On mutual fund performance evaluation

Maria do Céu Ribeiro Cortez (*)
Manuel José da Rocha Armada (**)

1 — Introduction

The evaluation of investment managers' performance has been a major topic in Finance for the past decades, concerning both practitioners and academics. It is of particular interest to managers (it obviously influentiates their compensation), clients/investors (who have the right to know how their money was applied) and regulators (in order to know how portfolio managers are allocating resources in the economy). Finally, it is of considerable interest to academics, since significant evidence of superior performance would violate the Efficient Market Hypothesis, which would have deep implications in Finance. The issue of measuring managers' performance has been discussed over three decades, and the debate still continues.

Based on the capital asset pricing model (CAPM), Treynor (1965), Sharpe (1966) and Jensen (1968) proposed risk-adjusted measures of performance — the so called traditional measures of performance evaluation, which have been widely used, inside and outside the academic circles. However, their effectiveness in providing precise has been questioned measures of performance, especially since the seventies.

Besides the conceptual and mainly econometric problems associated with these measures, another criticism that has been raised has to do with the fact that they only measure the performance in a global perspective, without attempting to analyze the components of timing and selectivity which contribute to overall performance. As Pfleiderer and Bhattacharya express, «some measurement techniques confound these two and thus produce poor indicators of true forecasting ability. By distinguishing these two sources of superior performance we may be able to produce more accurate measures of the total value of a manager's services» [Pfleiderer and Bhattacharya (1983, p. 2)]. Several researchers investigated this matter, in an attempt to develop models or methods which could provide, at least theoretically, separate measures of timing and selectivity.

In this context, we will investigate mutual funds investment managers' performance, in both overall terms as well as in terms of the components timing and selectivity, based on a sample of Portuguese mutual funds.

2 — A brief review of the traditional performance measures of Jensen, Treynor and Sharpe

According to the capital asset pricing model (CAPM), the expected return of a portfolio or asset p is positive and linearly related to the market return, as expressed by the security market line (SML):

$$E\,[R_{p,t}] = R_{f,t} + \beta_p\,(E\,[R_{m,t}] - R_{f,t}) \qquad \textbf{[1]}$$

where:

$E\,[R_{p,t}]$ = expected return on portfolio or asset $\boldsymbol{p}$ for period t;

$R_{f,t}$ = risk-free rate for period t;

$\beta_p = (\text{Cov}_{p,m})/\sigma^2_m$ = systematic risk measure of portfolio or asset p;

$E\,[R_{m,t}]$ = expected market return for period t.

(*) Assistant lecturer at the Department of Business and Public Administration, School of Economics and Business Administration, University of Minho, Braga, Portugal.
(**) Associate Professor of Finance at the Department of Business and Public Administration, School of Economics and Business Administration, University of Minho, Braga, Portugal.

In an ex-post perspective:

$$R_{p,t} - R_{f,t} = \beta_p (R_{m,t} - R_{f,t}) + \varepsilon_{p,t} \qquad [2]$$

where $\varepsilon_{p,t}$ is the residual term with the following properties:

$$E[\varepsilon_{p,t}] = 0,\ \text{Var}[\varepsilon_{p,t}],\ \sigma^2\varepsilon_{p,t},\ \text{Cov}[\varepsilon_{p,t}, R_{m,t}] = 0,$$
$$\text{Cov}[\varepsilon_{p,t}, \varepsilon_{j,t}] = 0$$

In equilibrium, this relationship should hold. However, if there are market inefficiencies, in which the portfolio manager believes some assets are not correctly priced, then, in this context, it is possible for him to obtain higher returns than those associated with their risk level. This being so, it is convenient not to constrain the regression to pass through the origin, allowing for an intercept α_p, which can be obtained from the following equation:

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon^*_{p,t} \qquad [3]$$

in which α_p is the measure of desequilibrium of portfolio or asset *p*.

The performance measure proposed by Jensen (1968) is precisely α_p of equation **[3]**, and can be interpreted as the return above (or below) the level predicted by the CAPM equilibrium relationship.

The technique proposed by Treynor (1965) is related to the previous one since it also uses the SML as the basis for performance comparison. Treynor's reward-to-variability ratio is:

$$T_p = (\overline{R}_p - \overline{R}_f) / \beta_p \qquad [4]$$

where:

T_p = Treynor's performance measure;
$\overline{R}_p$ = mean portfolio return for the period;
$\overline{R}_f$ = risk-free rate for the period;
β_p = portfolio beta for the period.

This indicator gives the excess return by unit of systematic risk. Higher values of T_p suggest, consequently, better performance.

The performance measure proposed by Sharpe (1966), unlike the two previous ones, uses as the measure of risk the total risk, expressed by the standard deviation of the portfolio's returns. The benchmark used for comparison is not the SML, but the capital market line (CML). Sharpe's reward-to-variability ratio, is given by:

$$S_p = (\overline{R}_p - \overline{R}_f) / \sigma_p \qquad [5]$$

where:

S_p = Sharpe's performance measure;
$\overline{R}_p$ and $\overline{R}_f$ as defined previously;
σ_p = standard deviation of the portfolio's returns.

3 — Criticisms to the traditional measures of performance evaluation

3.1 — Identification of the benchmark portfolio

One of the first criticisms pointed out to these traditional measures of performance evaluation has to do with the choice of the benchmark. The CAPM presents us a «natural» benchmark for comparison: the so called market portfolio. Since there is no exact quantifiable measure of such (market) portfolio, usually one uses a proxy (an index) as a substitute. Roll (1977, 1978, 1979, 1980, 1981) contests the use of these indexes, arguing that the resulting performance estimates can be biased, and showing that by using different indexes the ranking of portfolios by performance can be completely reverted.

Other authors also investigated this matter. For instance, Brown and Brown (1987), through the study of historical performance of portfolios relatively to different indexes, analyze the sensitivity of the performance measures to different specifications of the benchmark, having concluded that, in fact, the composition of the benchmark portfolio had influenced the results.

In order to overcome this potential problem, more recent literature has developed alternative techniques for evaluating performance, which eliminate the need of a benchmark portfolio for return/risk comparison. Grinblatt and Titman (1993) introduce a new performance measure which requires information about the portfolio composition, but not a benchmark.

3.2 — Correlation of the performance measures with risk

Another issue that can be raised relatively to the traditional measures of portfolio evaluation has to do with the extension to which they are related to risk. Theoretically, since they are risk-adjusted, their values shouldn't be related to the risk measures used (standard deviation or beta). However Friend and Blume (1970) demonstrate the opposite, that is, there is significant correlation (in their case negative). Other authors, such as [Klemkosky (1973), Ang and Chua (1979) and Wilson and Jones (1981), also found correlation between the risk measures and the performance measures.

3.3 — Time horizon for the calculation of returns

The influence of the chosen investment horizon for the calculation of the risk measures and, consequently, in the evaluation of performance has also been subject to research. The question is: is it indifferent to use the month, the quarter, the semester or the year as the time horizon for calculating returns? Fielitz and Greene (1980) and Levy (1981, 1984) studied this matter testing empirically the sensitivity of the risk measures to variations in the time horizon. The results obtained show that estimates of betas and performance may depend on the length of horizon over which they are calculated.

3.4 — Stability of risk measures

The traditional measures of portfolio evaluation assume that the risk measure remains stable over the evaluation period. In order to test the stability of risk, various authors carried out empirical studies. The results of studies on unmanaged portfolios show that there is some tendency for betas to regress to the mean [Blume (1971, 1975) and Levy (1971)]. Relatively to managed portfolios, we emphasize the study of Klemkosky and Maness (1978), Kon and Jen (1978, 1979) and Kon (1983), whose results suggest the unstability of portfolio's betas, as well as Fabozzi and Francis (1978, 1980), which support the idea that the betas move randomly through time rather than remain stable as the ordinary least squares model assumes.

4 — Alternative approaches to the CAPM: arbitrage pricing theory and stochastic dominance

The CAPM holds that the return on assets depends only on one factor: the market. Other authors, especially since the final seventies, defend that the return on assets may suffer the influence of more factors. These type of findings led some investigators to explore alternative theories to the CAPM, one being the arbitrage pricing theory (APT), developed by Ross (1976, 1977). APT can be synthesized through the following relation:

$$E\,[R_p] = R_f + \sum_{j=1}^{K} \beta_{p,j} \lambda j \qquad [6]$$

where:

$E\,[R_p]$ = expected return of asset (or portfolio) p;
R_f = risk-free rate (or rate of return of a zero beta portfolio);
$\beta_{p,j}$ = sensitivity measure of the return of asset (or portfolio) p to variations in factor j;
λ_j = risk premium relatively to factor j.

Although this idea — the search for factors that influence prices — concedes more flexibility (and attractiveness) to the model, however, many problems subsist at the empirical application level, which are due to the economic or statistical determination of the factors. Therefore, it was not the approach used in this study.

Stochastic dominance is an approach for ranking portfolios by preference order which uses the entire probability density function rather than a finite number of moments, such as the mean-variance approach (CAPM). This approach was initially developed by Quirk and Saposnik (1962) (1st degree) and later extended by Hadar and Russell (1969) and Hanoch and Levy (1969) (2nd degree), by Whitmore (1970) (3rd degree) and by Jean (1971, 1978) (with the mathematical proof for the n^{th} degree). However, it also raises problems at the application level, as it is not always possible to determine clearly a ranking (dominance) of portfolios, since the cumulative distribution probabilities can intercept for one or more levels of return. It also does not seem the convenient approach to investigate timing and selectivity capacities since its main objective is the ranking of portfolios by the comparison of the empirical distributions of returns.

5 — Timing and selectivity

As pointed out before, one of the assumptions inherent to the traditional measures of performance is that the portfolio's level of risk is stable.

They evaluate performance that is due solely to the manager's ability to forecast future security prices, that is, they show the manager's capacity to select mispriced securities: the capacity of *selectivity*. Such measures turn out to be insufficient since they do not consider the possibility that the manager could predict and anticipate general market price movements, in other words, the manager's *timing* capacity. It seems obvious that overall portfolio performance can be due both to stock selection (microforecasting), which is based on forecasts of company-specific events, and market timing (macroforecasting), which refers to the manager's capacity in predicting the direction of market movements which will affect all securities. In this case the manager will attempt to adjust the level of systematic risk in anticipation of the predicted market movements, thus allowing for higher returns. This adjustment can be achieved by 1) switching from high-beta stocks to low-beta stocks (or vice-versa); or 2) changing the proportions invested in the risk-free security. The point here to be emphasized is that the level of systematic risk (beta) is a decision variable that the manager can make use to increase the return of the portfolio.

The distinction between the part of return attributable to selectivity and that attributable to market timing has been receiving considerable interest in the literature. One of the early studies in this area was presented by Treynor and Mazuy (1966), which used a quadratic regression (in contrast to the linear relation sustained by the CAPM) to detect timing capacities of mutual fund managers.

Fama (1972) was the first to propose a formal (theoretical) methodology for the decomposition of the total return into the components of timing and selectivity. However, the implementation of his measures presents difficulties, since the required information is not easily available.

Kon and Jen (1978, 1979) criticize the use of the ordinary least squares technique to obtain the estimates of performance since it assumes that beta remains stable through time; yet, as we have explained, the manager can change the systematic risk level of the portfolio as a result of a timing strategy. In this context, these authors examine the existence of nonstationary risk levels through a technique of switching regression. An alternative procedure to test the variability of betas in bull and bear markets was proposed by Fabozzi and Francis (1979), through the use of dummy variables.

Merton (1981) and Henriksson and Merton (1981) develop nonparametric tests of market-timing forecasts without assuming a CAPM framework, and parametric ones showing that the theoretical structure of a timing strategy is similar to the return patterns of an option strategy (of the put-protective type). Their results were that, in general, fund managers have neither of those forecasting abilities.

By correcting an error made in Jensen (1972), Pfleiderer and Bhattacharya (1983) developed a regression technique based on the quadratic regression of Treynor and Mazuy (1966), which requires information on the portfolio returns and market returns, and provides separate timing and selectivity measures, at least theoretically.

Since the traditional measures of performance evaluation are insufficient (they only allow for estimates of selectivity not timing, at least as separated measures [1]) for investigating empirically the performance of mutual funds in terms of those two components, we chose, among the reviewed approaches, the model proposed by Lee and Rahman (1990) [based upon Pfleiderer and Bhattacharya (1983)] and tested in the USA by themselves and by Armada (1992) in the UK. The choice was due not only to the fact that it is a relatively recent approach, but also because the authors argue that one can *separate* timing from selectivity and that it has not been tested in the generality of the European countries, particularly in Portugal.

6 — Traditional measures of portfolio performance: empirical analysis for the Portuguese case

Once we have reviewed the literature we will start by the empirical examination of the Jensen, Treynor and Sharpe measures of performance on a sample of Portuguese mutual funds.

[1] Despite that fact, but for reasons stated previously, we also empirically studied these measures.

6.1 — Data description

Our sample, presented in the table 1, consists of bi-weekly returns on seven mutual funds over the four year period from 15th March 1989 to 15th February 1993.

TABLE 1

The sample of mutual funds

1. FIP	2. FUN	3. INV	4. MUL	5. PRI	6. UNI	7. VAL

These mutual funds were chosen according to the following criteria: 1) composition of the portfolio (that is, classified as equity funds); and 2) availability of information for the period. It would have been desirable to chose a larger time horizon, however, regarding the specificity of the Portuguese mutual fund market as well as its very recent development, the consideration of a larger time period would have dramatically reduced the number of funds available. In view of such circumstances, we used bi-weekly returns which allowed for 95 observations for each fund.

The information required for the calculation of returns was obtained from the *Bulletin of Quotations* of the Oporto Stock Exchange and also directly from the mutual fund companies, in such a way to guarantee the credibility of the inputs [2]. The returns, adjusted for dividends, were computed as follows:

$$R_{p,t} = (P_{p,t} + D_{p,t} - P_{p,t-1}) \;/\; P_{p,t-1} \qquad \textbf{[7]}$$

where:

$R_{p,t}$ = total return for fund p in period t;
$P_{p,t}$ = price of fund p at the end of period t;
$D_{p,t}$ = Dividend per unit paid by fund p during period t.

For the calculation of total market return the general index of the Lisbon Stock Exchange (BVL) and the Banco Totta e Açores index (BTA), both adjusted to dividends, were used. In this way, besides the determination of the traditional measures of performance, we can test empirically the question raised by Roll (1980, 1981) about the impact of using different indexes in the ranking of portfolios, already discussed in the review of the literature.

The market return, adjusted for dividends was computed as following:

$$R_{m,t} = (I_{m,t} - I_{m,t-1}) \;/\; I_{m,t-1} \qquad \textbf{[8]}$$

where:

$R_{m,t}$ = market return for period t;
$I_{m,t}$ = value of the index for period t.

The returns on the risk-free rate were calculated from the medium 30-day Treasury Bill returns obtained from the Bank of Portugal.

6.2 — Empirical evidence

Tables 2 and 3 below present the estimated measures of Jensen ($\hat{\alpha}_p$) and systematic risk ($\hat{\beta}_p$) obtained from equation **[3]**. The funds were ranked from highest to lowest performance.

TABLE 2

Estimates of the regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon^*_{p,t}$ for 7 mutual funds over the period from 89-03-15 to 93-02-15, utilizing the BVL index

Funds	$\hat{\alpha}_p$	*t*-stat.	$\hat{\beta}_p$	*t*-stat.	R^2 (percentage)
VAL	− 0.001 657 0	− 1.395 27	0.092 280 5	2.790 13*	07.72
UNI	− 0.002 346 7	− 1.721 85*	0.296 866 0	7.821 10*	39.68
FUN	− 0.002 797 1	− 1.102 62	0.000 637 3	0.009 02	0.00
INV	− 0.002 885 2	− 1.644 68	0.123 991 0	2.537 86*	6.48
FIP	− 0.003 172 5	− 1.721 36*	0.136 874 0	2.666 64*	7.10
MUL	− 0.003 711 0	1.695 54*	0.124 801 0	2.047 38*	4.31
PRI	− 0.004 218 2	− 2.101 56*	0.277 350 0	4.961 50*	20.93

[2] This information was carefully checked as we have, inclusively, compared the data supplied by these sources of information.

TABLE 3

Estimates of the regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon^*_{p,t}$ for 7 mutual funds over the period from 89-03-15 to 93-02-15, utilizing the BTA index

Funds	$\hat{\alpha}_p$	t-stat.	$\hat{\beta}_p$	t-stat.	R^2 (percentage)
VAL	– 0.001 570 8	– 1.324 27	0.092 871 9	2.944 99*	8.53
UNI	– 0.002 080 4	– 1.572 17	0.297 686 0	8.461 82*	43.50
INV	– 0.002 087 6	– 1.597 39	0.120 960 0	2.588 60*	6.72
FUN	– 0.002 928 7	– 1.150 98	– 0.012 609 8	– 0.186 40	0.04
FIP	– 0.003 038 7	– 1.650 67	0.138 354 0	2.826 94*	7.91
MUL	– 0.003 449 5	– 1.584 42	0.140 140 0	2.421 16*	5.93
PRI	– 0.003 837 3	– 1.959 07*	0.291 356 0	5.594 96*	25.18

It immediately stands out the fact that all funds exhibit negative $\hat{\alpha}_p$, varying from – 0.00166 (VAL) to – 0.00422 (PRI) and from – 0.00157 (VAL) to – 0.00384 (PRI), depending if the index used is BVL or BTA, respectively. Funds UNI, FIP, MUL, PRI and fund PRI using, respectively, the BVL and the BTA index, have significantly negative estimates of $\hat{\alpha}_p$ at the 0.05 level. The t-stat values for the remaining funds lead credibility to the null hypothesis that $\alpha_p = 0$. Thus, these results do not show any evidence of mutual fund managers' capacity to forecast security prices. In other words, they did not show evidence of selectivity.

For reasons stated previously (particularly when reviewing the literature), it is of interest to compare the ranking of funds by performance according to both indexes. For doing so, we utilized a measure of rank correlation, the Spearman's Rank Correlation Coefficient (r_s) [3]. The correlation between the ranking of funds given by the BVL index (table 2) and the ranking of funds given by the BTA index (table 3) expressed by r_s is 96.4% [4], which is a very strong positive correlation, demonstrating that the ranking of funds by either indexes are very similar.

Relatively to the estimates of systematic risk ($\hat{\beta}_p$), it was at first surprising to observe very low values, ranging from 0.00006 (FUN) to 0.29687 (UNI) (table 2) and from – 0.00126 (FUN) to 0.29768 (UNI) (table 3). However, a closer analysis to the composition of the portfolios would easily justify these values. In fact, since these funds were classified as equity funds, it would be expected to find a high proportion of stocks in their composition! But this was not observed, as one can check in appendix 2.

It should be stressed that the portfolios which present the lower betas in both tables (FUN and VAL) are those who have, on average, a less percentage of stocks in their composition, which is an expected result. In particular, fund FUN presents betas of 0.00064 (BVL index) and even negative – 0.01261 (BTA index)! If we analyze its composition, we could verify that the percentage of stocks it holds in the beginning is 5.49%, and this number decreases gradually until June 1991, when these securities cease completely to be held by the portfolio[5]. The situation of a declining stock market in the period of consideration (as can be seen at the end of appendix 2) explains somehow that stocks did not make up for a significant portion of the portfolios' composition, in contrast with bonds and other low risk securities. In this context, it is therefore comprehensive that $\hat{\beta}_p$'s (which express the portfolios' sensitivity to the market) exhibit such low values. Following the same reasoning, it is curious to point out that the portfolios which presented higher $\hat{\beta}_p$'s (UNI and PRI) also had, on average, higher percentage of stocks, notwithstanding the fact that these values were only about 50 % (and less)!

The results persist if we divide the total period up into two subperiods [6]: the α_p estimates remain similar and negative whichever subperiod and market index considered. For the same subperiod

[3] The formula for the computation of r_s is $r_s = 1 - \frac{6\sum_{i=i}^{n}(Xi - Yi)^2}{n(n^2-1)}$ where n is the number of values for ranking, Xi is the rank order of the i^{th} variable and Yi is the order of the i^{th} variable Y.

[4] The calculation of r_s is presented in appendix 1.

[5] In spite of that, it continued to be classified as an equity fund!

[6] Since the results do not differ significantly, they are presented in appendix 3.

the rank correlation between both indexes is 96.4 % [7], which represents a very similar ranking, already observed for the global period. In relation to the $\hat{\beta}_p$ estimates, once again we observe identical results (very low values), which is not surprising for the reasons stated previously.

It was also of our concern to investigate the possibility of heteroscedasticity, which would question the least squares estimators. In this sense, we used White's (1980) method for eventual detection and correction of these estimates [8]. The differences in the results with and without correction for heteroscedasticity are not substantial [9]: negatives values of $\hat{\alpha}_p$, low values for $\hat{\beta}_p$, and a strong correlation between the ranking of results corrected and not corrected for heteroscedasticity (89% and 82% respectively for indexes BVL and BTA [10]).

It is well known that outliers (leverage points [11]) may totally spoil a least squares (LS) regression and, in particular, affect the t-based significance levels. We, in fact, carried out the robust regression as described in appendix 8 in order to detect potential outliers but the results we got were similar to those already obtained so that we did not even include them.

Although, for obvious reasons, Jensen's measure is more related with the methodology that we will follow, for studying timing as well as selectivity, the Treynor and Sharpe measures of performance were also computed for reasons also stated previously. The results are shown in tables 4 to 6 [12] below:

TABLE 4

Estimates of Treynor's measure ($T_p = (\bar{R}_p - \bar{R}_f) / \beta_p$) for the period 89-03-15 to 93-02-15, utilizing the BVL index

Funds	$\hat{T}_p$
UNI	– 0.040 23 92
PRI	– 0.049 17 78
FIP	– 0.082 66 63
INV	– 0.087 99 26
MUL	– 0.094 09 78
VAL	– 0.101 79 14

TABLE 5

Estimates of **Treynor's** measure ($T_p = (\bar{R}_p - \bar{R}_f) / \beta_p$) for the period **89-03-15** to **93-02-15**, utilizing the **BTA** index

Funds	$\hat{T}_p$
UNI	– 0.040 128 40
PRI	– 0.046 813 74
FIP	– 0.081 781 98
MUL	– 0.083 798 30
INV	– 0.090 197 53
VAL	– 0.101 143 23

TABLE 6

Estimates of **Sharpe's** measure ($S_p = (\bar{R}_p - \bar{R}_f)\ \sigma_p$) for the period **89-03-15** to **93-02-15**

Funds	$\hat{S}_p$
MUL	– 0.559 664 39
FIP	– 0.630 976 72
INV	– 0.641 348 13
PRI	– 0.644 334 39
UNI	– 0.725 906 57
VAL	– 0.810 219 76

Besides the negative estimates, we still continue to observe a strong correlation between the rankings achieved with both indexes, as we can confirm by the r_s value, which is 94.3% for the ranking of funds according to T_p [13].

7 — Timing and selectivity: empirical analysis for the Portuguese case

The traditional measures of portfolio performance, particularly Jensen's measure, are insufficient since they do not allow the possibility of changing the portfolio's systematic risk level as a result of timing strategies. This being so, and for reasons presented before, it becomes necessary to adopt measures that allow for the decomposition of total return up into its timing and selectivity components. In this context, we will apply the Lee and Rahman (1990) model [based upon Pfleiderer and Bhattacharya (1983)] to the sample of mutual

[7] The calculations of r_s are presented in appendix 4.
[8] A description of White's method can be shown in appendix 5.
[9] The results corrected for heteroscedasticity are presented in appendix 6.
[10] See appendix 7 for the computation of r_s.
[11] That is, cases for which (X_{i1}, ... , X_{ip}, Y_i) deviates from the linear relation followed by the majority of the data, taking into account both the explanatory variables (X_{ip}) and the response variable (Y_i) simultaneously.
[12] Although not affecting Sharpe's measure, we did not include portfolio FUN since the beta estimates it presents are close to zero.
[13] The calculations are identical to those presented previously.

funds presented before, in such a way to obtain separate measures of timing and selectivity.

Pfleiderer and Bhattacharya (1983) sustain that it is possible to *separate* the stock selection ability from timing ability as described below.

From the following quadratic regression [Lee and Rahman (1990)]:

$$(R_{p,t} - R_{f,t}) = \eta'_0 + \eta'_1(R_{m,t} - R_{f,t}) + \eta'_2 (R_{m,t} - R_{f,t})^2 + \omega'_{p,t} \quad [9]$$

whose large-sample coefficients are:

$$\text{plim } \hat{\eta}'_0 = \alpha_p \quad [10]$$

$$\text{plim } \hat{\eta}'_1 = \theta E(R_{m,t} - R_{f,t})(1 - \psi) \quad [11]$$

$$\text{plim } \hat{\eta}'_2 = \theta\psi \quad [12]$$

where α_p is the estimate for selectivity. From the residual term $\omega'_{p,t} = \theta\psi\varepsilon_{p,t}(R_{m,t} - R_{f,t}) + u_{p,t}$, we run the following regression:

$$(\omega'_{p,t})^2 = \theta^2\psi^2\sigma^2_\varepsilon (R_{m,t} - R_{f,t})^2 + \zeta_{p,t} \quad [13]$$

which produces the estimate of $\theta^2\psi^2\sigma^2_\varepsilon$. Since we know $\theta\psi$, recovered through **[12]** we can obtain σ^2_ε. This, coupled with knowledge about σ^2_π [14], allows us to estimate $\psi = \sigma^2_\pi/(\sigma^2_\pi + \sigma^2_\varepsilon) = \rho^2$. Finally, we calculate ρ, which is the measure of timing.

7.1 — Empirical evidence

Utilizing the same sample of funds as well as all the other inputs described before [15], the estimated obtained by applying the model were the following:

TABLE 7

Estimated parameters (α_p e ρ) for 7 funds over the period Mar 89-Feb 93

Funds	Selectivity ($\hat{\alpha}_p$)	Timing ($\hat{\rho}$)
FIP	– 0.003 178	0.000 705 77
FUN	– 0.002 927	0.021 711 60
INV	– 0.003 770*	0.091 129 32
MUL	– 0.003 781	0.011 472 25
PRI	– 0.004 163*	0.004 395 86
UNI	– 0.001 851	0.055 169 89
VAL	– 0.002 600*	0.240 001 67*

* Statistically significant at 0.05 level.

As we can observe, all funds present negative $\hat{\alpha}_p$, with three funds (INV, PRI and VAL) having significant negative values of $\hat{\alpha}_p$ at the 0.05 level. Thus, these results do not show evidence of selectivity. As for timing, the results suggest that the managers also could not anticipate market movements in such a way to obtain higher returns: only one fund (VAL) has a $\hat{\rho}$ statistically different from zero at the 0.05 level [16].

Finally, we also observed a negative correlation between selectivity and timing, of about 38%, a phenomenon already observed within the context of other financial models for the same purpose [see Armada (1992)], but we did not investigate this puzzling issue in finance in this paper.

8 — Conclusions

The results obtained suggest that, for the period March 89- February 93, the mutual fund managers could not forecast individual security prices in order to beat the market, as suggested by the Jensen (1968), Treynor (1965) and Sharpe (1966) negative estimates of performance. On the other hand, when separate measures of timing and

[14] Jensen (1972) defines $\pi_t = (R_{m,t} - R_{f,t}) - E(R_{m,t} - R_{f,t})$. Merton (1980) shows how to obtain estimates of the variance of π_t through:

$$\hat{\sigma}^2_\pi = \frac{\sum_{t=1}^{n}[1n[1 + (R_{m,t} - R_{f,t})]^2]}{n}$$

[15] Since, on the one hand, there exists a very strong positive relation between the ranking of funds based on either indexes and that, on the other hand, heteroscedasticity does not seem to be a problem, we will consider from now on only one index (BVL) and not heteroscedasticity.

[16] We did also consider the possible influence of ouliers but, again, the results were not substantially altered.

selectivity were applied, according to the model developed by Pfleiderer and Bhattacharya (1983) and later developed and implemented by Lee and Rahman (1990), the results persisted. Besides the persisting negative levels of selectivity, the timing estimates were low: only one fund showed that capacity significantly; the others did not show evidence of successful systematic risk-adjustment to market movements.

One possible explanation was already anticipated: these mutual funds, though classified as equity funds, held during this period [17] a relatively small percentage of these assets. So, we should raise the pertinent question of the choice of the proper benchmark (clearly a urgent task to carry out in Portugal), taking into consideration, more than the classification of the funds, the structure/ composition of their portfolios and apply adequate frameworks for assessment.

Some other possible explanations can be suggested: the existence of transaction costs (which would not be compensated with the potential gains coming up from changing the composition of the portfolio), legal restrictions, and even no forecasting capacities by the managers. In this case, if there are no forecasting capacities, either at the macro or micro level, then one should consider following a passive strategy through, for instance, the construction of portfolios that reflect the market composition (index portfolios). Here, the managers' efforts would concentrate in providing a diversification service to the client (instead of concentrating in efforts of stocks selection and market timing) with obvious savings from research and transaction costs.

However, we believe other reasons (inclusive, at a theoretical level) might lie behind these unfavorable results. It has been a practice to carry out this type of research utilizing only series of prices (returns) as inputs, which assumes symmetry of information between the portfolio managers and the performance evaluators. This usually is not true. Thus, it would be important to consider the composition of the portfolios into the analysis. It is in this direction that the line of investigation proposed by Elton and Gruber (1991) points out. These authors attempt to analyze timing and selectivity considering not only the portfolio returns as inputs, but also other elements from the portfolio composition. Also it should be pointed out the work from Ferson and Schadt (1993), who propose conditional models.

Finally, a possible direction to follow in future investigation would be the evaluation of portfolio managers containing, besides the «traditional» assets, other products, namely derivative instruments, such as futures and options (clearly theoretical structures which assume that normality is not appropriate). Following these lines of investigation would certainly result in a research project of an enormous potential.

[17] Perhaps for comprehensive reasons!

Appendices

APPENDIX 1

Calculation of r_s between the ranking of funds according to the two indexes BVL and BTA (global period — Mar. 89-Feb. 93)

BTA			BVL		
Funds	α	Rank	Funds	α	Rank
VAL	– 0,001 570 8	1	VAL	– 0,001 657 0	1
UNI	– 0,002 080 4	2	UNI	– 0,002 346 7	2
INV	– 0,002 807 6	3	FUN	– 0,002 797 1	3
FUN	– 0,002 928 7	4	INV	– 0,002 885 2	4
FIP	– 0,003 038 7	5	FIP	– 0,003 172 5	5
MUL	– 0,003 449 5	6	MUL	– 0,003 711 0	6
PRI	– 0,003 837 3	7	PRI	– 0,004 218 2	7

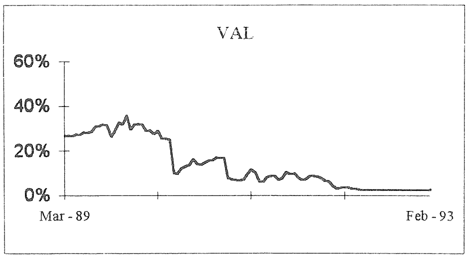
	Rank BVL — *Xi*	Rank BTA — *Yi*	*Xi-Yi*	(*Xi-Yi*)2
VAL	1	1	0	0
UNI	2	2	0	0
INV	4	3	1	1
FUN	3	4	– 1	1
FIP	5	5	0	0
MUL	6	6	0	0
PRI	7	7	0	0

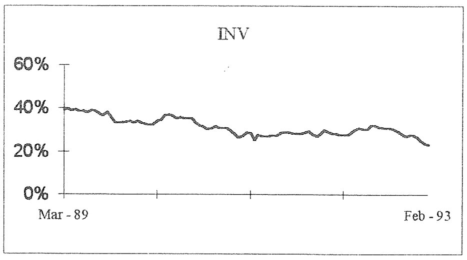
Total = 2.

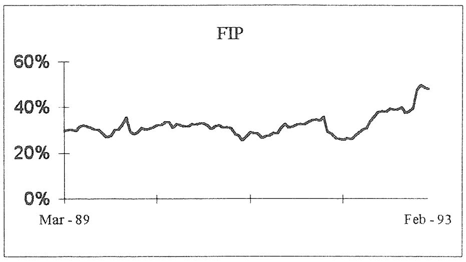
r_s = 0.964 285 714.

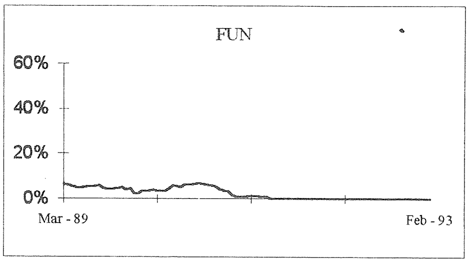
APPENDIX 2

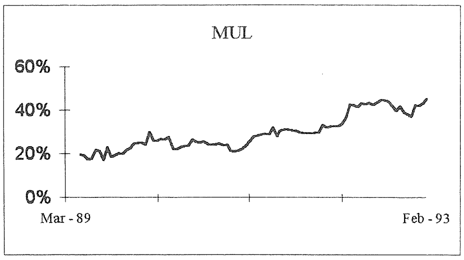
Percentage of stocks in the composition of the mutual funds

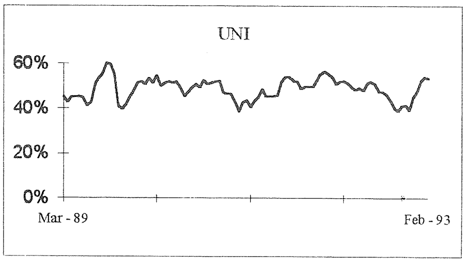


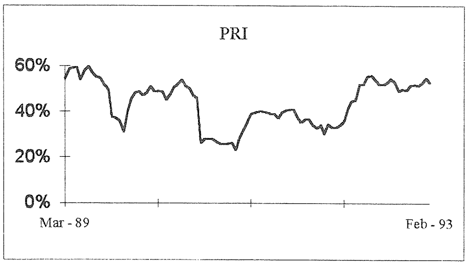






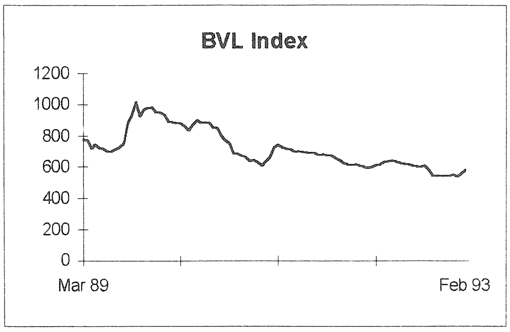


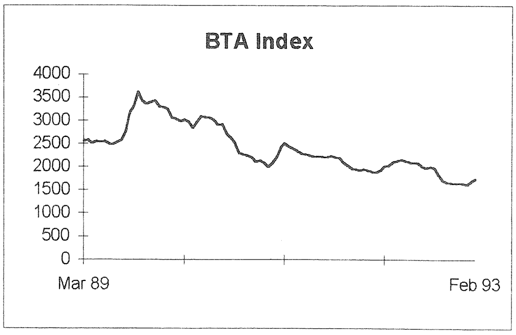




Source: *Bulletin of Quotations* of the Oporto Stock Exchange.

Market indices and returns in the evaluation period





APPENDIX 3

Estimates of Jensen's measure ($\hat{\alpha}_p$) and the systematic risk measure ($\hat{\beta}_p$) for the subperiods 89-03-15 to 91-02-15 and 91-03-01 to 93-02-15

Funds	$\hat{\alpha}_p$	*t*-stat.	$\hat{\beta}_p$	*t*-stat.	R^2 (percentage)
VAL	– 0.001 475 8	– 0.853 17	0.108 352 0	2.798 22	14.82
INV	– 0.001 667 9	– 0.803 89	0.119 864 0	2.580 80	12.89
UNI	– 0.002 109 0	– 0.978 76	0.230 729 0	4.783 30	33.71
FUN	– 0.002 277 4	– 0.639 23	0.033 418 9	0.419 02	0.39
FIP	– 0.002 983 7	– 0.082 36	0.113 722 0	1.402 22	4.19
MUL	– 0.002 993 7	– 0.922 72	0.110 388 0	1.519 89	4.88
PRI	– 0.005 933 8	– 1.659 59	0.220 168 0	2.750 74	14.39

Estimates of the regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon^*_{p,t}$ for 7 mutual funds in the period 89-03-15 to 91-02-15, using the BVL index.

Funds	$\hat{\alpha}_p$	*t*-stat.	$\hat{\beta}_p$	*t*-stat.	R^2 (percentage)
UNI	– 0.000 753 4	– 0.515 12	0.542 200 0	9.063 98	64.11
PRI	– 0.000 916 8	– 0.524 89	0.493 887 0	6.913 20	50.96
VAL	– 0.002 283 4	– 1.334 53	0.032 110 7	0.458 80	0.46
FIP	– 0.002 720 4	– 3.559 59	0.222 521 0	7.118 30	52.42
INV	– 0.003 989 9	– 1.339 43	0.136 542 0	1.120 62	2.66
MUL	– 0.004 030 3	– 1.296 09	0.176 741 0	1.389 56	4.03
FUN	– 0.004 225 2	– 1.112 28	– 0.122 438 0	– 0.787 99	1.33

Estimates of the regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon^*_{p,t}$ for 7 mutual funds in the period 91-03-01 to 93-02-15, using the BVL index

Funds	$\hat{\alpha}_p$	*t*-stat.	$\hat{\beta}_p$	*t*-stat.	R^2 (percentage)
VAL	– 0.001 414 8	– 0.829 36	0.118 276 0	3.047 74	17.11
INV	– 0.001 605 8	– 0.782 52	0.130 190 0	2.788 74	14.74
UNI	– 0.002 058 5	– 0.972 68	0.242 307 0	5.032 82	36.02
FUN	– 0.002 246 0	– 0.630 88	0.038 002 2	0.469 23	0.49
MUL	– 0.002 810 8	– 0.876 77	0.135 031 0	1.851 48	7.08
FIP	– 0.003 006 0	– 0.829 39	0.113 754 0	1.379 66	4.06
PRI	– 0.005 721 7	– 1.636 69	0.250 948 0	3.155 41	18.12

Estimates of the regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon^*_{p,t}$ for 7 mutual funds in the period 89-03-15 to 91-02-15, using the BTA index.

Funds	$\hat{\alpha}_p$	*t*-stat.	$\hat{\beta}_p$	*t*-stat.	R^2 (percentage)
UNI	– 0.001 060 2	– 0.697 57	0.426 289 0	8.452 85	60.83
PRI	– 0.001 140 2	– 0.647 11	0.393 117 0	6.723 53	49.56
VAL	– 0.002 209 5	– 1.299 92	0.033 186 1	0.588 39	0.75
FIP	– 0.002 609 1	– 3.842 41	0.195 386 0	8.671 24	62.04
MUL	– 0.004 009 6	– 1.298 79	0.149 350 0	1.457 88	4.42
INV	– 0.004 216 3	– 1.417 66	0.094 499 9	0.957 54	1.95
FUN	– 0.004 583 1	– 1.219 48	– 0.133 072 0	– 1.067 05	2.42

Estimates of the regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon^*_{p,t}$ for 7 mutual funds in the period 91-03-01 to 93-02-15, using the BVL index

APPENDIX 4

Calculation of r_s between the ranking of funds according to the two indexes: BVL and BTA (for both subperiods)

First subperiod

BTA			BVL		
Funds	α	Rank	Funds	α	Rank
VAL	– 0,001 414 8	1	VAL	– 0,001 475 8	1
INV	– 0,002 605 8	2	INV	– 0,001 667 9	2
UNI	– 0,002 058 5	3	UNI	– 0,002 109 0	3
FUN	– 0,002 246 0	4	FUN	– 0,002 277 4	4
MUL	– 0,002 810 8	5	FIP	– 0,002 983 7	5
FIP	– 0,003 006 0	6	MUL	– 0,002 993 7	6
PRI	– 0,005 721 7	7	PRI	– 0,005 933 8	7

	Rank BTA — X_i	Rank BVL — Y_i	X_i-Y_i	$(X_i$-$Y_i)2$
VAL	1	1	0	0
INV	2	2	0	0
UNI	3	3	0	0
FUN	4	4	0	0
MUL	5	6	– 1	1
FIP	6	5	1	1
PRI	7	7	0	0

Total = 2.

r_s = 0.964 285 714.

Second subperiod

BTA			BVL		
Funds	α	Rank	Funds	α	Rank
UNI	– 0,001 060 2	1	UNI	– 0,000 753 4	1
PRI	– 0,001 140 2	2	PRI	– 0,000 916 8	2
VAL	– 0,002 209 5	3	VAL	– 0,002 283 4	3
FIP	– 0,002 609 1	4	FIP	– 0,002 720 4	4
MUL	– 0,004 009 6	5	INV	– 0,003 989 9	5
INV	– 0,004 216 3	6	MUL	– 0,004 030 3	6
FUN	– 0,004 583 1	7	FUN	– 0,004 225 2	7

	Rank BTA — X_i	Rank BVL — Y_i	X_i-Y_i	$(X_i$-$Y_i)2$
UNI	1	1	0	0
PRI	2	2	0	0
VAL	3	3	0	0
FIP	4	4	0	0
MUL	5	6	– 1	1
INV	6	5	1	1
FUN	7	7	0	0

Total = 2.

r_s = 0.964 285 714.

APPENDIX 5

Description of White's (1980) method for correction of heteroscedasticity

1) From the original model: $(R_{p,t} - R_{f,t}) = \alpha_p + \beta_p(R_{m,t} - R_{f,t}) + \varepsilon_{p,t}$, obtain estimates of α_p e β_p.

2) Compute the respective residuals: $\hat{\varepsilon}_{p,t} = (R_{p,t} - R_{f,t}) - \hat{\alpha}_p - \hat{\beta}_p(R_{m,t} - R_{f,t})$ and square them.

3) Regress $\hat{\varepsilon}^2_{p,t}$ against a constant, all the previous variables as well as their squares and cross-products. From this regression save their predicted values as $\hat{\sigma}^2_t$, which is an estimate of their variances: σ^2_t. If any of these values is not positive, take the logarithm of $\hat{\varepsilon}^2_{p,t}$ and regress ln $(\hat{\varepsilon}^2_{p,t})$ against the variables just mentioned above. Then, from the antilog of the predicted values of ln $(\hat{\varepsilon}^2_{p,t})$ we will obtain $\hat{\sigma}^2_t$, which will be positive.

4) Set the weights $w_{p,t}$ to $1/\sqrt{\hat{\sigma}^2_t}$ and multiply each variable in the original model by $w_{p,t}$ (including the constant). Then obtain estimates of this new regression:

$$(R_{p,t} - R_{f,t})^* = \alpha_p w_{p,t} + \beta_p(R_{m,t} - R_{f,t})^* + \varepsilon^*_{p,t}$$

where:

$$(R_{p,t} - R_{f,t})^* = w_{p,t}\,(R_{p,t} - R_{f,t})$$

$$(R_{m,t} - R_{f,t})^* = w_{p,t}\,(R_{m,t} - R_{f,t})$$

The weighted least squares (WLS) thus obtained are consistent and asymptotically efficient and so are the estimated variances and covariances of the estimates.

APPENDIX 6

Estimates of Jensen's measure ($\hat{\alpha}_p$) and the systematic risk measure ($\hat{\beta}_p$) corrected for heteroscedasticity

Funds	$\hat{\alpha}_p$	*t*-stat.	$\hat{\beta}_p$	*t*-stat.	R^2 (percentage)
FIP	– 0.003 118	– 1.746 20*	0.140 590 0	2.903 80*	12.29
FUN	– 0.002 601	– 1.182 00	0.001 755 0	0.049 70	1.51
INV	– 0.003 012	– 1.815 60*	0.167 667 0	3.242 30*	11.83
MUL	– 0.003 635	– 1.679 20*	0.143 270 0	1.858 00*	7.56
PRI	– 0.003 909	– 2.055 80*	0.344 911 0	5.122 50*	29.77
UNI	– 0.002 026	– 1.568 50	0.360 422 0	7.744 20*	48.79
VAL	– 0.002 924	– 2.187 60*	0.025 560 0	0.705 80	11.68

* Statistically significant at 0.05 level

Estimates of the regression $(R_{p,t} — R_{f,t})^* = \alpha_p w_{p,t} + \beta_p(R_{m,t} — R_{f,t})^* + \varepsilon^*_{p,t}$ for 7 mutual funds in the period 89-03-15 to 93-02-15, using the BVL index.

Funds	$\hat{\alpha}_p$	*t*-stat.	$\hat{\beta}_p$	*t*-stat.	R^2 (percentage)
FIP	– 0.002 789	– 1.299 60	0.151 613 0	2.506 60*	16.28
FUN	– 0.002 818	– 0.094 51	– 0.007 890 0	– 0.127 10	1.43
INV	– 0.003 252	– 1.723 30*	0.086 487 0	1.557 10	3.90
MUL	– 0.003 429	– 1.742 80*	0.136 751 0	2.541 10*	9.10
PRI	– 0.003 465	– 1.803 00*	0.330 350 0	5.387 00*	31.10
UNI	– 0.001 962	– 1.385 80	0.312 416 0	7.801 60*	52.24
VAL	– 0.002 329	– 1.692 10	0.055 067 0	1.532 00	12.88

* Statistically significant at 0.05 level

Estimates of the regression $(R_{p,t} — R_{f,t})^* = \alpha_p w_{p,t} + \beta_p(R_{m,t} — R_{f,t})^* + \varepsilon^*_{p,t}$ for 7 mutual funds in the period 89-03-15 to 93-02-15, using the BTA index.

APPENDIX 7

Calculation of r_s between the ranking of funds with and without correction for heteroscedasticity

IBVL

Not corrected			Corrected for heteroscedasticity		
Funds	α	Rank	Funds	α	Rank
VAL	– 0,001 657 0	1	VAL	– 0,002 924 0	3
UNI	– 0,002 346 7	2	UNI	– 0,002 026 0	1
FUN	– 0,002 797 1	3	FUN	– 0,002 601 0	2
INV	– 0,002 885 2	4	INV	– 0,003 012 0	4
FIP	– 0,003 172 5	5	FIP	– 0,003 118 0	5
MUL	– 0,003 711 0	6	MUL	– 0,003 635 0	6
PRI	– 0,004 218 2	7	PRI	– 0,003 909 0	7

	Rank no cor. — *Xi*	Rank cor — *Yi*	*Xi-Yi*	(*Xi-Yi*)2
VAL	1	3	– 2	4
UNI	2	1	1	1
FUN	3	2	1	1
INV	4	4	0	0
FIP	5	5	0	0
MUL	6	6	0	0
PRI	7	7	0	0

Total = 6.

r_s = 0.892 857 143.

IBTA

Not corrected			Corrected for heteroscedasticity		
Funds	α	Rank	Funds	α	Rank
VAL	– 0,001 570 8	1	VAL	– 0,002 329 0	2
UNI	– 0,002 080 4	2	UNI	– 0,001 962 0	1
INV	– 0,002 807 6	3	INV	– 0,003 252 0	5
FUN	– 0,002 928 7	4	FUN	– 0,002 812 0	4
FIP	– 0,003 038 7	5	FIP	– 0,002 789 0	3
MUL	– 0,003 449 5	6	MUL	– 0,003 429 0	6
PRI	– 0,003 837 3	7	PRI	– 0,003 465 0	7

	Rank no cor. — *Xi*	Rank cor — *Yi*	*Xi-Yi*	(*Xi-Yi*)2
VAL	1	2	– 1	1
UNI	2	1	1	1
INV	3	5	– 2	4
FUN	4	4	0	0
FIP	5	3	2	4
MUL	6	6	0	0
PRI	7	7	0	0

Total = 10.

r_s = 0.821 428 571.

APPENDIX 8

Robust regression

In order to deal with potential outliers, we followed here the approach suggested by Rousseeuw (1984) and later developed by Rousseeuw and Zomeren (1990).

In practice, robust regression is carried out by firstly computing the *least median of squares* (LMS), which corresponds to:

$$\text{Minimize median } r^2_i \quad i = 1,...,N$$

where r_i is the residual of the i^{th} observation.

Afterwards, the outliers can be identified as those points that lie faraway from this robust fit, i. e., points with large positive or large negative residuals. However, in general, the Y_i (and hence the residuals) may be in any unit of measurement so, in order to decide if a residual r_i is «large», we need to compare it to an estimate of the error scale. Of course, this scale estimate has to be robust itself so that it should depend only on the «good» data and does not get blown up by outliers. For the least median of squares (LMS) one can use the following procedure [Rousseeuw and Zomeren (1990)]: we first calculate a preliminary scale estimate S^0 based on the value of the objective function and multiplied by a finite sample correction factor C as follows:

$$S^0 = C\sqrt{\text{median } r_i^2}$$

where:

r_i is the residual of case i with respect to the LMS fit; and

$$C = 1.4826[1 + 5/(N - p)]$$

is the above mentioned correction factor, which depends upon the number of observations N as well as the number of parameters p. The factor 1.4826 was chosen in order to achieve consistent Gaussian error distributions [Rousseeuw (1984)].

Next, with this *preliminary* scale estimate, the standardized residuals r_i/S^0 are computed and used to determine the weight w_i for the i^{th} observation as follows *:

$$W_i = \begin{cases} 1 & \text{if } \left|r_i / S^0\right| \le 2{,}5 \\ 0 & \text{if } \left|r_i / S^0\right| > 2{,}5 \end{cases}$$

The next step is to calculate the *final* scale estimate for the LMS regression which is given by:

$$\sigma^* = \sqrt{\left(\sum_{i=1}^{N} w_i * r_i^2\right) \Big/ \left(\sum_{i=1}^{N} w_i - p\right)}$$

The advantage of this formula is that outliers do not influence the scale estimate any more. Now, if the standardized residual $|\, y_i/\sigma^* \,|$ is large (> 2,5), observation i will be disregarded because it is considered as an outlier.

Finally, in order to improve on the crude LMS and to obtain standard quantities like t-values, confidence intervals and the like, we can use the so-called *reweighted least squares* (RLS) regression. This corresponds to minimizing the sum of squared residuals multiplied by a weight w_i:

$$\text{Minimize } \sum_{i=1}^{N} w_i * r_i^2$$

The weights are determined as previously but with the final scale estimate σ^* instead of S^0. The effect of the weights, which can only take the values of 0 or 1, is identical to deleting the cases for which w_i equals zero. Therefore RLS can be seen as ordinary LS on a «reduced» data set consisting of only those observations that received a nonzero weight. Because this «reduced» data set does not contain regression outliers any more, the statistics (and inferences) are more trustworthy than those associated with LS on the whole data set.

* The bound 2,5 is, of course, arbitrary but quite reasonable because in a Gaussian situation there will be very few residuals larger than 2,5. Instead of a «hard» rejection of outliers as we did for determining the weights w_i, one could also apply «smooth» rejection, for instance by using continuous functions of $|\, y_i/S^0 \,|$, thereby allowing for a region of doubt (e. g., points with $2 \le |\, y_i/S^0 \,| \le 3$ could be given weights between 1 and 0).

Bibliography

ANG, S.; CHUA, J. H. (1979) — «Composite measures for the evaluation of investment performance», *Journal of Financial and Quantitative Analysis*, vol. XIV, no. 2, June, pp. 361-384.

ARMADA, M. R. (1992) — «On the investigation of timing and selectivity in portfolio management», PhD dissertation, Manchester Business School.

BLUME, M. (1971) — «On the assessment of risk», *The Journal of Finance,* vol. XXVI, no. 1, March, pp. 1-10.

—— (1975) — «Betas and their regression tendencies», *The Journal of Finance*, vol. XXX, no. 3, June, pp. 785-795.

BROWN, K.; BROWN, G. D. (1987) — «Does the composition of the market portfolio really matter?», *The Journal of Portfolio Management,* Winter, pp. 26-32.

ELTON, E. J.; GRUBER, M. J. (1991) — «Differential information and timing ability», *Journal of Banking and Finance,* no. 15, pp.117-131.

FABOZZI, F.; FRANCIS, J. C. (1978) — «Beta as a random coefficient», *Journal of Financial and Quantitative Analysis,* vol. 13, no. 1, March, pp. 101-116.

—— (1979) — «Mutual fund systematic risk for bull and bear markets: an empirical examination», *The Journal of Finance,* vol. XXXIV, no. 5, December, pp. 1243-1250.

—— (1980) — «Stability of mutual fund systematic risk statistic», *Journal of Business Research,* vol. 8, no. 2, June, pp. 263-275.

FAMA, E. (1972) — «Components of investment performance», *The Journal of Finance,* vol. XXVII, no. 3, June, pp. 551-567.

FERSON, W. E.; SCHADT, R. W. (1993) — «Measuring fund strategy and performance in changing economic conditions», paper presented at the 20th Annual Meeting organized by the European Finance Association (Copenhagen).

FIELITZ, B. D.; GREENE, M. T. (1980) — «Shortcomings in portfolio evaluation via M. P. T.», *Journal of Portfolio Management,* Summer, pp. 13-19.

FRIEND, I.; BLUME, M. (1970) — «Measurement of portfolio performance under uncertainty», *The American Economic Review,* vol. LX, September, pp. 561-575.

GRINBLATT, M.; TITMAN, S. (1993) — «Performance measurement without benchmarks: an examination of mutual fund returns», *Journal of Business*, vol. 66, no. 1, pp. 47-68.

HADAR, J.; RUSSELL, W. R. (1969) — «Rules for ordering uncertain prospects», *The American Economic Review,* vol. XLIX, March, pp. 47-68.

HANOCH, G.; LEVY, H. (1969) — «The efficiency analysis of choice involving risk», *Review of Economic Studies*, vol. XXXVI, no. 3, July, pp. 335-346.

HENRIKSSON, R. D.; MERTON, R. C. (1981) — «On market timing and investment performance. II — Statistical procedures for evaluating forecasting skills» — *Journal of Business,* vol. 54, no. 4, pp. 513-533.

JEAN, W. H. (1971) — «The extension of portfolio analysis to three or more parameters», *Journal of Financial and Quantitative Analysis*, vol. VI, January, pp. 505-515.

—— (1978) — «Completely monotonic risk aversion stochastic dominance», The University of Alabama, working paper.

JENSEN, M. C. (1968) — «The performance of mutual funds in the period 1945-1964», *The Journal of Finance,* vol. 23, June, pp. 389-416.

—— (1972) — «Optimal utilization of market forecasts and the evaluation of investment performance», in *Mathematical Models in Investment and Finance,* edited by Szego and Shell, North-Holland Press, pp. 310-335.

KLEMKOSKY, R. (1973) — «The bias in composite performance measures», *Journal of Financial and Quantitative Analysis,* vol. VIII, June, pp. 505-514.

KLEMKOSKY, R.; MANESS, T. (1978) — «The predictability of real portfolio risk levels», *Journal of Finance,* vol. 33, May, pp. 631-639.

KON, S. J. (1983) — «The market-timing performance of mutual fund managers», *Journal of Business,* vol. 54, no. 3, pp. 323-347.

KON, S. J.; JEN, F. C. (1978) — «Estimation of the time-varying systematic risk and performance for mutual fund portfolios: an application of switching regression», *The Journal of Finance,* vol. XXXIII, no. 2, May, pp. 457-475.

—— (1979) — «The investment performance of mutual funds: an empirical investigation of timing, selectivity and market efficiency», *Journal of Business,* vol. 52, no. 2, April, pp. 263-289.

LEE, C. F.; RAHMAN, S. (1990) — «Market timing, selectivity and mutual fund performance: an empirical investigation», *Journal of Business,* vol. 63, no. 2, pp. 261-278.

LEVY, H. (1981) — «The CAPM and the investment horizon», *Journal of Portfolio Management,* Winter, pp. 32-40.

—— (1984) — «Measuring risk and performance over alternative investment horizons», *Financial Analysts Journal,* March-April, pp. 61-68.

LEVY, R. (1971) — «On the short-term stationary of beta coefficients», *Financial Analysts Journal,* November-December, pp. 55-62.

MERTON, R. C. (1980) — On estimating the expected return on the market», *Journal of Financial Economics*, 8, pp. 323-361.

—— (1981) — «On market timing and investment performance. I — An equilibrium theory of value for market forecasts», *Journal of Business*, vol. 54, no. 3, pp. 363-406.

PFLEIDERER, P.; BHATTACHARYA, S. (1983) — A note on performance evaluation», technical report no. 714, October, Stanford University, Graduate School of Business.

QUIRK, J. P.; SAPOSNIK, R. (1962) — «Admissibility and measurable utility functions», *Review of Economic Studies,* vol. XXXI, no. 79, February-March, pp. 140-146.

ROLL, R. (1977) — «A critique of the asset pricing theory's tests. Part I — On past and potential testability of the theory», *Journal of Financial Economics*, vol. 4, no. 2, March, pp. 129-176.

—— (1978) — «Ambiguity when performance is measured by the securities market line», *The Journal of Finance*, vol. XXXIII, no. 4, September, pp. 1051-1064.

—— (1979) — «A reply to Mayers and Rice (1979)», *Journal of Financial Economics*, 7, pp. 391-400.

—— (1980) — «Performance evaluation and benchmark errors (I)», *The Journal of Portfolio Management*, 6, Summer, pp. 5-12.

—— (1981) — «Performance evaluation and benchmark errors (II)», *The Journal of Portfolio Management,* Winter, pp. 17-22.
ROSS, S. A. (1976) — «The arbitrage theory of capital asset pricing», *Journal of Economic Theory*, vol. 13, December, pp. 341-360.
—— (1977) — Risk, return and arbitrage», in *Risk and Return in Finance,* Friend & Bicksler Ed., Cambridge, MA, pp. 189-218.
ROUSSEEAUW, P. J. (1984) — «Least median of squares regression», *American Statistical Association Journal*, vol. 79, no. 388, December, pp. 871-880.
ROUSSEEUW, P. J.; ZOMERAN, B. C. V. (1990) — «Unmasking multivariate outliers and leverage points», *American Statistical Association Journal*, vol. 85, no. 411, September, pp. 633-639.
SHARPE, W. F. (1966) — «Mutual fund performance», *Journal of Business,* vol. 39, no. 1, pp. 119-138.
TREYNOR, J. L. (1965) — «How to rate management of investment funds», *Harvard Business Review,* vol. 43, no. 1, January--February, pp. 63-75.
TREYNOR, J. L.; MAZUY, K. K. (1996) — «Can mutual funds outguess the market?», *Harvard Business Review,* July-August, pp. 131-136.
WHITE, H. (1980) — «A heteroscedasticity-consistent covariance matrix estimation and a direct test for heteroscedasticity», *Econometrica,* vol. 48, no. 4, May, pp. 817-838.
WHITMORE, G. A. (1970) — «Third-degree stochastic dominance», *The American Economic Review*, vol. LX, no. 3, June, pp. 457-459.
WILSON, J. W.; JONES, C. P. (1981) — «The relationship between performance and risk: whence the bias», *The Journal of Financial Research*, vol. IV, no. 2, Summer, pp. 103-107.