GRETA

WORKING PAPER n. 00.06

September 2000

Performances and Performance Persistence of Italian Equity Founds

Roberto Casarin[1]
Loriana Pelizzon[2]
Andrea Piva[1]

[1] GRETA Associati, Venice
[2] University of Padova

Introduction

Fifteen years after the development of the italian funds market, a detailed analysis of fund's performances can supply interesting questions about the way the market works in this firstly phase of development. In addition, the implications of these empirical researches does not affect only this sector of study, but also the entire capital market with indications on its efficiency.

This work evaluates the presence of "superior abilities", considered both in terms of capability to generate "extra-returns" and ability to beat constantly the competitors.

In particular, referring to this last topic (performance persistence) it is not yet evaluated on the italian market.

In addition, the persistence analysis developed in this paper captures relation with other variables, such as the lag of evaluation, the performance measure, the number of performing class and the statistical test.

This empirical research refers to two main topics:

- Evaluation of the ability to generate extra-performances (in terms of "market timing" and "stock picking" ability);
- Analysis of performance persistence on different lags, with different measure of performance and with several statistic techniques.

Referring to the first objective, we appreciated the statistic significance of the Jensen's indicator, of the market timing coefficients (estimated with both the quadratic equation of Treynor and Mazuy and with that of Henriksson-Merton) and of the related "total performance" measures, determined in a six factors Arbitrage Pricing Theory contest. In particular, we adopted the Chen, Roll and Ross's model (1986) plus the "market factor".

Referring to the second issue, we studied the "hot-hand effect" and the "long-run persistence"; we appreciated the phenomenon with Jensen's indicator, with total performance coefficients and with the measures developed by Sharpe, Treynor and Sortino (in addition to row returns) and we evaluated the persistence using stability class of different length (e.g. divisions in octiles or in two macro-classes). In addition, as in Agarwal and Naik (2000), we applied a multiperiod persistence test. By this technique, we performed the analysis on windows of three or for subperiod (not only of two periods, as in "traditional" test of persistence).

The “theoretical framework” of the analysis is the study of Cesari and Panetta (1998) (for the performance evaluation part) and the entire set of studies produced in the last ten years, with particular regard to the contributes of Hendriks, Patel and Zeckhauser (1993) and Agarwal and Naik (2000) (for the persistence analysis).

The organisation of the paper is as follows: in section one we describe the sample of valuation; in section two we estimate risk adjusted performance measure; in last section we describe the methodology used to perform the persistence tests and we evidence the results.

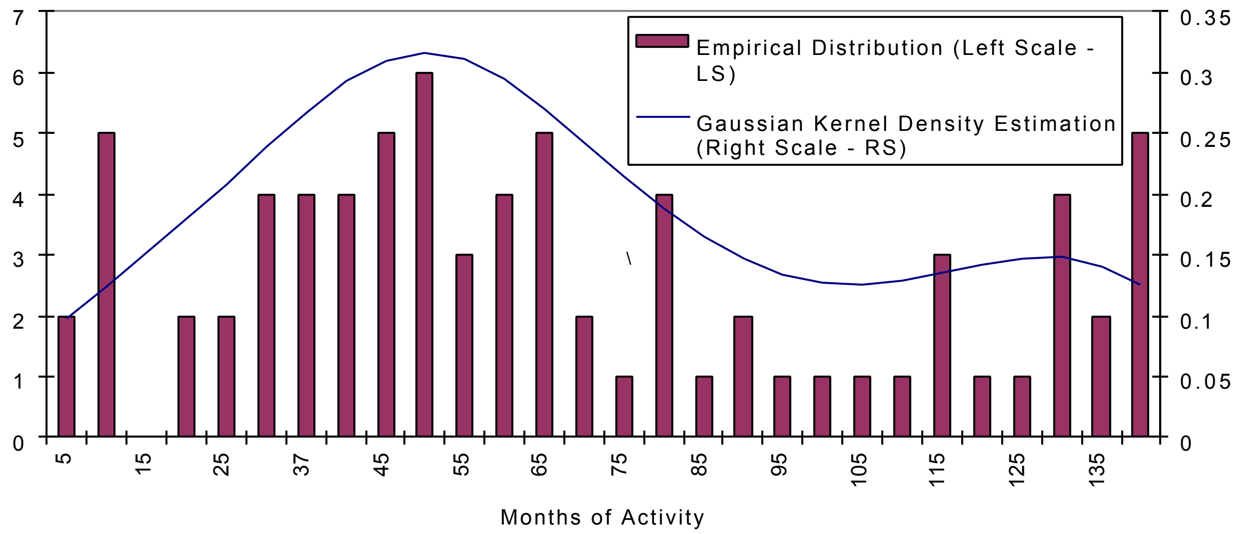
1. The sample

We investigate the presence of “superior abilities” on a sample of funds classified as “Azionari Italia” by Assogestioni[1] at August the 31st 1999, which invest 84% of the managed funds on italian stocks (some descriptive statistics are reported in tables AI and AII of the appendix).

The data set contains monthly quotations of all the selected funds, since their constitution and through changes (eventual) in their name. It is composed by “journal quotation”: data correspond to those commonly published on the press and they are extracted from the archive of “Il Sole 24 Ore”.

The sample is determined combining the above-mentioned criteria with another constrain: mutual funds must have become active on the market at least at July the 31st 1996. This new restriction has been introduced because, to perform our analysis, we need at least 36 data (36 months of returns or, in other words, 37 months of quotations). In this way, the initial sample has been reduced from 76 to 57 elements (Chart 1.1 reports fund’s distribution for months of activity).

Chart 1.1 – Fund’s distribution for month of activity



[1] A mutual fund is classified by Assogestioni (the italian mutual funds association) as “Azionario Italia” if the fraction invested on italian stocks is at least 70%.

The sample does not present attrition rate, but it is affected, even though in a very limited way, by survivorship bias. In fact, even if the data set captures changes in the names of mutual funds, the selection procedure does not include funds that changed investment politic, or that retired from activity, etc. However, it seems rationale to say that these operations affect the italian market only marginally: these are typical operations of a full market, which presents high levels of competition (one of the main purposes of these operations is to hide bad performances from view of the public).

2. Performance measures and empirical results

Mutual fund's performance can not be evaluated without considering the level of risk. However, if researchers agree on the necessity to associate a risk measure to returns, they do not agree on the most suitable risk indicator. Sometimes we utilise the standard deviation of past returns (as in the Sharpe's indicator); other times we employ a volatility measure which considers only the left side of the distribution (downward volatility, half variance, etc., as in the Sortino's measure); other times again we consider the covariance of mutual fund returns with those of a market indicator (measures of Treynor and Jensen).

In this work we applied several performance measures, in order to appreciate the phenomenon from different perspectives and to point out relations among them. Given N mutual funds, for the p-th fund we implemented the following measures.

2.1 Sharpe's indicator

$$S_p = \frac{\left(\bar{r}_p - \bar{r}_f\right)}{s_P} \tag{2.1}$$

This measure, also called "reward to variability ratio", expresses the trade off between the excess return on the portfolio and its standard deviation (which represents an indicator of *total risk*). Hence, Sharpe's indicator expresses the *Capital Asset Line*'s slope.

2.2 Traynor's indicator

$$T_p = \frac{\left(\bar{r}_p - \bar{r}_f\right)}{b_P} \tag{2.2}$$

Like Sharpe's measure, it evaluates fund's performance by dividing its excess return for the level of risk, but in this case, returns are corrected for a *systematic risk* measure.

2.3 Sortino's indicator

$$So_p = \frac{\bar{r}_P - \bar{r}_f}{DD_p} \tag{2.3}$$

It expresses the trade-off between the excess return on the fund and a downside risk measure defined as:

$$DD_p = \sqrt{VAR\left[\min\left(0, \bar{r}_p - \bar{r}_f\right)\right]} \tag{2.4}$$

The risk free rate represents the comparison term, but it could be set to other values such as the average or median of returns.

In this work we computed all the returns in the logarithmic way and we approximated the risk-free rate, r_f, with the *Buoni Ordinari del Tesoro*'s return (BOT).

2.4 Jensen's indicator

$$a = \left(\bar{r}_p - \bar{r}_f\right) - b_{pM}\left(\bar{r}_M - \bar{r}_f\right) \tag{2.5}$$

This indicator (also called a -coefficient) is equal to the difference between the excess return (risk premium) on the fund and the theoretical excess return (expressed by the CAPM) which should have been earned by the portfolio, given its risk (b_{pM}).

The measure developed by Jensen (1968, 1969) requires some other specification. In fact, it can be interpreted as the intercept of the following regression:

$$r_{pt} - r_{ft} = a_p + b_{pM}\left(r_{Mt} - r_{ft}\right) + e_{pt}, \quad \text{where } E\left(e_{pt}\right) = 0, V\left(e_{pt}\right) = s^2. \tag{2.6}$$

Interpreting the measure as the intercept of an OLS regression, enables us to perform statistical tests on the coefficient estimated in order to appreciate its significance. In general, a positive coefficient indicates superior abilities of the manager, whereas negative alphas indicate insufficient performance (to compensate the risk assumed).

To calculate a-Jensen, the choice of the benchmark is a fundamental matter because it can affect the performance measure[2]. For this reason it is extremely important to choose efficient benchmarks (in the mean-variance way).

In general, the problem does not change if we assess that two or more risk factors affect simultaneously fund's returns. Indeed, a multifactor model was implemented in this work: we applied the following variant of the Chen, Roll and Ross's model (1986)[3]:

$$r_{pt} - r_{ft} = a_p + \sum_{k=1}^{6} b_{pk}\left(r_{kt} - r_{ft}\right) + e_{pt}, \quad \text{where } E\left(e_{pt}\right) = 0, V\left(e_{pt}\right) = s^2 \tag{2.7}$$

Where the risk factors, r_k, are the following:

1. monthly variations on annual series of industrial production;

[2] For example Lehman and Modest (1987) and Grimblatt and Tietman (1994) studied the effect of different benchmarks on performance evaluation.

[3] Factors 1-5 are derived from the study of Chen, Roll and Ross (1986).

2. unexpected inflation, defined as the difference between the expect (at the end of the precedent period) and the effective (ex-post) rate of inflation;
3. variation on the expect inflation rate;
4. risk premium, defined as the difference between the return of a "BAA" rated bond fund and a long run government bond;
5. changes on term structure's slope, approximated by the difference between returns of bonds with different duration;
6. market portfolio's returns;

We approximated the above mentioned factors as it follows:

1. for the industrial production none approximation was required;
2. for the unexpected inflation rate, we approximated the expect inflation rate with the variation on M2-money aggregate and we measured the ex-post inflation rate with the variation on the consumer price index (cpi);
3. once we have determined the expect inflation rate (as described above) its variation is easily calculated;
4. the risk premium is determined by subtracting from a bond index's return (JP Morgan Bond Italy) the ten-year BTP's yield ("Buoni Poliennali del Tesoro");
5. we captured changes on term-structure's yield by making the difference between the ten-year BTP and the BOT;
6. we approximated market index with the Comit Globale index[4].

2.5 Market timing

As demonstrated by Grinblatt and Tietman (1989), α-Jensen's measure is not distorted if the fund manager is not a market timer. To consider the effect of timing ability, we applied the two models developed in literature.

Treynor and Mazuy (1966) noted that, if fund manager possesses timing ability, the portfolio characteristic line could not be straightforward. The way by which they captured the non-linearity of the Security Market Line, is the following:

$$r_{pt} - r_{ft} = \alpha_p + \beta_p \left(r_{Mt} - r_{ft}\right) + \gamma_p \left(r_{Mt} - r_{ft}\right)^2 + \varepsilon_{pt}, \quad \text{where } E\left(\varepsilon_{pt}\right) = 0, V\left(\varepsilon_{pt}\right) = \sigma^2 \qquad (2.8)$$

where α, β and γ are the regression's coefficients[5]. If the estimated γ is significantly positive there is timing ability.

[4] Even if we utilised this index, it is well known that value weighted italian portfolios are not efficient (Ferretti and Murgia (1991) demonstrated that the *Comit Globale* and also the *Mib Storico* indexes are inefficient)

[5] The formula refers to "the one factor model", but the question does not change if we consider a multifactorial contest (APT). In addition, it is important to remark that, in multiple regression, timing ability may be evaluated for every risk factor; however, in practical terms, it is evaluated only for the market factor.

Merton (1981) and Henriksson and Merton (1981) developed an alternative approach. Given the definition of market timing and considered that a perfect timer manager will choose the risk level b_{p0} if $r_{Mt} \leq r_{ft}$ and $b_p (> b_{p0})$ if $r_{Mt} > r_{ft}$, the beta of this manager is as follows:

$$b_{pM}(t) = b_{p0} + (b_{pi} - b_{p0})D_M \equiv b_{p0} + (b_{pi} - b_{p0})\left[\frac{\max\left(0, r_{Mt} - r_{ft}\right)}{r_{Mt} - r_{ft}}\right] \tag{2.9}$$

with $D_M = \begin{cases} 1 \text{ se } \mathrm{r_{Mt}} > r_{ft} \\ 0 \text{ se } \mathrm{r_{Mt}} \leq r_{ft} \end{cases}$. Observing that $max(0,x) = x + max(0,-x)$, we obtain:

$$b_p + (b_p - b_{p0})\frac{\max\left[0 - \left(r_{Mt} - r_{ft}\right)\right]}{r_{Mt} - r_{ft}}. \tag{2.10}$$

So, the Henriksson and Merton's model is the following:

$$r_{Mt} - r_{ft} = a_p + b_p\left(r_{Mt} - r_{ft}\right) + g_{pi}\max\left[0, -\left(r_{Mt} - r_{ft}\right)\right] + e_{pt}, \text{ where } E(e_{pt}) = 0, V(e_{pt}) = s^2 \tag{2.11}$$

If the estimated g is significantly positive there is timing ability.

It is interesting to note that Henriksson and Merton interpret the timing ability as a put option on market portfolio with strike price set to the risk-free rate. For this reason, the return deriving from timing ability, $max\left[0, -\left(r_{Mt} - r_{ft}\right)\right]$, is the option's payoff. In particular, funds return is equal to the sum of the standard one factor model plus g_p put options on market portfolio. So, if market-timing activity can be interpreted as a put option, it is also possible to determine its theoretic value.

In addition, following Grinblatt and Tietman (1994), we estimated total performance measures by adding the average return of market timing activity to the selectivity parameter α:

$$p_p = a_p + g_{pi}Var\left(r_{Mt} - r_{ft}\right); \tag{2.12}$$

In order to test the significance of total performance measures, the standard error has been calculated using the procedure suggested by Grinblatt and Tietman (1994)[6]. In fact, for p_p, the standard error was the following:

$$s_p = \sqrt{q' V q} \tag{2.13}$$

[6] Grinblatt and Tietman developed a methodology to estimate the significance of total performance measure only for the Treynor and Mazuy approach (see pag 440, B appendix of the mentioned paper). However, following Cesari and Panetta (1998) we applied it also to the Henriksson and Merton approach.

where $q' = \left(1 \quad 0 \quad Var\left(r_{Mt} - r_{ft}\right)\right)$ and V is the variance covariance matrix of regressors. Subsequently, the t-test for p_p was calculated observing that p / s_p follows a t-Student function with N-k-1 degree of freedom (N is the number of observations, k is the number of factors).

2.6 Empirical results

Table 2.1 reports the fund orders for every adopted measure.

We determined Jensen's indicator and timing coefficients (in addition to various β_k) by a "Recoursive Least Squares" (RLS) approach. This methodology, instead of an "Ordinary Least Squares" (OLS), allow us to test coefficient's stability and convergence (CUSUM test).

The residual's heteroscedasticity is a well-known problem in literature[7], for this reason we adopted the White's correction (1980) to determine t-statistics.

We performed numerous tests to evaluate the fitting of the model and its stability during time. Even if the model describes quite well the phenomenon (the average $\overline{R^2}$ coefficient is 0.94, with a minimum of 0.74), its parameters are rarely significant and only the Comit index ("market factor") is always different from zero at 99% probability.

Considering residuals, both the test on the AR(1) term and the Lagrange Multiplier test at one lag reject the presence of autocorrelation. Instead, the hypothesis of normality (Jarque-Bera's test) and omoscedasticity (White's test, ARCH(1) test and ARCH(4) test) could not be accepted: 41% times the normality is rejected at 95% level of probability; 32% times the omoscedasticity is rejected at the same level of probability.

The CUSUM test's results confirm the stability of the model and the convergence of parameters toward the determined value.

[7] See, for example, Cesari and Panetta (1998).

Table 2.1 – Mutual fund's performances

Fund	p T-M	p H-M	a -Jensen	Sharpe	Treynor	Sortino
Fund 2	48	48	49	29	21	20
Fund 3	26	14	27	9	4	2
Fund 4	14	17	22	51	52	51
Fund 5	29	39	13	41	39	40
Fund 7	33	46	36	25	28	26
Fund 9	43	40	54	44	50	49
Fund 10	53	44	53	15	13	14
Fund 11	11	15	18	54	54	54
Fund 12	51	53	41	50	51	52
Fund 13	45	52	56	18	24	22
Fund 14	37	41	21	3	6	5
Fund 15	58	56	47	12	12	12
Fund 16	28	31	51	56	56	56
Fund 19	16	11	19	6	10	9
Fund 20	36	26	45	32	34	34
Fund 22	32	6	37	5	7	6
Fund 23	54	55	50	16	25	24
Fund 25	24	12	42	14	14	13
Fund 26	27	18	35	13	15	15
Fund 27	38	35	25	58	58	58
Fund 28	44	49	7	2	5	7
Fund 29	3	4	5	19	23	23
Fund 30	5	3	3	21	17	17
Fund 31	17	29	4	1	1	1
Fund 32	20	28	32	35	35	35
Fund 33	22	25	26	43	41	41
Fund 34	30	32	39	20	18	16
Fund 35	50	45	44	10	11	11
Fund 36	6	5	20	33	26	30
Fund 37	47	50	34	49	48	48
Fund 39	31	20	38	28	30	25
Fund 41	49	43	58	39	45	45
Fund 42	34	33	24	57	57	57
Fund 45	9	7	9	23	19	19
Fund 46	55	36	46	31	22	27
Fund 48	40	38	40	37	38	38
Fund 52	41	47	43	26	27	29
Fund 53	12	10	30	40	42	42
Fund 56	7	8	11	42	40	39
Fund 57	21	9	16	7	9	8
Fund 58	39	54	17	27	32	32
Fund 59	46	42	48	34	36	36
Fund 60	18	22	29	46	43	43
Fund 61	10	16	14	45	44	44
Fund 62	8	13	15	38	37	37
Fund 63	42	34	52	22	20	18
Fund 65	52	51	55	30	29	28
Fund 66	19	21	10	24	31	31
Fund 67	35	37	23	52	49	50
Fund 68	1	1	2	11	3	4
Fund 69	57	57	57	17	16	21

The table reports the funds ranking for every adopted measure.

Table 2.1(follows) – Mutual fund's performances

Fund	p T-M	p H-M	a -Jensen	Sharpe	Treynor	Sortino
Fund 70	25	30	28	36	33	33
Fund 71	13	23	8	53	53	53
Fund 73	4	24	1	4	2	3
Fund 74	23	27	31	55	55	55
Fund 75	56	58	33	8	8	10
Fund 76	2	2	6	47	47	47
EQW	15	19	12	48	46	46

The table reports the funds ranking for every adopted measure.
"EQW" represents an "equally weighted" of all the fund of the sample.

Table 2.2 shows the results for the fund's alpha estimated with equation 2.7. It points out that fund managers were not able to generate extra-performances during the considered period. In fact, only six funds (which represent 10% of the sample) realised positive and significant alphas.

Table 2.2 – Mutual funds performances: *a* -Jensen

	a -J positive	a -J negative
Significant	6	0
Non significant	40	12

One of the portfolios which realised "superior performances" is the "Equally Weighted" (indicated as EQW in table 2.1)[8].

These results seem to be equal to those of Cesari and Panetta (1998) because, in this study as in that paper, net returns are not significantly different from zero.

In table 2.3 we reports market timing coefficients determined with equations 2.8 and 2.11. As for the *a* -Jensen, timing coefficients are not statistically different from zero, denoting that managers can not predict macro-movements on market.

Table 2.3 –"Market timing" ability

Quadratic Model (T-M)	*g positive*	*g negative*
Significant	5	0
Non significant	33	20
Henriksson-Merton's Model (put)		
Significant	2	0
Non significant	30	26

[8] Probably, this is attributable to the few funds which realised superior performances: they significantly incremented the average.

In addition, when we included a timing factor, the selectivity parameter (α-Jensen) has been reduced of 0.17% per month and of 0.14% per month (on average), respectively, for the quadratic equation and for that of Henriksson and Merton. Table 2.4 reports the results for the stock-picking parameter in the two equations:

Table 2.4 –"Stock selection" ability

Quadratic Model (T-M)	α *positive*	α *negative*
Significant	1	0
Non significant	36	21
Henriksson-Merton's Model (put)		
Significant	1	0
Non significant	34	23

Finally, following Grinblatt and Tietman's procedure, we estimated total performance measures. The results confirm the absence of superior abilities between fund managers: considering both equations 2.8 and 2.11, the estimated π are never significant (see tab. 2.5).

Table 2.5 – "Total performance" indicator

Quadratic Model (T-M)	$\pi = \alpha + \gamma$ *positive*	$\pi = \alpha + \gamma$ *negative*
Significant	0	1
Non significant	37	20
Henriksson-Merton's Model (put)		
Significant	0	0
Non significant	35	23

Tables 2.2 - 2.5 provide results in favour of market efficiency by asserting that fund managers can not earn "superior returns". In other words, empirical results suggest that a passive management of the benchmark produce higher performances than those realised by fund managers, in line with numerous studies produced on italian market[9] (see for example Cesari and Panetta (1998), Panetta and Zautzik (1991) and Ferretti and Murgia (1991)).

Nevertheless, this could represent a crucial point for the analysis developed here: if active managed funds cannot beat the market, why should investors put their money on this kind of investment that is also more expensive? The question is related to the embarrassing alternative that involves every empirical research on mutual fund's return: if market is efficient, then investors are

[9] Following Gruber (1996), we use indifferently the expressions "market return" and "passive fund return" (also called "index fund"). In fact, the two expressions are not too different because "index funds" have the objective to replicate a benchmark portfolio (ex.: MIB30, Mibtel, Standard and Poor's, etc.).

irrational while they buy active managed funds quote; on the other hand, if active managed funds realise superior performance, then market is not efficient.

Some studious, as Cesari and Panetta (1998), solve the diatribe by asserting that funds presents extra-performances when gross returns are compounded, but, when net returns are considered, fund managers are unable to beat the market (models of Grossman and Stiglitz (1980) and Cornell and Roll (1981)). Other studious, as Gruber (1996), formulate a different thesis, based on performance persistence.

In the opinion of Gruber, the incapacity of fund manager to beat market does not necessarily imply his total incapacity (he could always be the "most skilled" of all the competitors). So, if abilities exist, funds return will be foreseeable (at least in relative terms) and, if this is true, rational investors can realise superior performance by purchasing top performer funds and selling bottom performers. This is the crucial point of this work and we will dedicate it third section.

2.6.1 Relations between performance measures

Referring to performance evaluation, an interesting topic is to appreciate correlation among different measures. In other words, it is extremely interesting to verify if various performance measures show the same evaluation capacity.

It is well known that several performance measures present different characteristics (in terms of risk measure, in terms of their intrinsic meaning, for the possibility to appreciate their statistic significance, etc.), however it is particularly interesting to see if they produce analogous rankings of funds.

Table 2.6, which reports correlation among performance measure, indicates, in general, a good relation between regression based indicators (α-J and total performance). It also suggests high correlation between the coefficients of Sharpe, Traynor and Sortino, while it indicates little correlation between them and regression based measures.

Table 2.7, which reports rank order correlation[10] between different measures, seems to indicate a weak relation between their rankings.

Table 2.6 – Correlation between various performance measures

	p di T-M	p di H-M	a-Jensen	Sharpe	Treynor	Sortino
p di T-M	1	0,940542	0,644155	-0,1353	-0,1116	-0,0919
p di H-M		1	0,512132	-0,1041	-0,0608	-0,0314
a-Jensen			1	0,292	0,31275	0,2743
Sharpe				1	0,95849	0,9492
Treynor					1	0,9939
Sortino						1

[10] We appreciated rank order correlation by Spearman's coefficient (see below, footnote 25).

Table 2.7 – "Rank Order Correlation" between various performance measures

	p **di T-M**	p **di H-M**	a **-Jensen**	**Sharpe**	**Treynor**	**Sortino**
p **di T-M**	1	0,047511	0,120236	0,1186	-0,0501	-0,1921
p **di H-M**		1	0,041418	-0,2541	0,15738	-0,0622
a **-Jensen**			1	0,1434	0,05937	0,006
Sharpe				1	0,13171	0,0839
Treynor					1	0,3641
Sortino						1

3. Performance persistence[11]

3.1 The estimation method and statistical tests

We evaluated the level of persistence on italian equity funds by adopting several criteria of analysis.

The first dimension is *return*. We appreciated persistence on raw and risk-adjusted returns (with performance measures described above, at the precedent section).

Second, we evaluated the effect of *temporal lag* on persistency's level. While we investigated only for short run persistence on raw returns (on four months and annual intervals), on risk-adjusted performance we evaluated also long run persistence. More precisely, if we estimated the hot hand phenomenon on four months and annual lags, long run persistence has been evaluated on interval of 5.5 and of 2.5 years. These periods have been determined in this way: in the first case we considered performances on two periods of the same length (of the entire sample period); in the second case we evaluated the phenomenon on last five years of sample with two intervals of the same length.

Finally, we appreciated performance persistence with different *statistical tests*. Some of these (such as "*cross product ratio*" and *independence* c^2) have been performed on two macro-classes ("winners" and "losers"[12]); some other have been performed on group (octiles) of funds (such as *Spearman's rank order correlation* and *transition matrixes*); some other have been developed on single funds (*cross-sectional regressions*).

Considering single dimensions of analysis, we have:

- for *returns* and for *temporal lags*, we can say the followings:

1. raw returns, determined on monthly or annual base, are continuously compounded;

---

[11] In spite of its important implications, performance persistence is a phenomenon that has been studied only in last decade (unique exception is Jensen (1968)). In fact, principal contributes on subject are referred to Grinblatt and Titman (1992), Hendriks, Patel and Zeckhauser (1993), Goetzmann and Ibbotson (1994), Malkiel (1995), Brown and Goetzmann (1995), Carhart (1995), Elton and others. (1996), Brown, Goetzmann and Ibboltson (1996), Gruber (1996), Wermers (1997), Kent and others (1997) and Agarwal and Naik (2000).

[12] We defined a fund as "winner" if it excesses median return or, in other case, if it excesses the performance of the 75th percentile.

2. for risk-adjusted measures we have to discern between each single indicator. We estimated Jensen's indicator on the six factors Arbitrage Pricing Model (see above, second section) with regressions of 36-months[13]. While this constraint did not produce any problem for long-run analysis, for the hot hand phenomenon it caused evaluation windows overlapping. This effect could generate spurious persistence, therefore, as we will illustrate below, we tried to implement a model that overweighs the last observations, which are not overlapped. The same observations are valid for total performance indicators. For the measures of Sharpe, Traynor and Sortino, when we computed four months or annual returns, the risk indicators have always been calculated on 36-months windows[14]. This decision, which is assumed to produce consistent estimations, does not present the same problems encountered for Jensen's measure and for total performance coefficients. In fact, the numerator of these indicators (the excess return) is calculated on non-overlapped intervals.

- referring to *statistical tests*, we can assert the following:

1. *Cross product ratio test and Chi-squared independence test*. Schematically, we worked as it follows: at every date (for example at the end of every quarter) we considered all the funds that were active at the end of the precedent interval[15] and we computed the contingency table of WW ("winner" in both periods), LL ("loser" in both periods), WL ("winner" in the first period, "loser" in the second) and LW ("loser" in the first period, "winner" in the second). So, having the number of funds for every class, we computed the *cross-product ratio* as it follows:

$$CPR = \frac{WW * LL}{WL * LW} \tag{2.14}$$

This coefficient captures the fraction of funds that manifest persistence in spite of those that do not: the null hypothesis of no-persistence corresponds to a CPR close to one. In other words, the null hypothesis corresponds to four classes of the same number of funds (if we define "winner" a fund that excesses the median return). Then, we appreciated the statistical significance of this coefficient by considering the standard deviation of its natural logarithm (see Christensen (1990)):

$$s_{\ln(CPR)} = \sqrt{\frac{1}{WW} + \frac{1}{LL} + \frac{1}{WL} + \frac{1}{LW}}\,. \tag{2.15}$$

[13] As mentioned above, in the first section, we considered this interval as the minimum length sufficient to have good estimations.

[14] When we utilised these indicators to evaluate long run persistence, we estimated their risk measure considering the same lag (which is generally longer than 36 months) utilised to determine the excess return at the numerator.

[15] We remark that, in our sample, funds did not become active at the same data.

In fact, it is possible to demonstrate that the statistic $Z = \frac{\ln(CPR)}{s_{\ln(CPR)}}$ is normally distributed. In addition, we implemented a *test (Chi-squared)* to verify the independence of the distributions WW, WL, LL and LW. We performed this test at the end of the evaluation period (by summing between various dates) and calculating the c^2 statistic as it follows:

$$c^2 = \frac{(WW - D1)^2}{N} + \frac{(WL - D2)^2}{N} + \frac{(LW - D3)^2}{N} + \frac{(LL - D4)^2}{N} \tag{2.15}$$

where: $D1 = \frac{(WW + WL)*(WW + LW)}{N}$; $D2 = \frac{(WW + WL)*(WL + LL)}{N}$;

$D3 = \frac{(LW + LL)*(WW + LW)}{N}$; $D4 = \frac{(LW + LL)*(WL + LL)}{N}$.

This statistic, defined as Pearson's statistic, follows a χ^2 distribution with one degree of freedom[16]. Finally we performed these two tests by modifying the concept of "winner": in this case a fund was put in that category if it exceeded the 75[th] percentile's return.

2. *Spearman's rank order correlation and transition matrixes*. At every date of the selection period (ex.: at every quarter, if we are evaluating persistence on four months returns) we proceeded, first of all, to order the existing mutual funds on their realised performance. Then, we composed eight portfolios of funds (the first was composed as an equally weighted of top performer funds, the eighth of bottom performers[17]) and we evaluate their performance[18] in the subsequent period (evaluation period). Finally, we reordered the portfolios and we verified, by performing Spearman's test[19], if the ranking has been changed or not. This coefficient follows a t-Student distribution (asymptotically) with N-2 degree of freedom[20], so it is possible to appreciate its

---

[16] Although we evaluated the independence on the entire distribution of the sample, this c^2 test, like CPR, could have been calculated at every date.

[17] The exact allocation procedure, following Hendriks and others (1993), is as following. Lets assume N_t the number of existing funds at that date and $rank(r_i)$ the position of the i-th fund after ordering (in decreasing sense), then every fund will be assigned to the j-th portfolio, which satisfies: $(j-1)\left\lfloor N_t/8 \right\rfloor + \sum_{k=1}^{j-1} d_k \le rank(r_i) < j\left\lfloor N_t/8 \right\rfloor + \sum_{k=1}^{j} d_k$, where: $d_k = 1$ if $k \le (N_t \bmod 8)$ and $d_k = 0$ otherwise. ($\lfloor . \rfloor$ indicates the integer part of the fraction).

[18] The portfolio's performance is the average performance of its funds.

[19] Spearman's ranking order correlation assumes values between -1 (the first ranking is the opposite of the second) and 1 (the two rankings are identical), through 0 (the two rankings are completely independent). This coefficient captures rank order correlation and it is defined as: $r_s = 1 - \frac{6\Sigma D_i^2}{N(N^2 - 1)}$, where D indicates ranking difference in the two dates and N indicates the number of components (in our case N=8).

[20] Let's assume r Spearman's coefficient, its test statistic will be: $t = r\left[(N-2)/(1-r^2)\right]^{\frac{1}{2}}$.

statistical significance (see Kendall and others (1952)). In particular, we verified the absence of persistence by testing the null hypothesis of a Sperman's coefficient close to zero. This procedure has been recalculated at every date, so we take account when funds remained in the same octile (both in selection and evaluation periods) or when they changed their position (and toward which octile). In this way, we estimated the transition matrix on octiles between selection and evaluation periods.

3. *Cross-sectional regressions*. When we evaluated long run persistence, we utilised, as independent variable, the first lag of the dependent variable; while, when we evaluated the hot hand phenomenon, we utilised, as independent variables, the lags 1-6 of the dependent variable (four months returns[21]). This statistical technique evaluates the persistence by testing if returns on different periods are dependent. For these reasons, the null hypothesis of no persistence corresponds to a regression coefficient statistically close to zero.

In synthesis, we evaluated performance persistence following the scheme (on a two-period base) in figure 3.1.

Figure 3.1 – Experimental scheme of performance persistence's evaluation

Raw returns	**Short run persistence:** ➢ quarterly returns; ➢ annual returns;	**Statistical tests:** ➢ CPR e c^2 (on "winners" and "losers"); ➢ Spearman' s coefficient and transition matrixes (on octiles); ➢ cross-sectional regressions (on single funds);
Risk-adjusted returns: ➢ $a - J$; ➢ p^{T-M}; ➢ p^{H-M}; ➢ Sharpe; ➢ Treynor; ➢ Sortino;	**Short run persistence:** ➢ quarterly returns; ➢ annual returns;	**Statistical tests:** ➢ CPR e c^2 (on "winners" and "losers"); ➢ Spearman' s coefficient and transition matrixes (on octiles); ➢ cross-sectional regressions (on single funds);
	Long run persistence: ➢ on two periods of 5.5 years; ➢ on two periods of 2.5 years;	

[21] Unlike other tests (explained above), we did not perform cross-sectional regressions on yearly returns, because the third lag of the dependent variable corresponds exactly to the precedent year.

3.2 Empirical results[22]

3.2.1 *Raw returns*

The empirical results evidence the absence of the hot-hand effect on raw returns. In fact, even if the CPR exposed in table AI is significantly different from one, it refers only to the aggregate contingency table: considering the results of this test at every date, we observed that the periods of persistence are very fewer than those of no-persistence. In addition, also considering the aggregate contingency table, the χ^2 test (which is powerful and robust in presence of survivorship bias, as demonstrated by Carpenter and Lynch) asserts the absence of persistence.

3.2.2 *Long run persistence*

As in the precedent case, our results deny long run persistence (on risk-adjusted returns), whereas they evidence a weak tendency to reversal.

Reversal effects over long periods could be a reasonable matter for italian fund market. In fact, in a relative new market it appears rational that fund's disappearance would be related to performances over long periods. Besides, some characteristics of the analysed market (such as investor's insusceptibility to short periods of under-performance or the rigidity of supply's structure) seem to confirm the precedent intuition. Indeed, the market of fund's shares presents attrition costs that contrast the switching between funds[23].

So, if this is true, it could suggest numerous considerations about italian market and about its effects on the present analysis[24].

Performance persistence evaluation depends on analysis interval. Considering sample period length, empirical results evidence a weak tendency to reversal, whereas, taking into consideration the last five years of sample, fund's ranking is very independent between periods. In fact, if we are concerned about the values of CPR-test in table AII panel A (relative to sample period length), it is always below one (which indicates reversal); whereas, when we consider table AIII panel A (relative to the last five years of sample), the values of CPR are around one (see, for example, the values of CPR test for Sortino's indicator).

The relation between persistence and analysis period length is confirmed by the other tests. Considering B-panels of the tables above mentioned, we can observe the larger presence of negative

[22] Because of limited space, we present the most significant results of the research. In particular, we evidence the χ^2 results because, as demonstrated by Carpenter and Lynch (1999), it is a powerful and robust test in presence of survivorship bias.

[23] Consider that banks are the principal vendors of mutual fund's shares and that every bank does not supply shares of all the existing mutual fund (on italian market).

[24] It is well known that sample's characteristics may affect performance measurement and performance persistence (see for example Carpenter and Lynch (1999)).

signs in the first (tab. AII, relative to sample period length) than in the second (tab. AIII, relative to the last five years of sample length).

In every case, the conclusions suggested by empirical results present a low level of significance and the emerging point is the absence of long-run persistence on risk-adjusted returns.

3.2.3 Hot hand effect on risk-adjusted performances

Before analysing empirical results, some methodological specification is required.

As above mentioned, when we consider four-month risk-adjusted returns (with measures of Jensen and total performance) there is an overlapping on evaluation windows, which could generate spurious persistence.

To solve this problem and simultaneously to estimate these measures on 36-months lags, we implemented a model that assigns a higher weight to the non-overlapped observations.

This model, that is called "Generalised Least Squares", estimates a different value of the weight for any regression, which would be in the interval 0.05-0.3[25] (Harvey (1990)).

According to the experience and research on statistical series matured by J.P. Morgan, we utilised a λ-value of 0.03 (on the basis of the empirical research on monthly series, see Zangari and Longerstaey (1996)).

For this reason, when we refer to GLS measures (α-Jensen and total performance indicators), we refer to a measure in which both regressors and dependent variable are weighted for $(1-l)^{T-j}$ (where $T = 35$ and $j = 0,...,35$).

Taking into account empirical results, we observed, first of all, that they are related to the evaluation interval (yearly lags or four-month lags). In fact, while on yearly returns there is no persistence, on four-month returns we find evidence of persistence.

However, there is a weak uncertainty on yearly returns. Considering the results of χ^2-test for α-J and for Henriksson-Merton's total performance measure, we observe (tab. AIV, panel A) that, at a confidence interval of 92%, we can not reject the persistence hypothesis. But, while we introduce the correction for "windows overlapping", the same test does not reject the null hypothesis of no-persistence at a confidence interval of 85% (see tab. AIV, panel A, GLS measures).

We found evidence of persistence on four-month intervals. In fact, the larger part of these tests provides results in favour of persistence.

In particular, considering some measure of performance, all the implemented tests confirm the thesis of persistence; on the other hand, some test refuses the null hypothesis of no-persistence for

---

[25] The higher the coefficient, the lower the weight of far-off observations (the model is $m_t = l \sum_{j=0}^{T-1} (1-l)^j y_{t-j}$).

every measure. So, all the statistical tests on α-J and total performance (also after the correction) provide results in favour of persistence. On the other side, χ^2-test on "winners" and "losers" refuse, at confidence level of 7%, the null hypothesis of no-persistence for all the utilised indicators (except for Sharpe's coefficient; see tab. AIV)[26].

If we consider that: i) on the basis of Carpenter and Lynch (1999), χ^2-test is robust and powerful in presence of survivorship bias (as in our sample, also considering the mentioned limitations); ii) according with Brown and others (1992), when mutual fund survival is related to long-run persistence (as on italian market), statistical tests could be influenced by spurious reversal effects; then we can assert the presence of persistence on four-month performances.

In addition, it is important to observe that we found the highest level of persistence when "winners" and "losers" are defined with respect to the median return. However, such performance measure (α-J and total performance coefficients) realise stable orders also considering octiles subdivision (not only subdivision in two macro-classes).

Indeed, there are significant differences between the tests on α-J and total performance coefficients and tests on the other measures. Considering that this difference remains after the "overlapping correction", it seems explained by a structural difference on performance measure.

This is, in our opinion, a crucial point: performance persistence results may be related to the performance indicator. For this reason, it is important in empirical studies of performance persistence, to analyse the phenomenon with several indicators (which present different characteristics). In this sense, we consider that Agarwal and Naik (2000) could have utilised another measure in addition to α-J and appraisal ratio[27] (analogously for studies based on a single performance indicator); while the study of Kent and others (1997) may be robust in this sense[28].

In conclusion, when we considered four-month risk-adjusted returns, we found a degree of repeat performance. In particular, we found that mutual funds repeat their position relatively to the median return.

3.3 Multi-period tests of persistence

In this section, following Agarwal and Naik (2000), we extend our investigation from the traditional two-period framework to a multi-period framework.

In this paper as in Agarwal and Naik (2000), we utilise Kolmogorov-Smirnov's test to implement a multi-period analysis of persistence. Kolmogorov-Smirnov's test evaluates statistical differences

[26] We observed that the Spearman's test refuse the null hypothesis 23 times on 23 for α-J, 15 times on 23 for π^{H-M}, 11 on 23 for π^{T-M}, 1 on 23 for Sharpe, 1 on 23 for Treynor and 2 on 23 for Sortino.
[27] The appraisal ratio, while it is derived from the α-coefficient, it is high related to Jensen's indicator.
[28] The study of Kent and others is based on α-J, on raw returns and on a measure based on portfolio composition.

between the empirical distribution of "winners" and "losers" and the theoretical distribution in absence of persistence (binomial distribution).

Toward that end, we constructed, for all the fund of sample, a series of "winners" and "losers" and we computed the number of windows of four (three) years (four-months) extractable. Then, we computed the number of zero-winners-windows, of one-winner-windows, and so on[29]. On the other side, having the total number of extractable windows, binomial distribution provides us the theoretical number of these windows.

We must evaluate the results of this test with particular attention, because they could suggest groundless conclusions. In fact, some problem could arise from the low number of classes (only four classes for three-period-windows and five classes for four-period-windows): when it is too low, Kolmogorov-Smirnov's test refuses very easily the hypothesis of no-persistence.

For example, consider table AVII and chart AI. Observing only table AVII (which reports Kolmogorov-Smirnov's test on four year's windows for Sortino's measure), the test refuses the null hypothesis of no-persistence (at a significance of 5%). But, if we analyse the frequencies of "winners" (chart AIII), we observe that empirical distribution and theoretical distribution are not different.

In our opinion, this effect derives from the limited number of classes, which causes also the instability of the test. For this reason, conclusions based only on Kolmogorov-Smirnov's test are not significant.

4. Concluding remarks

In this work we evaluates the presence of “superior abilities” during the period: march 1988 – august 1999.

We investigated the presence of “talented managers” in two way. In the first case, we measured their capacity to generate extra-returns (given the risk level); in the second, we investigated the presence of funds that are always the top performers (persistence). In addition, with particular regard to this last topic, we captured relations between persistence and several variables such as: analysis period length, performance indicator, performing class (e.g. subdivision in octiles or in two macro-classes labelled “winner” and “loser”) and statistical test.

The analysis asserts that, in general, fund managers are not able to score extra-performances and to remain top performers during considerable periods. Instead, on four month lags, we found evidence of the “hot hand effect”.

[29] We performed multi-period tests on "winners" and "losers" defined with respect to the median return.

More precisely, referring to the capacity to generate extra-returns, mutual fund managers does not possess significant “stock picking” or “market timing” abilities (the performances does not differ from the equilibrium level).

Considering persistence analysis, we can summarise results as following:

1. there is not the hot hand phenomenon on raw returns;
2. there is not long run persistence on risk-adjusted returns (we found a weak evidence of reversal effect);
3. we found evidence of the hot hand effect on risk-adjusted returns on four-month intervals. In fact, while considering some measure of performance all the implemented tests confirm the thesis of persistence; some test refuses the null hypothesis of no-persistence for every measure. In every case, any evidence of persistence (which could be profitably exploited) “is loss” as soon as we analyse yearly intervals.

These results contrast with market efficiency. In fact, they evidence both the presence of “talented managers” and the possibility to realise superior performance by implementing adequate strategies (e.g. buy “winners” and sell “losers”).

But paradoxically, also when we found persistence, the realisation of the extra-returns is not an easily matter. Indeed, banks are the principal vendors of mutual fund's shares and every bank does not supply shares of all the existing mutual fund. In this way, transaction costs are always higher than the researched profits. All this considerations brigs us to conclude that, paradoxically, capital market inefficiency (in the sense of Fama) is explained by frictions of the market (transaction costs) that limits the switching between funds[30].

Our results are coherent with those of the precedent researches on italian fund market: in this work as in the studies of Ferretti and Murgia (1991) and of Cesari and Panetta (1998), there is no evidence of superior abilities.

Considering performance persistence, the empirical evidence on four-month intervals is coherent with italian market’s structure[31]. In fact, market’s rigidity limits the possibility to implement strategies that exploit the hot hand effect.

We retain that this study offers an exhaustive analysis of the considered phenomenon. In fact, in addition to analyse relation between returns, we appreciated the effect of several variables (such as: stability class, performance measure, evaluation lag and statistical test) on the levels of persistence. These factors, in explicit or implicit way, are always included in an analysis of performance

[30] The absence of friction, transaction costs, taxes, etc., is not required to define the market efficiency (in the sense of Fama), while it is required to define perfect markets.

[31] We remark that the phenomenon is not yet investigated on italian fund market.

persistence, but rarely (unique exception for the evaluation lag) researchers evaluated their effect on the levels of persistence[32]. Instead, this study evidences the higher level of performance when:

- we assumed a temporal lag of four-months;
- we evaluated funds performance using Jensen's indicator;
- we performed statistical tests on two macro-classes ("winners" and "losers" with respect to median return)[33].

[32] In our opinion, this is a relevant gap because, as we demonstrated, everyone of these variables affects the level of persistence.

[33] While we provides these results, we notify that, before any conclusion, it is important to perform several statistical tests on different combinations of these variables.

Bibliography

AGARWAL, V and N. Y. NAIK (2000), "Multi-period Performance Persistence Analysis of Hedge Funds", forthcoming paper on: *Journal of Financial and Quantitave Analisys*.

BROWN, S. J. and W. N. GOETZMANN (1995), "Performance Persistence", *Journal of Finance*, Vol. 50, No. 2, pp. 679-98.

BROWN, S. J., W. N. GOETZMAN, R. G. IBBOTSON and S. A. ROSS (1992), "Survivorship Bias in Performance Studies", *Review of Financial Studies*, Vol. 5, No. 4, pp. 553-580.

BROWN, S. J., W. N. GOETZMANN and R. G. IBBOTSON (1996), "Offshore Hedge Funds: Survival & Performance 1989-1995", *Journal of Business*, No. 72, pp. 91-117.

CARHART, M. M. (1997), "On Persistence in Mutual Fund Performance", *Journal of Finance*, Vol. 52, No.1, pp.57-82.

CARPENTER, J. N. and A. W. LYNCH (1999), "Survivorship Bias and Attrition Effects in Measures of Performance Persistence", *Journal of financial economics*, No. 54, pp. 337-374.

CESARI, R and F. PANETTA (1998), "Style, Fees and Performances of Italian Equity Funds", Temi di discussione del servizio studi Banca d'Italia, No. 325.

CHEN, N., R. ROLL. and S. ROSS (1986), "Economic Forces and the Stock Market", *Jurnal of Business*, Vol. 59, pp. 383-404.

CHRISTENSEN, R. (1990), "Log-linear Models", Springer-Verlag, New York.

CORNELL, B. and R. ROLL (1981), "Strategies for Pairwise Competitionsin Markets and Organizations", *Bell Journal of Economics*, pp. 201-213.

ELTON, E. J., M. J. GRUBER and C. R. BLAKE (1996), "The Persistence of Risk-adjusted Mutual Fund Performance", *Journal of Businnes*, Vol.69, No.2, 133-157.

FAMA, E. F. (1970), "Efficient Capital Markets: a Review of Theory and Empirical Work", *Journal of Finance*, Vol. 25, pp.383-417.

FAMA, E. F. (1991), "Efficient Capital Markets: II", *Journal of Finance*, Vol. 64, No. 5, pp. 1575-1617.

FERRETTI, R. and M. MURGIA (1991), 'Fondi Comuni di Investimento", in A. Penati (ed.), *Il rischio azionario e la borsa*, Milano, EGEA.

GOETZMANN, W. N. and R. G. IBBOTSON (1994), "Do Winners repeat? Patterns in Mutual Fund Performance", *Journal of portfolio managment*, Vol. 20 (spring), pp. 9-18.

GRINBLATT, M. and S. TITMAN (1989), "Mutual Fund Performance: An Analisys of Quarterly Portfolio Holdings", *Journal of Business*, Vol. 62, No. 3, pp. 393-416.

GRINBLATT, M. and S. TITMAN (1992), "The Persistence of Mutual Funds", *Journal of Finance*, Vol. 47, pp. 1977-1984.

GRINBLATT, M. and S. TITMAN (1994), "A Study of Montly Mutual Funds Returns and Performance Evaluation Techniques", *Journal of Financial and Quantitave Analisys*, Vol. 29, No. 3, pp. 419-444.

GROSSMAN, S. J. and J. STIGLITZ (1980), "The Impossibility of Informationally Efficient Markets", *American Economic Review*, pp. 393-408.

GRUBER, M. J (1996), "Another Puzzle: The Growth in Actively Managed Mutual Funds", *Journal of Finance*, Vol. 51, No. 3, pp. 783-810.

HENDRICKS, D., J. PATEL and R. ZECKHAUSER (1993), "Hot Hands in Mutual Funds: Short-Run Persistence of Performance", *Journal of Finance*, Vol. 48, pp.93-130.

HENRIKSSON, R. and R. MERTON (1981), "On Market Timing and Investment Performance II: Statistical Procedures for Evaluating Forecasting Skill", *Jurnal of Business*, Vol. 54, No. 3, pp. 513-534.

JENSEN, M. (1968), "The performance of mutual funds in the period 1945-1964", *Journal of Finance*, Vol. 23, No.1, pp. 389-416.

JENSEN, M. (1969), "Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolios", *Jurnal of business*, Vol. 42, No. 1, pp. 167-247.

KENDALL, M. G. and A. STUART (1952), "The Advanced Theory of Statistics", London, C. Griffin, vol. I.

KENT, D., M. GRINBLATT, S. TITMAN and R. WERMERS (1997), "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks", *Journal of Finance*, Vol. 52, no. 3, pp. 1035-1058.

LEHMAN, B. N. and D MODEST (1987), "Mutual Fund Performance Evaluation: a Comparison of Benchmarks and Benchmarks Comparison", *Journal of Finance*, Vol. 42, pp. 233-265.

MALKIEL, B. G. (1995), "Returns from Investing in Equity Funds", *Journal of Finance*, Vol. 50, pp.549-72.

MERTON, R. (1981), "On Market Timing and Investment Performance I: An Equilibrium Theory of Value and Market Forecasts", *Jurnal of Business*, Vol. 54, No. 3, pp. 363-406.

PANETTA, F. and E. ZAUTZIK (1990), "Evoluzione e Performance dei Fondi Comuni Mobiliari Italiani", Temi di Discussione del Servizio Studi della Bancad'Italia, No.142.

TREYNOR, J. L. and K. MAZUY (1966), "Can Mutual Funds Outgess the Market?", *Harvard Business Review*, Vol. 44, No. 4, pp.131-136.

WERMERS, R. (1997) "Momentum Investment Strategies of Mutual Funds, Performance Persistence and Survivorship Bias", Working Paper, Graduate School of Business and Administration, University of Colorado at Boulder, Boulder, Col. (http://bus.colorado.cdu/faculty/wermers/).

ZANGARI, P. and J. LONGERSTAEY (1996), "RiskMetrics Technical Document", 4th ed., New York: Morgan Guaranty Trust Company.

APPENDIX

Chart AI – Managed fund's distribution

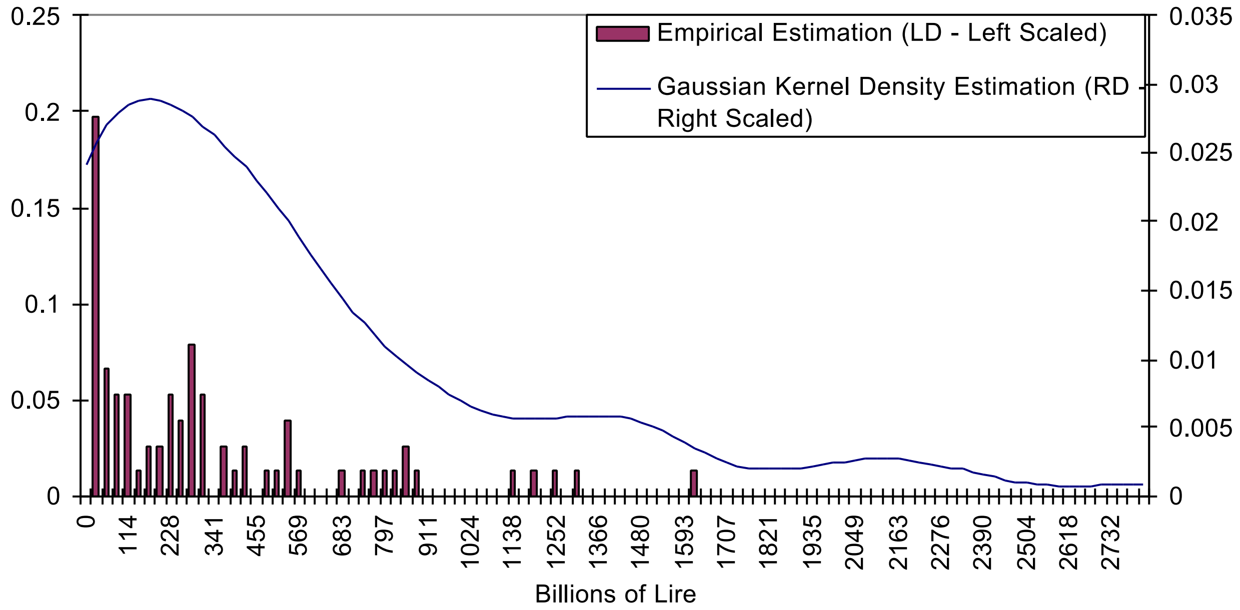
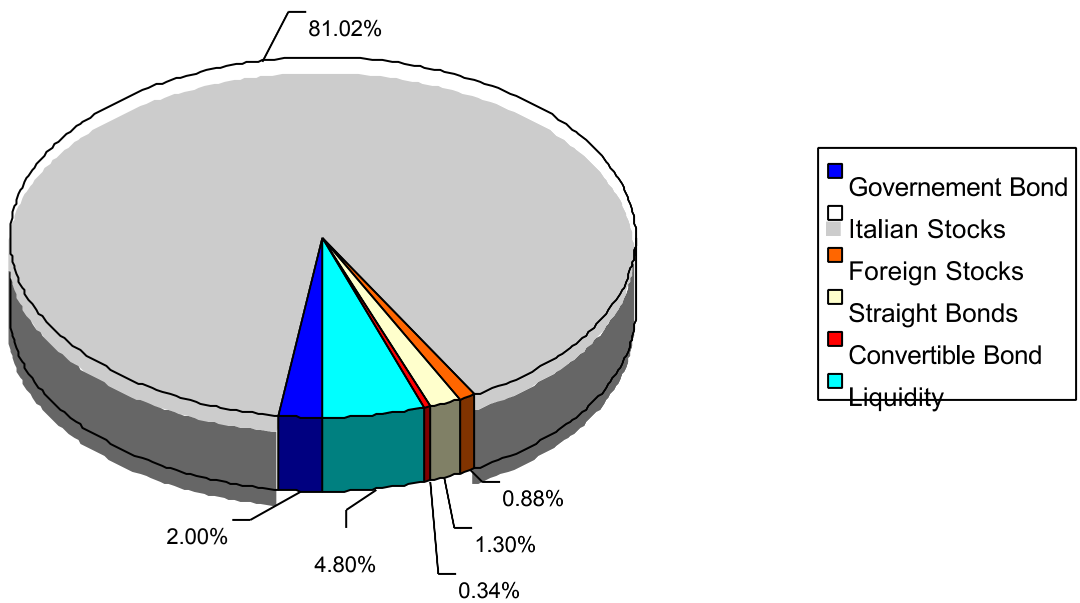


Chart AII – Managed fund's composition (category's average)



Managed fund's composition: average of the category "Azionari Italia" at august, 31st 1999.

Table AI – "Hot hand phenomenon" on raw returns

Panel A: "Winner" if it excess median return

Lag	N°	WW	LL	LW	WL	CPR	Z-statistic	c^2	p-value
4 months	1037	280,75	280,75	237,75	237,75	1,394	2,668 **	1,783	0,1818
Year	263	79,75	79,75	52,75	50,75	2,376	3,427 **	2,989	0,0838

Panel B: "Winner" if it excess 75th percentile

Lag	N°	WW	LL	LW	WL	CPR	Z-statistic	$c2$	p-value
4 months	1037	84,06	611,56	170,69	170,69	1,765	3,574 **	2,215	0,1367
Year	263	25,62	163,12	37,63	36,63	3,033	3,526 **	2,329	0,1270

* indicates 5% significance, ** indicates 1% significance.

"WW" indicates the number of persistence cases on "winners"; "LL" indicates the number of persistence cases on "losers"; "LW" e "WL" express the number of reversal cases (respectively from "loser" to "winner" and from "winner" to "loser"). "CPR" expresses the value of "Cross-Product Ratio"; "Z-statistic" indicates the value of statistic test on "CPR"; "χ^2" indicates the value of this test and "p-value" refers to χ^2 test.

Table AII – Long-run persistence (on sample period length)

Panel A[1]: "Cross-product ratio" and c^2

Measure	WW	LL	LW	WL	CPR	Z-statistic	c^2	p-value
α-J	4	4	5	5	0,64	-0,47	0,056	0,8129
$\pi^{H\text{-}M}$	3	3	6	6	0,25	-1,386	0,500	0,4795
$\pi^{T\text{-}M}$	5	5	4	4	0,64	0,47	0,056	0,8129
Sharpe	4	4	5	5	0,64	-0,47	0,056	0,8129
Treynor	4	4	5	5	0,64	-0,47	0,056	0,8129
Sortino	3	3	6	6	0,25	-1,386	0,5	0,4795

Panel B: Spearman's test and cross-sectional regression

Measure	Spearman		Cross-sectional coeff..		R^2 cross-sect. regr.
α-J	-0,5	(-1,3229)	-0,106	(-2,013)	0,2021
$\pi^{H\text{-}M}$	-0,524	(-1,386)	-0,096	(-1,326)	0,099
$\pi^{T\text{-}M}$	-0,214	(-0,567)	-0,085	(-1,108)	0,071
Sharpe	0	(0)	0,0014	(0,0225)	0,00003
Treynor	-0,0476	(-0,1256)	-0,0199	(-0,2503)	0,0039
Sortino	0,0238	(0,0623)	-0,0273	(-0,2614)	0,0042

[1] "Winners" e "Losers" are defined with respect to median return;

* indicates 5% significance, ** indicates 1% significance.

Cross-sectional regression has 16 degree of freedom.

"WW" indicates the number of persistence cases on "winners"; "LL" indicates the number of persistence cases on "losers"; "LW" e "WL" express the number of reversal cases (respectively from "loser" to "winner" and from "winner" to "loser"). "CPR" expresses the value of "Cross-Product Ratio"; "Z-statistic" indicates the value of statistic test on "CPR"; "χ^2" indicates the value of this test and "p-value" refers to χ^2 test.

"Spearman" reports the value of the test on octiles subdivision (in brackets the value of t-Student); "Cross-sectional coeff." expresses the regression coefficient (in brackets the value of t-Student).

Table AIII – Long-run persistence (last five years of sample)

Panel A[1]: "Cross-product ratio" e c^2

Measure	WW	LL	LW	WL	CPR	Z-statistic	c^2	p-value
α-J	6	6	11	11	0,248	-1,883	0,9317	0,337
π^{H-M}	9	9	7,5	7,5	1,44	0,52151	0,068	0,794
π^{T-M}	8,5	8,5	8	8	1,129	0,174	0,008	0,9287
Sharpe	8,5	8,5	8	8	1,129	0,174	0,008	0,9287
Treynor	8,5	8,5	8	8	1,129	0,174	0,008	0,9287
Sortino	8,25	8,25	8,25	8,25	1	0	0	1

Panel B: Spearman's test and cross-sectional regression

Measure	Spearman		Cross-sectional coeff..		R^2 cross-sect. regr.
α-J	-0,214	(-0,056)	-0,038	(-0,8118)	0,0190
π^{H-M}	-0,310	(-0,819)	-0,192	(-2,879)	0,196
π^{T-M}	-0,381	(-1,008)	-0,215	(-3,273)	0,240
Sharpe	0,5238	(1,3858)	0,129	(1,1585)	0,0415
Treynor	0,0238	(0,0623)	0,1523	(1,2531)	0,0482
Sortino	0,0952	(0,252)	0,1528	(0,6594)	0,0138

[1] "Winners" e "Losers" are defined with respect to median return;

* indicates 5% significance, ** indicates 1% significance.

Cross-sectional regression has 16 degree of freedom.

"WW" indicates the number of persistence cases on "winners"; "LL" indicates the number of persistence cases on "losers"; "LW" e "WL" express the number of reversal cases (respectively from "loser" to "winner" and from "winner" to "loser"). "CPR" expresses the value of "Cross-Product Ratio"; "Z-statistic" indicates the value of statistic test on "CPR"; "χ^2" indicates the value of this test and "p-value" refers to χ^2 test.

"Spearman" reports the value of the test on octiles subdivision (in brackets the value of t-Student); "Cross-sectional coeff." expresses the regression coefficient (in brackets the value of t-Student).

Table AIV – "Hot hand phenomenon" on risk-adjusted measures[1].

Panel A: Yearly interval

Measure	WW	LL	LW	WL	CPR	Z-statistic	c^2	p-value
α-J	50,25	50,25	28,75	28,75	3,055	3,377 **	2,926	0,0871
α-J GLS	47,75	47,75	31,25	31,25	2,335	2,606 **	1,723	0,1893
π^{H-M}	52	52	27	27	3,709	3,907 **	3,956	0,0467 *
π^{H-M} GLS	48,5	48,5	30,5	30,5	2,529	2,838 **	2,051	0,1521
π^{T-M}	49,25	49,25	29,75	29,75	2,741	3,070 **	2,407	0,1279
π^{T-M} GLS	44,25	44,25	34,75	34,75	1,621	1,508	0,571	0,4499
Sharpe	42,25	42,25	36,75	36,75	1,322	0,874	0,191	0,6620
Treynor	44,25	44,25	34,75	34,75	1,621	1,508	0,571	0,4498
Sortino	47	47	32	32	2,157	2,372 *	1,424	0,2327

Table AIV(follows) – "Hot hand phenomenon" on risk-adjusted measures[1].

Panel B: Four-month intervals

Measure	WW	LL	LW	WL	CPR	Z-statistic	c^2	p-value
α-J	222,75	222,75	67,75	67,75	10,81	12,132 **	41,351	0 **
α-J GLS	209,5	209,5	81	81	6,69	10,271 **	28,42	0 **
$\pi^{H\text{-}M}$	230,75	230,75	59,75	59,75	14,914	13,164 **	50,329	0 **
$\pi^{H\text{-}M}$ GLS	206,25	206,25	84,25	84,25	5,993	9,792 **	25,618	0 **
$\pi^{T\text{-}M}$	220,25	220,25	70,25	70,25	9,830	11,794 **	38,726	0 **
$\pi^{T\text{-}M}$ GLS	204,75	204,75	85,75	85,75	5,701	9,569 **	24,373	0 **
Sharpe	161,25	161,25	129,25	129,25	1,556	2,65 **	1,762	0,1843
Treynor	166,75	166,75	123,75	123,75	1,816	3,555 **	3,182	0,0744
Sortino	167	167	123,5	123,5	1,829	3,596 **	3,257	0,0711

[1] "Winners" e "Losers" are defined with respect to median return;

* indicates 5% significance, ** indicates 1% significance.

Cross-sectional regression has 16 degree of freedom.

"WW" indicates the number of persistence cases on "winners"; "LL" indicates the number of persistence cases on "losers"; "LW" e "WL" express the number of reversal cases (respectively from "loser" to "winner" and from "winner" to "loser"). "CPR" expresses the value of "Cross-Product Ratio"; "Z-statistic" indicates the value of statistic test on "CPR"; "χ^2" indicates the value of this test and "p-value" refers to χ^2 test.

"Spearman" reports the value of the test on octiles subdivision (in brackets the value of t-Student); "Cross-sectional coeff." expresses the regression coefficient (in brackets the value of t-Student).

Table AV – Spearman's coefficient on four-month returns ($\pi^{H\text{-}M}$ GLS)

Data	Spearman's coefficient	t-Student
29/02/92	-0,309524	-0.79733
30/06/92	-0,642857	-2.05575
31/10/92	0,5	1.41421
28/02/93	0,571429	1.70561
30/06/93	0,833333	3.692745 *
31/10/93	-0,547619	-1.60314
28/02/94	0	0.00000
30/06/94	0,428571	1.16190
31/10/94	0,785714	3.11127 *
28/02/95	0,952381	7.650921 **
30/06/95	0,904762	5.203364 **
31/10/95	0,97619	11.023524 **
29/02/96	0,97619	11.023524 **
30/06/96	0,97619	11.023524 **
31/10/96	0,761905	2.881441 *
28/02/97	0,714286	2.5 *
30/06/97	0,833333	3.692745 *
31/10/97	0,095238	0.23435
28/02/98	0,833333	3.692745 *
30/06/98	0,857143	4.076197 **
31/10/98	0,880952	4.560147 **
28/02/99	0,785714	3.11127 *
30/06/99	0,904762	5.203364 **

* indicates 5% significance, ** indicates 1% significance.

Table AVI – Transition matrixes on four-month returns (π^{H-M} GLS)

	1° octile	2° octile	3° octile	4° octile	5° octile	6° octile	7° octile	8° octile
1° octile	54,32%	13,92%	6,49%	10,81%	5,88%	3,03%	0,00%	4,84%
2° octile	24,69%	22,78%	16,88%	13,51%	14,71%	10,61%	3,13%	1,61%
3° octile	3,70%	24,05%	27,27%	18,92%	13,24%	6,06%	9,38%	0,00%
4° octile	4,94%	18,99%	20,78%	24,32%	20,59%	9,09%	3,13%	1,61%
5° octile	6,17%	7,59%	19,48%	14,86%	14,71%	16,67%	14,06%	8,06%
6° octile	2,47%	7,59%	6,49%	12,16%	20,59%	25,76%	15,63%	3,23%
7° octile	2,47%	3,80%	1,30%	1,35%	7,35%	25,76%	32,81%	20,97%
8° octile	1,23%	1,27%	1,30%	4,05%	2,94%	3,03%	21,88%	59,68%
Total	*100,00%*	*100,00%*	*100,00%*	*100,00%*	*100,00%*	*100,00%*	*100,00%*	*100,00%*

In column headline there is the ranking position on selection period; in raw headline there is the ranking position on evaluation period.
The matrix expresses <u>average probabilities</u>.

Table AVII – Kolmogorov-Smirnov's test on four-year periods (Sortino's measure)

Interval	Empirical distrib.	Theoretical distrib.	Absolute diff.
-inf. < x < 2.1875	0	0	0
2.1875 <= x < 3	0	0.4	0.4
3 <= x < 4	0.2	0.4	0.2
4 <= x < 7	0.4	0.4	0
7 <= x < 8	0.6	0.4	0.2
8 <= x < 8.75	0.6	0.4	0.2
8.75 <= x < 13	0.8	0.8	0
13 <= x < 13.125	1	0.8	0.2
13.125 <= x < inf.	1	1	0
		Max:	0.4

"Empirical distrib." is the "empirical distribution function" of Kolmogorov-Smirnov's test determined on observed frequencies; "Theoretical distrib." is the same distribution function. determined on theoretical frequencies (binomial distribution).
If some absolute difference excess 0.4 value, we accepted the null hypothesis of no-persistence at 5% significance.

Chart AIII – Frequency of "winners" on four-years returns (Sortino's measure)