach		
	Multi-Strategy		
	Others		
	Relative Value		
	Top-Down		