MARIA DO CÉU RIBEIRO CORTEZ

ON THE PERSISTENCE OF MUTUAL FUND PERFORMANCE IN SMALL MARKETS

Ph.D. THESIS

UNIVERSIDADE DO MINHO

1998

MARIA DO CÉU RIBEIRO CORTEZ

ON THE PERSISTENCE OF MUTUAL FUND PERFORMANCE IN SMALL MARKETS

THESIS SUBMITTED TO THE *UNIVERSIDADE DO MINHO*
FOR THE DEGREE OF PHILOSOPHY DOCTOR
IN BUSINESS ADMINISTRATION

SUPERVISOR:
PROFESSOR DOUTOR MANUEL JOSÉ DA ROCHA ARMADA

JOINT SUPERVISOR:
PROFESSOR DOUTOR DEAN ALLEN PAXSON

1998

"Nothing in the world can take the place of persistence"

Calvin Coolidge

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

RESUMO

A persistência do desempenho dos gestores de investimentos tem assumido um papel de destaque na literatura financeira, revestido-se de grande importância não só ao nível da investigação académica, mas também ao nível do processo de tomada de decisão do investidor. Estudos recentes sugerem que o desempenho futuro de fundos de investimento é previsível a partir do seu desempenho passado, isto é, fundos com um desempenho superior/inferior tenderão a manter esse bom/mau desempenho no futuro. Esta pesquisa aborda a persistência do desempenho de fundos de investimento num contexto de mercado de pequena dimensão (mercado de fundos de acções em Portugal).

Neste âmbito, procedemos à identificação de alguns dos problemas decorrentes da avaliação do desempenho dos fundos num contexto de amostra reduzida. Seguindo a metodologia baseada em tabelas de contingência, apresentamos, comparamos e discutimos os vários critérios de avaliação da persistência do desempenho em termos de significância estatística ajustada para o enviesamento decorrente da dimensão da amostra.

Os resultados da análise efectuada às rendibilidades dos fundos portugueses evidenciam, em geral e de forma significativa, persistência do desempenho (numa base trimestral). Contudo, este fenómeno desaparece quando as rendibilidades são ajustadas a várias dimensões do risco. A investigação também revela evidência significativa de persistência do risco. Podemos ainda observar que, para maiores e menores intervalos de tempo, a persistência do desempenho em termos globais desaparece, embora alguns fundos, individualmente, evidenciem um desempenho consistentemente superior/inferior.

ABSTRACT

The issue of persistence in fund performance is a major topic of debate in the finance literature, assuming great importance not only in terms of academic research but also in terms of practical investors' decision making. Recent evidence suggests that future performance is predictable from past performance, that is, funds with superior (inferior) performance in the past are likely to remain good (bad) performers in the future. This research addresses the persistence of mutual fund performance in a small market (the Portuguese equity fund market).

We identify some of the problems in evaluating fund persistence in the context of limited sample size and using the peer group median as benchmark for contingency table analysis of performance persistence. The criteria for assessing performance persistence based on the contingency table methodology of repeated winners and losers are presented in terms of significance statistics, adjusted for small sample bias. The appropriateness of each criteria under different circumstances is also discussed.

The analysis of the returns of all Portuguese domestic equity funds, since a representative number was established, shows significant performance persistence (on a quarterly basis). The persistence, however, disappears when the returns are controlled for the various dimensions of risk. We also have documented significant risk persistence. Furthermore, for more or less frequent intervals of measurement, the industry persistence is rejected, although individual funds exhibit superior/inferior performance.

ACKNOWLEDGMENTS

It is a pleasure to be able to express my acknowledgments to all those who, in many ways, through encouraging words or helpful suggestions, made this research possible.

Firstly, I would like to express my deepest gratitude to my supervisor Professor Manuel José da Rocha Armada, for his continuous encouragement and assistance. His PERSISTENT enthusiasm, support and dedication have not only helped me to develop research skills but, above all, have influenced my academic trajectory in this area of research, always constituting a motivation for positive PERFORMANCES.

I also wish to express my greatest recognition to my joint supervisor, Professor Dean Paxson, for his willingness to give his time and criticize every manuscript. Always encouraging and challenging me, his suggestions and insightful views were an invaluable source of help throughout this research.

I also wish to express a special gratitude to the encouragement given by colleagues and friends. In particular to Carlos Machado-Santos, I would like to thank him not only for his helpful comments and assistance on some parts of the thesis, but also for sharing difficulties and uncertainties in the most arduous times. To Nelson Areal, I thank his kind availability to assist me on computer intensive methods of hypothesis testing as well as his valuable observations and friendly words of support. To my colleagues at the finance group and at the University, I am glad to work with such friendly and helpful peers.

I am also very grateful to those who have provided the data used in this research: Dr. António Santiago and Carla Miranda from the Unit Trust Association (APFIN), Dra. Bernardina Ribeiro and Dr. Diamantino Leite from the Oporto Derivatives Exchange (BDP) and Dr. Luis Oliveira from the Lisbon Stock Exchange (BVL). Grateful thanks must also go to many managers of the mutual fund companies. Without the help of these entities, it would have been very difficult to collect the data.

I would also like to thank the *Universidade do Minho* for having provided me with a leave of absence during my doctoral research and also the School of Economics and Management (EEG), the Research Centre in Economics and Management (CEEG) and the Department of Business and Public Administration (DGAP) for financial support.

Finally, I can never thank enough my family for the continuing support always received. To my parents, brothers and specially to my sister Susana for baby-sitting for me so many times. To my husband Armando, whose presence, patience and comprehension has given me the necessary inspiration and strength to pursue this research and, last but not least, to my children Rafaela and Zé Pedro, for filling my life with (turbulence and) happiness.

To my parents, to Armando, Rafaela and Zé Pedro

THE AUTHOR

Maria do Céu Ribeiro Cortez obtained a Licentiate degree (1989) in Business Administration from the *Universidade do Minho*, Braga, Portugal. She subsequently joined the Management Department of the School of Economics and Management of the *Universidade do Minho*, as an Assistant Lecturer, where she taught several courses, in particular Mathematics of Finance. In 1994 she was awarded the MA degree in Management from the *Instituto Superior de Economia e Gestão, Universidade Técnica de Lisboa*, Lisbon. In October 1995 she was provided with a leave of absence from the University, starting her doctoral research programme.

CHAPTER 1

INTRODUCTION

1.1. GENERAL AREA, PURPOSE AND JUSTIFICATION OF THE RESEARCH[1]

The diversification and professional management services provided by mutual funds are frequent reasons cited for this type of financial intermediary having become a major vehicle through which individuals invest in capital markets. The performance of mutual funds, as a consequence, has always been a significant issue. The evaluation of portfolio performance is an integral part of the investment management process, being the ultimate feedback for managers' and investors' decision making. Moreover, the issue of assessing whether fund managers add value is a challenging one, since it is closely related to questions (not easily answered) about market efficiency and information dissemination in capital markets.

Over the last decades, portfolio evaluation has evolved considerably: the sophistication of markets, instruments and agents and the proliferation of performance data and techniques to analyze these data have raised the competitiveness among investment managers and reduced the opportunity for value added. Notwithstanding, the problem of evaluating the performance of portfolio managers remains largely unsolved for over than three decades, and is still an evolving subject.

The early traditional measures of TREYNOR [1965], SHARPE [1966] and JENSEN [1968] go back to the beginning of asset pricing

---

1 Equivalently, this section could be entitled *"Why bother with performance?"*, citing the question posed by ELTON and GRUBER [1997, p.26].

theory and generated an intense debate in the financial community. Issues such as the appropriateness of benchmarks, the use of beta as the risk measure and the correlation of performance measures with risk questioned the validity of these measures (and the underlying theoretical frameworks, Capital Market Theory and Capital Asset Pricing Model) and led to the development of alternative approaches incorporating multidimensional benchmarks (such as the Arbitrage Pricing Theory). Academic research on mutual fund performance includes such diverse topics as market timing, differential information, flows into and out of mutual funds and the categorization of investment styles. More recently, a matter of particular interest that has attracted the attention of researchers is the degree of persistence of performance over time. In a perfectly efficient market superior performance is the result of chance and does not reflect managers' skill. If this hypothesis were true, although some fund managers might outperform a passive strategy and others underperform it, we would expect performance to be random over time. However, a manager with superior investment abilities might be expected to continue to exhibit above-average abilities in subsequent periods. Hence, if fund managers do have sustainable superior skills, it should be reflected in persistence of performance.

While most research on mutual funds has been devoted to testing whether funds exhibit abnormal performance, recent studies published in the nineties (e.g. HENDRICKS, PATEL and

ZECKHAUSER [1993] and BROWN and GOETZMANN [1995]) have directly examined the persistence of performance and found evidence supporting the idea that past performance is related to future performance. However, the overall results on performance persistence are somewhat mixed, with some studies providing evidence of persistence and other attributing this phenomenon to biases present in the data which would overstate the persistence effect.

As reflected in a recent burst of studies (reviewed in chapter 2), there is an ongoing debate in progress on whether mutual funds can consistently outperform the market or not. This type of evidence would have implications at various levels: not only it would be inconsistent with market efficiency, but it would have practical value to investors, since it suggests that they can realize abnormal returns by investing in the recent top performing funds.

On the other hand, the Portuguese capital market is clearly in need of research. Unlike the extensively studied U.S and U.K. markets, characterized by a long history of data, the Portuguese mutual fund industry is a very recent one.[2] However, in the last few years it has experienced a significant growth such as to become an important component of the market and an important alternative investment product for individual investors. In fact, as of December 1993, the total assets invested in mutual funds represented approximately 17 percent of

[2] The first equity mutual fund was introduced in 1986.

the value of bank deposits and 12 percent of the Gross National Product.[3] By December 1997, these figures have risen to 30 and 22 percent, respectively.

In this context, the purpose of this investigation is to study the performance persistence of a sample of Portuguese stock funds. This is a challenging task since a priori we face a number of limitations related with the small size of the sample. Among the methodologies used to assess performance persistence (described in chapter 2), and for reasons exposed in chapter 3, we considered most appropriate to pursue the study of performance persistence following the methodology of two-way contingency tables of winners and losers. Any of the methodologies (and in our particular case, the methodology based on contingency tables, described in chapter 3) used to assess persistence involve tests which are appropriate in the context of large samples but might generate biased results in the case of limited sample size. Hence, we will contribute to the mutual fund performance literature by identifying some of the problems in quantifying performance persistence with a small sample and correcting for the bias. The two-way contingency table methodology of repeat winners and losers is presented in terms of significance statistics adjusted for the small sample bias problem. Moreover, our approach is innovative in the sense that it constitutes a multifaceted examination of performance

---

[3] All statistics presented here (and forward) relative to mutual funds exclude real estate investment trusts. All statistics have been obtained from the Unit Trust Association (APFIN) reports.

persistence, based on the comparison of the essentials points of each performance persistence criteria. Also, and to the best of our knowledge, ours is the first study to consider the persistence of risk characteristics of mutual funds. Furthermore, we also study the persistence of fund performance over various measurement intervals - term structure of performance persistence - and analyze the economic significance of investing in performance persistence strategies (over alternative investment horizons).

1.2. PLAN OF PRESENTATION

This research is organized into six chapters. In chapter 2 we review and discuss the literature. The chapter presents an overview of the performance evaluation methodologies (and the underlying theoretical frameworks) and main issues currently in debate, in particular the question of persistence of performance. In chapter 3 we present the methodology of analysis of performance persistence within the context of contingency tables and describe the various forms of tests applied to assess performance persistence. In chapter 4 we analyze the performance and persistence of our sample of funds. Having described the database, we report and discuss the empirical results relative to the performance persistence of returns (adjusted and unadjusted for risk) and of risk characteristics. We also address the term structure of performance persistence and analyze the performance persistence of individual funds. In all cases the adjustments required due to limited sample size were carried out and compared with the significance of the asymptotic tests. The analysis allowed for reflection on the appropriateness of each performance persistence criteria.

In chapter 5 the analysis is complemented by examining the implications of strategies that exploit the persistence phenomenon. Finally, in chapter 6 we summarize the main results and conclusions and suggest possible lines for future research.

CHAPTER 2

REVISION AND DISCUSSION OF PRIOR RESEARCH

2.1. INTRODUCTION

Performance evaluation is mainly concerned with two issues: (1) determining whether the portfolio manager added value by outperforming an established benchmark and (2) identifying the sources of such performance, trying to determine whether it was due to skill or luck. However, there are many difficulties associated with carrying out these tasks. Accordingly, this chapter will present and discuss the various approaches for evaluating portfolio performance and the problems that may arise with their use.

The first studies on mutual fund performance evaluation were based solely on the rate of return achieved by a portfolio over the holding period relative to the returns of other portfolios having similar objectives or to some general market index.[4]

Developed by MARKOWITZ [1952], Modern Portfolio Theory (MPT) describes the principles for constructing efficient portfolios and selecting the optimal portfolio based on two key parameters: risk and return. Furthermore, Markowitz demonstrated the importance of diversifying a portfolio, the goal of which is to reduce the portfolio's risk without sacrificing return. Markowitz's analysis was further expanded by TOBIN [1958], who explored the role of the risk-free asset in the portfolio.

[4] For a review of performance studies based on return information only see ARMADA [1992, p.24-25].

Based on this normative mean-variance framework, SHARPE [1964], LINTNER [1965] and MOSSIN [1966] developed the Capital Asset Pricing Model (CAPM), which is a positive theory that describes the relationship between risk and return, or equivalently, it is an equilibrium model for the pricing of capital assets in a competitive market.

A fundamental principle resulting from MPT / CAPM is that investors require compensation for bearing risk. It is acknowledged that in evaluating performance of portfolios, the effect of differential risk must be taken into account. Superior returns achieved by the portfolio manager simply by increasing the portfolio's risk cannot be considered skill or superior performance.

Although there is consensus about the need of including risk in performance evaluation, still controversial is what constitutes risk and how it should be measured. Indeed, "*there is no single universally applicable risk measure*" [BALZER, 1995, p.5]. In the portfolio selection context, having defined risk in terms of the variance of expected returns, MARKOWITZ [1959] suggested that the semivariance would be theoretically more robust.[5] Also, HARLOW [1991], MARMER and NG [1993], ROM and FERGUSON [1994a, 1994b], MERRIKEN [1994] and BALZER [1995] propose semivariance as a more appropriate measure of risk. These studies show that the theoretical assumptions required to justify variance as a measure

[5] However, in light of computational problems associated with calculating semivariance, Markowitz adopted variance as the risk measure in his analysis.

of risk are very restrictive, particularly in the context of investments that generate asymmetric return distributions.

On the other hand, the validity of the equilibrium models in which the risk measures are inserted constitutes another issue in debate. Alternative equilibrium models incorporating multidimensional benchmarks, such as the Arbitrage Pricing Theory, maintain that there is more than one source of risk affecting security returns. As a result of this debate, the financial literature has proposed different approaches for evaluating performance. The growth in academic research along with the application of more sophisticated econometric techniques added new dimensions to performance evaluation. Studies on mutual fund performance have investigated such topics as performance attribution (v.g. timing and selectivity), investment style and more recently, performance persistence. The ability of fund managers to generate consistently superior performance is perhaps one of the most interesting questions mutual fund researchers currently investigate, and constitutes the primary focus of this study. Academics and practitioners have recognized the importance of this issue and of its implications: if performance does persist, pertinent questions arise relative to market efficiency and its economic significance to investors. Recent studies (e.g.: GRINBLATT and TITMAN [1992], HENDRICKS, PATEL and ZECKHAUSER [1993], BROWN and GOETZMANN [1995], ELTON, GRUBER and BLAKE [1996a] and CARHART [1997]) have found evidence of performance predictability, that is, funds which perform well or poorly in the past will tend to continue to do so in the future. Despite this type of evidence, the persistence phenomenon remains a controversial and unresolved matter. There are reasons for such a controversy: risk-adjustment of returns, appropriateness of benchmarks, survivorship bias and, of course, the conflicting empirical results of previous studies.

Since the questions that may arise relatively to performance persistence are intimately related to questions of performance evaluation, we will review the literature of portfolio performance evaluation focusing on the main issues that are susceptible of being discussed in the context of performance persistence. In this chapter, and in the context of what was said immediately before, we will briefly review some of the most relevant methodologies for evaluating performance and their underlying theoretical structures, giving particular attention to the benchmark problem, which constitutes a major source of debate and criticisms in this area of research. Since most empirical evidence on mutual fund performance does not support the hypothesis that fund managers have forecasting skills (either at the micro and macro level), we will emphasize the role of investment style in performance evaluation, in particular the importance of evaluating performance relative to benchmarks that reflect the manager's investment style. Finally, and most importantly, we will discuss the topic of persistence of performance and present the state-of-the-art on this matter. Besides a brief description of the most relevant methodologies used to assess persistence, we will present the main evidence found in the literature and discuss the most pertinent issues outstanding.

2.2. PERFORMANCE MEASURES BASED ON CAPITAL MARKET THEORY

Three different but related measures of performance which explicitly incorporate both risk and return into a single composite measure have been developed from the Sharpe - Lintner - Mossin theory of asset prices in equilibrium, in this way allowing the comparison of portfolios with different risk policies. These so-called traditional measures of performance are briefly described next.

According to the CAPM, the expected return on a portfolio or security is positive and linearly related to the expected return on the market, as follows:

$$E\left[R_{p,t}\right] = R_{f,t} + \beta_p\left(E\left[R_{m,t}\right] - R_{f,t}\right) \qquad [2.1]$$

where:

$E\left[R_{f,t}\right]$ = Expected return of security or portfolio **p** over period **t**;

$R_{f,t}$ = Risk-free rate over period **t**;

$\beta_p = \frac{Cov_{p,m}}{\sigma_m^2} =$ Systematic risk measure of security or portfolio **p**;[6]

$E\left[R_{m,t}\right]$ = Expected return on the market portfolio over period **t**.

The performance measure suggested by JENSEN [1968] is based on this equilibrium relationship[7], which is an ex-ante model stated in terms of expected returns on the security (or portfolio) and expected returns on the market portfolio. Since these expectations are unobservable, JENSEN [1968] recasts [2.1] in terms of ex-post returns, and allows for detection of forecasting ability simply by not constraining the estimation regression to pass through the origin:

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p\left(R_{m,t} - R_{f,t}\right) + \varepsilon_{p,t} \qquad [2.2]$$

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

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

---

6 In a CAPM context, the total risk of a security can be decomposed into two components: the *non-systematic* risk (or *unique* risk), which is related to company-specific influences and can be eliminated by diversification and the *systematic* (or *market*) risk, which measures the sensitivity of the security to general movements in the market.

7 This equilibrium relationship between expected return and risk is termed Security Market Line (SML).

The intercept α_p is the measure of performance proposed by Jensen and can be interpreted as the return above (or below) the CAPM equilibrium return implied by its level of systematic risk. Hence, significant positive estimates of α_p suggest superior performance relative to the market. In this case, "*it represents the average incremental rate of return on the portfolio per unit of time which is due solely to the manager's ability to forecast security prices*" [JENSEN, 1968, p.394]. Conversely, significant (statistically) negative estimates of α_p indicate inferior performance relative to the market. From a random selection buy-and-hold policy is expected a zero intercept.

Unlike the previous, the techniques developed by TREYNOR [1965] and SHARPE [1966] are relative and not absolute measures of performance which, in essence, measure the excess return per unit of risk.[8] The measure of performance suggested by TREYNOR [1965], also known as the *reward-to-volatility ratio*, gives the excess return per unit of systematic risk:

$$T_p = \frac{\bar{R}_p - \bar{R}_f}{\beta_p} \qquad [2.3]$$

where:

T_p = Treynor's performance measure;

$\bar{R}_p$ = Average return on the portfolio over period **p**;

[8] The excess return is defined as the difference between the portfolio's return and the risk-free rate of return over the same evaluation period.

$\overline{R}_f$ = Average risk-free return over period **p**;

β_p = Systematic risk measure for the portfolio over period **p**.

Treynor's measure is closely related to Jensen's since it also uses the ex-post SML as a benchmark to portfolio evaluation. In this sense, they will always give the same assessment of a portfolio's performance relative to the market.

The measure of performance proposed by SHARPE [1966], unlike the two previous ones described above, is a measure of risk adjusted performance that uses a benchmark based on the ex-post Capital Market Line (CML).[9] Being so, the *reward-to-variability ratio* measures the excess returns per unit of total risk:

$$S_p = \frac{\overline{R}_p - \overline{R_f}}{\sigma_p} \qquad [2.4]$$

where:

S_p = Sharpe's performance measure;[10]

$\overline{R}_p$ and $\overline{R}_f$ as described previously;

σ_p = Standard deviation of the portfolio's returns over period **p**.

[9] The Capital Market Line consists of alternative combinations of risk and return obtainable by combining the market portfolio with riskfree borrowing or lending. Assuming homogeneous expectations and perfect markets, the CML represents the linear efficient set.

[10] This performance measure is also termed the "Sharpe ratio" (see SHARPE [1994]).

The above reviewed techniques for evaluating performance present major similarities. In fact, it can be demonstrated that, under certain market conditions, the three measures are approximately linear transformations of each other.[11] Also, these similarities can be further illustrated by looking at the results of various empirical studies, which show that the correlation coefficients between the rankings of portfolios by these three measures are very high.[12] Moreover, all of these measures assume the normality and stationarity of returns. The main difference concerns the risk measure used: the Jensen and Treynor measures use systematic risk, or beta, whereas Sharpe's ratio uses total risk as measured by the standard deviation of returns. In this context, the appropriate risk measure depends on the level of diversification of the portfolio. In the presence of a well-diversified portfolio (the unsystematic risk is close to zero), the variance of the portfolio's returns is totally explained by the variance of the market returns, then beta is the relevant measure of risk, and performance should be based on the SML. On the other hand, if the investor's portfolio is poorly diversified, the rankings could be quite different due to the large amount of unsystematic risk. To such an investor, standard deviation is the relevant measure of risk, and performance should be based on the Sharpe ratio.

---

11 See, for example, MOSES and CHENEY [1989, p.204].

12 See for example BOWER and WIPPERN [1969] and SMITH and TITO [1969].

2.3. PROBLEMS WITH PERFORMANCE EVALUATION

2.3.1. General Critiques

The classical studies of TREYNOR [1965], SHARPE [1966] and JENSEN [1968] reinforced the market efficiency paradigm that would dominate the finance literature for the next two decades. In fact, the large majority of the empirical results of these first-generation mutual fund papers failed to find evidence of consistently superior investment performance by professional fund managers and suggest that mutual funds do not generate rates of return enough to offset their expenses.

However, the efficacy of these composite measures in providing accurate inferences of performance has been questioned, as they have been subject to a number of criticisms and objections, widely reported and discussed in the literature. The following discussion briefly outlines some of the most relevant theoretical considerations and empirical findings on the basis of the controversy.

FRIEND and BLUME [1970] raised the issue of a "systematic bias" in these empirical measures of performance because of the observed tendency of the empirical measures to be inversely correlated with risk. Other studies (e.g., KLEMKOSKY [1973], ANG and CHUA [1979]) have shown that the performance measures are significantly correlated with portfolio risk measures, but document a positive relationship. Also, WILSON and JONES [1981] address this apparent conflict and CHEN and LEE [1981] investigate possible sources of the bias.

The effects of the interval of time used to calculate returns on estimates of performance measures have also been investigated in the literature. FIELITZ and GREENE [1980] and LEVY [1981, 1984] provide empirical evidence on the sensitivity of performance measures to variations in the time horizon for calculation of returns, and point out that the rankings of portfolios by performance may change when different investment horizons are used to calculate returns. More recently HANDA, KOTHARI and WASLEY [1993] and GUNTHORPE and LEVY [1994] demonstrate that changes in the return measurement interval can affect inferences about the risk/return relation.

2.3.2. The benchmark problem: Roll's critique

A major source of controversy regarding the evaluation of portfolio performance is that these techniques, in particular the JENSEN [1968] and TREYNOR [1965] measures of performance, require the use of a benchmark portfolio. As described previously, the CAPM is a general equilibrium model based on the existence of a market portfolio that is defined as the value-weighted portfolio of all investment assets. This market portfolio is considered the appropriate benchmark. However, since it cannot be observed, market indices are used as proxies for the market and as benchmarks for performance. It is precisely from the non-observability of the *true* market portfolio and the use of proxies for the market that stem ROLL's [1977, 1978, 1979, 1980, 1981] critiques to CAPM and CAPM-based performance measures. Following Roll's argument, if the market index is efficient (in Markowitz's sense) all portfolios will lie on the SML, invalidating any attempts for evaluating relative performance. If however, there are deviations from the SML, this would imply that the index was not mean-variance efficient, and the deviations can tell us nothing about the relative performance portfolios. Hence, Roll submits that any empirical test of the CAPM is really a joint test of the validity of the pricing model and the mean-variancy of the index. Given the difficulty of finding a *true* proxy for the market portfolio, performance evaluation with these methods will be sensitive to the benchmark choice. Moreover, benchmarks that are mean-variance inefficient will provide erroneous inferences. As Roll demonstrates, since different benchmarks can produce different estimates of betas, manager rankings could be completely reversed simply by choosing a different benchmark to represent the market. The ambiguity of the SML analysis results from the fact that "(...) *for every asset (or portfolio) judicious choice of the index can produce any desired measured performance (positive or negative) against the securities market line*" [ROLL, 1978, p. 1056].[13] Roll concludes that from these estimates of

[13] In this line of thought, FERGUSON [1980, 1986] considers the SML procedure for performance evaluation arbitrary, because the choice of the reference portfolio is also arbitrary.

performance we cannot infer about the manager's true performance, given that what is perceived as superior performance may be due simply to benchmark error. The sensitivity of performance evaluation measures to different benchmarks was also examined by DYBVIG and ROSS [1985a], BROWN and BROWN [1987], ZIMMERMANN and ZOGG-WETTER [1992] and FLETCHER [1995], who support Roll's assertion and essentially verified that the rankings of portfolios by performance are sensitive to the index used. In this line of research, LEE and JEN [1978] point out that the utilization of indices as proxies for the market portfolio can lead to measurement errors on the market return, which will affect estimates of the systematic risk of the portfolio. "*Hence, interpretation of the empirical results of CAPM should be done with extreme care*" [LEE and JEN, 1978, p. 309].

MAYERS and RICE [1979] contend that "*although there are potential problems (...) they are valid tests*" [MAYERS and RICE, 1979, p. 23] under certain conditions. In fact, they propose an information framework in which the appropriate index is efficient relative to the probabilities assessed by the market. This interpretation assumes an economy that is dominated by agents who share homogeneous beliefs and an individual with better assessments than the market. Since this individual has no weight in the economy, the assumptions of the CAPM are not violated. If the informed investor's probability beliefs are correct, he will earn a higher average return then the uninformed investor expects him to, and this will position him above the SML drawn by the uninformed investors. Although the index used was not efficient relative to the *true* joint probability density function, it is mean-variance efficient relative to the probability density function of returns assessed by the market.

Roll's criticisms are strengthened and expanded by DYBVIG and ROSS [1985a, 1985b], who examine the validity of the SML analysis under certain theoretical frameworks. On one side, assuming symmetric information and an inefficient index, DYBVIG and ROSS [1985a] show that the SML can be misleading, since both efficient and inefficient portfolios can plot above and below the SML. Abnormal returns relative to the SML may simply reflect the misspecification of the index and not superior performance. On the other hand, and contrary to MAYERS and RICE [1979], DYBVIG and ROSS [1985b] sustain that differential information disrupts the validity of the SML analysis, since it falls outside the domain of mean-variance analysis. Dybvig and Ross show that a manager with superior information can plot above, on, or below the SML, calling attention to the fragility of this procedure for evaluating performance. As these authors suggest: "*the mounting evidence against the validity of the SML analysis prompts a call for a new performance measurement technique*" [DYBVIG and ROSS, 1985b, p. 397].

The benchmark problem is even more pertinent if we consider an extensive body of empirical research contradicting CAPM's prediction that there is a positive and linear cross-sectional relation between expected returns and risk. Motivated by the

CAPM, early empirical studies on mutual fund performance made use of general stock indices as benchmarks. Yet, since the 1970's, a number of CAPM anomalies were uncovered, suggesting that there are variables other than beta that have power to explain stock returns. These include a price-earnings factor (BAZU [1977]), a firm-size factor (BANZ [1981]) and both a firm-size and book-to-market factor (FAMA and FRENCH [1992]). In fact, FAMA and FRENCH's [1992] study was considered "*the most damaging blow to the CAPM*" [GRUNDY and MALKIEL, 1996, p. 39] as they empirically show that beta and average return are not correlated. These results were challenged on several basis: for example, CHAN and LAKONISHOK [1993] call attention to the noise in realized returns which can affect the relationship between returns and betas. ROLL and ROSS [1994] emphasize that the weak relation between the CAPM risk/return relationship could be attributed to the mean/variance inefficiency of the benchmark portfolio. BLACK [1993a, 1993b] further argues that the Fama and French results might be due to data mining, while KOTHARI, SHANKEN and SLOAN [1995] support that they were influenced by survivorship bias. Anyhow, and in light of this evidence concerning the risk/return relationship, the use of CAPM-based benchmarks appear to be inappropriate, for they provide the possibility for managers to game the performance evaluation process.

In the next section, we will present and discuss an alternative paradigm for describing asset prices, that arose in an attempt to overcome some of CAPM's problems described above.

2.4. ARBITRAGE PRICING THEORY AND PERFORMANCE EVALUATION

As we have seen, CAPM predicts that only one type of systematic risk affects security returns. The recognition that a variety of factors may affect expected returns, coupled with concerns on the testability of the CAPM, led researchers to explore an alternative asset pricing theory: the Arbitrage Pricing Theory (APT), developed by ROSS [1976, 1977]. Ross argued that systematic risk need not be adequately represented by a single common factor (market risk), but instead it explicitly assumes that there are **k** common sources of covariation affecting security returns. These **k** factors constitute another potential benchmark to which performance can be measured.

The APT relies on a factor model of asset returns which postulates that the actual return on any asset is a linear combination of the expected return and a number of factors, plus an asset-specific random variable, as follows:

$$R_{p,t} = E\left[R_p\right] + \sum_{j=1}^{k} \beta_{p,j} F_{j,t} + \varepsilon_{p,t} \qquad [2.5]$$

where:

$R_{p,t}$ = Return on portfolio (or security) **p** over period **t**;

$E\left[R_p\right]$ = Expected return of portfolio (or security) **p**;

$\beta_{p,j}$ = Sensitivity of the return on portfolio (or security) **p** to movements in the $\mathbf{j}^{th}$ common factor (also termed factor loading);

$F_{j,t}$ = Value taken by the **j**th common factor over period **t**;

$\varepsilon_{p,t}$ = Residual error term, or idiosyncratic return on the ith asset, which is assumed to have the following properties:

$$E\left[\varepsilon_{p,t}\right] = 0 \;\; ; \;\; Var\left[\varepsilon_{p,t}\right] = \sigma^2_{\varepsilon_{p,t}} \;\; ; \;\; Cov\left[\varepsilon_{p,t}, F_{j,t}\right] = 0$$

The "no-arbitrage" (equilibrium) equation that arises from this return-generating process can be summarized as:

$$E\left[R_p\right] = R_f + \sum_{j=1}^{k} \beta_{p,j}\lambda_j \qquad [2.6]$$

where:

R_f = riskless rate of return (or the rate of return on the zero-systematic risk portfolio);

λ_j = risk premium for the **j**th factor.

Supporters of this model sustain that the APT framework presents major advantages over the CAPM. First, it makes much less restrictive assumptions regarding investors preferences towards risk and return. Second, it makes no assumption on the joint normal distribution of returns. Finally, it does not require the identification of the "true" market portfolio, thus being free of Roll's arguments on the testability of the theory.

For these reasons, APT is considered more appealing and flexible than CAPM.

Despite its attractive features though, empirical applications of the APT are particularly difficult to implement. The theory does not specify the number and identity of the multiple factors that affect security returns. Consequently, tests of APT require a strategy for measuring the common factors. One means of accomplishing this is through factor analysis. Factor analysis is a statistical procedure for determining a specific set of factors and factor loadings ($\beta_{p,j}$) such that the covariance of residual returns is as small as possible. A set of factors is extracted that can best describe the behaviour of a sample of assets. However, there are some problems itself with employing this methodology, since the mathematics of factor analysis is particularly complex in the context of large-scale covariance matrices. As a consequence, the use of small cross-sections can produce imprecise estimates of the common factors.[14] Also, there are different ways of carrying out factor analysis! In addition, it is not possible to be sure that one has captured all the relevant factors. DHRYMES, FRIEND and GULTENKIN [1984] found that the number of factors that appear significant is an increasing function of the size of the sample analyzed.[15] An alternative approach to testing APT is to prespecify a set of

---

[14] CHEN [1983] developed a procedure that would accommodate for large cross-sectional samples. However, this procedure has been criticized by DHRYMES, FRIEND, and GULTENKIN [1984].

[15] For an example of empirical application of factor analysis see LEHMANN and MODEST [1988]. An extensive examination of the literature on this methodology can be found in CONNOR and KORAJCZYK [1995].

observable variables as proxies for systematic factors. But even here there is no consensus on the appropriate factors that determine returns. Several researchers have investigated security returns and estimated that there are anywhere from three to five pervasive factors. CHEN, ROLL and ROSS [1986] specify ex-ante four macroeconomic variables as being the pervasive factors[16] and conclude that these prespecified factors provide a reasonable specification of the sources of systematic risk in the economy. Also BERRY, BURMEISTER and McELROY [1988] identify five similar factors.[17] Rather than specifying certain macroeconomic variables as being the pervasive factors, other authors have specified ex-ante sets of portfolios whose returns are assumed to be maximally correlated with the factors. For example, FAMA and FRENCH [1993] found stock returns to be related to three factors: a market factor, a size factor and a book-to-market equity factor. They also identify two bond-market factors, related to maturity and default risk. Fama and French conclude that "*at a minimum, our results show that five factors do a good job explaining (a) common variation in bond and stock returns and (b) the cross-section of average returns.(...) But the choice of factors, especially the size and book-to market factors, is motivated by empirical experience. Without a theory that specifies the exact*

[16] The four prespecified factors are: rate of growth in industrial production, inflation, spread between long-term and short-term interest rates and spread between low-grade and high grade bonds.

[17] The five factors are: the difference in returns of long-term corporate bonds and long-term government bonds plus a constant, the difference in returns on long-term government bonds and short term Treasury bills, a measure of unexpected deflation, a measure of growth of sales in the economy and the rate of return on a market index.

form of the state variables or common factors in returns, the choice of any particular version of the factors is somewhat arbitrary" [FAMA and FRENCH, 1993, p. 53].

Just as the CAPM, APT has been the theoretical basis for studies on the evaluation of managed portfolios. In particular, given the excess returns on factor portfolios, an abnormal performance (α) can be obtained as a multifactor or APT analogue of Jensen's measure. The study of LEHMANN and MODEST [1987] was among the first studies to apply APT to performance evaluation. CONNOR and KORAJCZYK [1991] also applied APT-based estimates of Jensen's measures. Their findings will be presented in section 2.5.

Of the two major equilibrium theories we have discussed, although CAPM is clear about the factor to be measured (the market), it is unmeasurable; on the other hand, APT postulates factors that could be easily measured, but neglects to identify them. Before proceeding, it appears convenient at this point to synthetize the relation between factor models, CAPM and APT.

A factor model describes the return-generating process that represents the behaviour of security returns. It is intended to identify the pervasive factors that affect securities' returns and to capture the sensitivity of security returns to the movements of those common factors or indices. One type of factor model is the market model, which assumes that there is only one factor - the return on a market index. On the other hand, multiple factor models assume that there is more than one pervasive factor that affects security returns. A factor model may represent a market that is efficient or inefficient. Therefore, it is not, per se, an equilibrium theory of security prices. CAPM and APT are both equilibrium theories which make a statement about securities' expected returns. Unlike the APT, the CAPM does not require that returns are generated by a factor model. However, this does not mean that it is inconsistent with a world in which returns are generated by a factor model. It is possible for the CAPM to hold even considering that returns are generated by a multiple factor model.[18] Since APT and CAPM are not necessarily inconsistent with each other, *"it is natural to consider a world in which (1) returns are generated by a factor model, (2) the remaining assumptions of APT hold, and (3) the assumptions of CAPM hold"* [SHARPE, 1984, p. 23].

2.5. EVIDENCE ON MUTUAL FUND PERFORMANCE - SINGLE PORTFOLIO BENCHMARKS VERSUS MULTIPLE PORTFOLIO BENCHMARKS

While the debate on the "true" model for asset pricing has not yet been resolved, most empirical studies on mutual fund performance have diverged on the use of single and multiple portfolio benchmarks.

Motivated by CAPM, the preponderance of the early literature on performance evaluation has concentrated on single-index

[18] It is the case that beta will be a linear combination of its sensitivities to the factors. In this context, SHARPE [1995, p. 336] and ELTON and GRUBER [1995, p.387] show that the APT with multiple factors is fully consistent with the Sharpe-Lintner-Mossin form of the CAPM.

measures of performance for many years. The mutual fund studies of TREYNOR [1965], SHARPE [1966] and JENSEN [1968] failed to find evidence of consistently superior investment performance. More recently IPPOLITO [1989], applying the methodology of these first-generation mutual fund studies and using comparable data found, for the 20-year period 1965-1985, positive average estimates of α even after accounting for transaction costs and expenses. Ippolito interprets his results as indicating that informed managers are able to produce abnormal returns. These conclusions contrast clearly with the earlier studies, in particular Jensen's, who found, on average, a negative alpha for his sample of mutual funds over the 1945-1964 period. However, the Ippolito conclusions are called into question by ELTON, GRUBER, DAS and HLAVKA [1993], for they are heavily dependent on the benchmark used.

The problems raised by inefficient benchmarks together with the development of the Arbitrage Pricing Theory led some researchers to move away from the single-index theoretical framework. Examples of studies in this line of research include CONNOR and KORAJCZYK [1986], LEHMANN and MODEST [1987] and GRINBLATT and TITMAN [1989a]. In particular, LEHMANN and MODEST [1987], using APT-based measures of performance, found negative average performance for mutual funds over the period 1968-1982. Furthermore, these researchers also examine the impact of alternative (CAPM and a variety of APT) benchmarks and found that performance was very sensitive to the benchmark chosen. In light of these results, Lehmann and Modest stress the need for determining the set of benchmarks that reflect the common factors that condition security returns. CONNOR and KORAJCZYK [1991], using a five factor model, evaluate the performance of the same sample used by Lehmann and Modest and reached similar conclusions about the effects of using CAPM or APT benchmarks in estimates of performance.[19] Contrary opinion though was shared by CHEN, COPELAND and MAYERS [1987], who found that performance estimates were unaffected by single or multi-index methodologies.

The controversy was further stimulated by Grinblatt and Titman, who discuss the problem of the appropriate benchmark in performance evaluation (GRINBLATT and TITMAN [1989b]) and introduce a multiple portfolio benchmark formed on the basis of security characteristics - firm size, dividend yield and past returns (GRINBLATT and TITMAN, [1988, 1989a]). The rationale for forming this benchmark (referred to as 'P8') is that various firm characteristics are correlated with their stock's factor loadings. As a result, these characteristics may be better proxies for the true factors than factors formed with statistical factor analysis. In addition, a benchmark formed in such a way cannot be gamed by strategies based on the well known CAPM anomalies. By examining mutual fund performance against this and other three alternative benchmark portfolios, GRINBLATT and TITMAN [1994] confirm LEHMANN and MODEST's [1987] findings on the sensitivity of performance inferences to the benchmark choice.

---

[19] Both studies also address the problem of the bias in Jensen's measure caused by the effect of timing decisions. This issue will be discussed in section 2.6.

Indeed, *"in addition to affecting average performance, benchmarks have a large effect on how funds perform relative to each other"* [GRINBLATT and TITMAN, 1994, p.438]. Using the 'P8' benchmark, GRINBLATT and TITMAN [1989a] found abnormal performance of hypothetical returns of funds, but since these hypothetical returns are computed without deducting expenses of the funds and transaction costs, these results do not imply that investors could realize abnormal returns from investing in these funds. More surprising were the results obtained in a subsequent study, in which GRINBLATT and TITMAN [1994] found evidence of abnormal performance when measuring actual returns net of transaction costs.[20]

The importance of the benchmark portfolio in performance evaluation studies is also illustrated by ELTON, GRUBER, DAS and HLAVKA (EGD&H) [1993]. These researchers re-examine the IPPOLITO [1989] study, which concluded that mutual funds generate abnormal returns, in this way contrasting with a series of mutual fund studies. EGD&H propose the application of a three-index benchmark containing stock portfolios as well as bond portfolios - a broad market index, a firm size index and a bond index. Applying this multiple portfolio benchmark to Ippolito's sample of mutual funds, they verified that superior performance disappeared, with negative average α results similar to Jensen's [1968]. Hence, EGD&H attributed Ippolito's results to the use of an incorrect

[20] GRINBLATT and TITMAN [1993] also found positive performance when applying a measure of performance that does not require the use of a benchmark portfolio, but the observation of portfolio holdings.

benchmark[21] rather than to superior security selection abilities on the part of fund managers. As they comment: "*even a very parsimonious description of a three-factor model of the proxies generating returns on portfolios of bonds and stocks can lead to very different and superior inferences about the attributes of active portfolio management compared to a single-index*" [ELTON, GRUBER, DAS and HLAVKA, 1993, p. 2]. These findings were confirmed by ELTON, GRUBER and BLAKE [1995], in which the use of a one-index model led to a substantial overestimate of fund performance relative to the three-index model.[22] Similar conclusions were reached by SHUKLA and TRZCINKA [1994], who found that alphas using multiple factor benchmarks are lower than those using single factor benchmarks. These researchers argue that this happens because multiple benchmarks control for risk better, causing the risk-adjusted performance to be lower. It is still worthwhile pointing out the study of BLAKE, ELTON and GRUBER [1993] on the evaluation of bond mutual funds. In this investigation, these researchers found that the application of single-index and various multi-index models result in virtually the same rankings of funds, which is in contrast to the results achieved in the evaluation of equity and balanced portfolios.

More recently, ELTON, GRUBER and BLAKE [1996a] introduce an additional index to account for the performance of value/growth

---

21 In particular, they attribute Ippolito's results to fund holdings of small stocks and bonds, not weighted in the index used (S&P 500).

22 BROWN and GOETZMANN [1995] also apply the three-index model of EGD&H and conclude that it "performs well".

factor. The inclusion of this fourth factor is motivated by its correlation with book-to-market ratios, which have been shown by FAMA and FRENCH [1992, 1993, 1995] to be empirically important in explaining stock returns. This four-index model was also applied by GRUBER [1996], who found that it explained 89 percent of the variability of return for the average fund in the sample. ELTON, GRUBER and BLAKE [1996b] further examine several alternative models of the return-generating process, including the one-index model and the four-index model described above. Since this four-index model outperforms the two- and three-index versions in explaining variation in returns, Elton, Gruber and Blake explore whether a fifth index will still result in lower correlation between residuals. The first candidate they propose for the fifth index is derived from the data itself, via factor analysis on the residuals from the four-index model. The results indicate a substantial improvement relative to the four-index model. An alternative approach for the identification of a fifth possible index involves forming portfolios that represent sectors of the economy. However, none of these sector indices outperformed the four-index model in explaining variations in the returns of mutual funds. Finally, the last candidate was derived using data from mutual funds, through an index which represented mutual fund returns themselves rather than passive portfolios of stocks. The introduction of this index outperformed the four-index model and the results are indistinguishable from the five-index model using the factor as the fifth index. The explanation for this phenomenon is that many funds have a large number of common holdings so "*when the effect of common holdings is taken into consideration, a five-index model based on publicly available observable indices seems sufficient to account for the covariances between funds*" [ELTON, GRUBER and BLAKE, 1996b, p. 24].

The use of multi-index model in performance evaluation studies is also illustrated by CARHART [1995, 1997]. The four-factor model used results from FAMA and FRENCH's [1993] three factor model plus an additional factor included to capture JAGADEESH and TITMAN's [1993] one-year momentum anomaly. His results are consistent with the majority of findings in the literature, in that the application of the four-factor model substantially improves on the pricing errors of the CAPM and the three-factor model.

Besides providing better explanations of mutual fund returns, an additional application of the multi-index model is to explain the investment strategy followed by fund managers. This type of analysis was suggested by SHARPE [1988, 1992], which proposes a multifactor model of returns for portfolio evaluation, where the factors are defined to be various asset classes. This asset class factor model, which reflects the portfolio's exposure to variations in the return on the major asset classes, is the appropriate benchmark, since it represents the manager's investment style.[23] GRUBER [1996] enhances this feature of the

[23] The topic of style management and style benchmarks will be discussed in section 2.7.

multi-index model in addition to better explaining mutual fund return behaviour: "*the ability of the four index model to correctly capture the investment policies we know are associated with funds with differing stated objectives is additional evidence that employing the multi-index model leads to more accurate performance evaluation*" [GRUBER, 1996, p. 788-789].

In conclusion, the importance of the choice of the benchmark is illustrated by the contradicting results on average fund performance between single portfolio and multiple portfolio benchmarks. Nevertheless, although there is no consensus on the form of the return-generating process or on the number of influences that should be included in it, the recent evidence on mutual fund performance presented above indicates there is general agreement that multi-factor alternatives are a much more useful characterization of portfolio returns than single-index models.[24]

2.6. TIMING AND SELECTIVITY

Even in the absence of problems concerning the identification of the appropriate benchmark, additional difficulties persist when evaluating performance with the

[24] Obviously, the review of performance evaluation studies that employ multi-index models is not meant to be exhaustive. Other studies that we do not present here, but that do utilize multi-index models include, for example, HENDRICKS, PATEL and ZECKHAUSER [1993] and BROWN and GOETZMANN [1995], which we will survey in section 2.8.

traditional measures described earlier. In fact, this type of investigation focuses solely on the manager's security selection skills by assuming that the portfolio's risk levels are stationary over time. However, superior performance may be achieved as a result of timing (macroforecasting) as well as of security selection (microforecasting) skills of portfolio managers. Security selection involves the identification of individual securities which are under or overvalued and thus, according to the CAPM, lie off the security market line. Market timing refers to forecasts of general market movements. In this sense, managers may deliberately shift the overall risk levels of the portfolios in anticipation of general price movements, switching between low and high beta stocks or between risky and riskless assets according to market conditions. Various studies (for example KLEMKOSKY and MANESS [1978], KON and JEN [1978, 1979], and FABOZZI and FRANCIS [1978, 1979, 1980]) provide evidence that mutual funds do not maintain constant risk levels over time, which is consistent with the timing activities of investment managers. Therefore, the Sharpe-Lintner-Mossin model is likely to generate biased results when used to evaluate performance, since it assumes the systematic risk as a fixed coefficient rather than a decision variable.

The measurement problems involved in evaluating properly the constituents of investment performance when risk levels are nonstationary constitute another source of criticism to the traditional CAPM-based measures of performance. In particular, it has been shown (JENSEN [1972], GRANT [1977], DYBVIG and ROSS [1985b], GRINBLATT and TITMAN [1989b]) that the Jensen measure will be biased downward when the manager possesses and utilizes superior timing information. This contention has come to be empirically demonstrated by CHANG and LEWELLEN [1984], HENRIKSSON [1984] and LEE and RAHMAN [1990]. In response to this problem, GRINBLATT and TITMAN [1989b] propose the "Positive Period Weighting Measure" that is not subject to these timing-related problems of traditional techniques.

The distinction between the part of return attributable to selectivity and that attributable to market timing has received considerable interest in the literature. Next we will briefly discuss the various methods that attempt, at least theoretically, to separate these two types of performance ability.[25]

FAMA [1972] was the first to propose a formalized (theoretical) methodology for the decomposition of total returns into the components of timing and selectivity. In general terms, he divides the return into selectivity and risk components. From finer subdivisions of the risk component, FAMA [1972] develops a theoretical measure of timing, which requires information regarding the target risk level of the fund, a time series of expected returns on the market portfolio and a time series of risk level decisions by the fund manager. However, since the only direct information available to the evaluator is the time series

---

[25] For a more exhaustive analysis of the literature on timing and selectivity see ARMADA [1992].

of return of the market portfolio and the fund, Fama's measures are particularly difficult to implement.

TREYNOR and MAZUY [1966] develop a procedure for detecting timing ability that is based on regression analysis of the managed portfolio's realized returns. These researchers added a quadratic term to equation [2.2] arguing that, if fund managers could forecast general market movements, they will increase the portfolio's risk on the upside and decrease it on the downside, thereby altering the linear CAPM securities line to a nonlinear function as follows:

$$R_{p,t} - R_{f,t} = \alpha + \beta_{1p}\left(R_{m,t} - R_{f,t}\right) + \beta_{2p}\left[\left(R_{m,t} - R_{f,t}\right)^2\right] + \varepsilon_{p,t} \quad [2.7]$$

A positive and statistically significant estimate of the regression coefficient β_{2p} will indicate superior market timing ability.[26] The empirical results obtained by Treynor and Mazuy through the application of this test to a sample of managed funds show no statistical evidence that managers have outguessed the market. CHEN and STOCKUM [1986] develop a similar approach to TREYNOR and MAZUY [1966], but in their regression model beta nonstationarity is allowed to result not only from market timing activities but also from random fluctuations. LEHMANN and MODEST [1987] combine the APT based measures of performance evaluation

[26] PFLEIDERER and BHATTACHARYA [1983] and ADMATI, BHATTACHARYA, PFLEIDERER and ROSS [1986] develop formal conditions under which β_{2p} indicates market timing ability.

with the TREYNOR and MAZUY [1966] quadratic regression technique. More recent empirical examination of the TREYNOR and MAZUY [1966] procedure was also conducted by GRINBLATT and TITMAN [1988] and CUMBY and GLEN [1990], who report negative coefficients on the quadratic term.

JENSEN [1972] proposed a similar formulation for detecting timing and selectivity skills of managers. Under the assumption that the forecasted return and the actual return on the market have joint normal distributions, Jensen shows that market timing ability can be measured by the correlation between the market timer's forecast and the realized return on the market. However, Jensen concluded that, unless the manager's expectations are known, it is impossible to identify the separate contributions of timing and selectivity. By correcting an error made in JENSEN [1972], PFLEIDERER and BHATTACHARYA [1983] show that it is possible to obtain accurate measures of timing and selectivity ability from a simple regression technique. This model is a refinement of the TREYNOR and MAZUY [1966] quadratic regression, as it focuses on the coefficient of the squared excess market return as an indication of timing skill, and requires easily available information on market realized returns and portfolio realized returns. This model was empirically implemented in the U.S. by LEE and RAHMAN [1990], in the U.K. by ARMADA [1992] and in Portugal by CORTEZ and ARMADA [1997]. Although these studies reveal some timing ability, it should be pointed that, within the PFLEIDERER and BHATTACHARYA [1993] approach, timing is constrained not to be negative. COGGIN, FABOZZI and RAHMAN [1993], when applying this methodology to U.S. pension plans, allowed for negative timing. With this alteration, their results reveal the timing measure was negative, so being consistent with those from previous studies.

KON and JEN [1978, 1979] have proposed switching regression techniques to accommodate funds' changing risk levels. In an extension of this methodology, KON [1983] concludes that there is no evidence of timing performance within fund managers as a group. An alternative procedure to analyze market timing was employed by FABOZZI and FRANCIS [1979] and ALEXANDER and STOVER [1980]. A dummy variable regression model was used to fit for two characteristic lines: one for up markets ($R_{m,t}>R_{f,t}$) and one for down markets ($R_{m,t}<R_{f,t}$). The underlying argument is that a manager with timing ability will select a high up market beta and a low down market beta. In general, these researchers found additional evidence that fund managers did not shift their funds beta to take advantages of market movements.

An additional regression-based approach for estimating timing performance is the option approach developed by MERTON [1981] and HENRIKSSON and MERTON [1981]. By assuming that the market timer's forecasts take two possible predictions: either stocks will outperform bonds or bonds will outperform stocks, MERTON [1981] derives an equilibrium theory that shows that the pattern of returns resulting from a market timing strategy is similar to the returns pattern of an option strategy (of the putprotective type). Based on this model, HENRIKSSON and MERTON [1981] develop statistical procedures to test for market timing abilities of investment managers.[27] The regression used is similar to the TREYNOR and MAZUY [1966] regression but in this specification the quadratic term is replaced by $\max(0, R_{f,t}-R_{m,t})$, which is the return on a put option on the portfolio **p** with an exercise price equal to the risk-free rate of return. A significantly positive coefficient for this term indicates superior timing ability, whereas a significantly negative coefficient will indicate perverse timing activity. The results obtained by HENRIKSSON [1984] from the application of these techniques to a sample of mutual funds do not support the hypothesis that mutual funds are able to follow a strategy that successfully times the market. Empirical work with this procedure has also been conducted by CHANG and LEWELLEN [1984], ARMADA [1992] and FLETCHER [1995], who also report no evidence of timing.[28] CONNOR and KORAJCZYK [1991] extend the model to an APT framework,[29] and find similar results.

In conclusion, the majority of the empirical studies on timing seem to suggest that significant timing ability is rare.

---

[27] HENRIKSSON and MERTON [1981] develop both parametric and nonparametric tests of market timing ability. The parametric tests require the assumption that expected returns are priced according to the CAPM for the identification of the separate contributions from timing and selectivity. The nonparametric tests do not require the assumption of a CAPM framework, but do require knowledge of the actual forecasts or a good proxy for them, which is rarely available.

[28] If anything, there seems to be evidence of negative timing, as in ARMADA [1992] and FLETCHER [1995].

[29] Within an APT framework, managers can have timing information with respect to any of the **k** factors.

Furthermore, these studies found more evidence of negative market timing than positive.[30]

2.7. STYLE MANAGEMENT AND PERFORMANCE EVALUATION

Academic studies since the 1960s have examined mutual fund returns and, with few exceptions, found negative performance or no performance for the average mutual fund. In fact, the results of most empirical studies suggest that neither security selection nor market timing abilities are evident in return data on managed funds.

In this context, some literature shifted to the thesis that asset allocation is much more important than focusing on a particular stock or timing: "*the degree of efficiency in the investment marketplace, which makes superior selection increasingly difficult, has caused investors to rethink the role of asset allocation*" [HAMMER, 1991, p. v]. The idea that the return of a portfolio is dominated by investment policy decisions and less by active management was introduced by BRINSON, HOOD and BEEBOWER (BH&B) [1986, 1991]. According to BH&B, the design of a portfolio involves at least four steps: (1)deciding which asset classes to include and to exclude from the portfolio; (2)

[30] Also, a number of researchers (for example: HENRIKSSON [1984], CUMBY and GLEN [1990], ARMADA [1992] and COGGIN, FABOZZI and RAHMAN [1993]) found that there is a strong negative correlation between the measures of timing and selectivity. For a discussion of this phenomenon see HENRIKSSON [1984] and ARMADA [1992].

deciding upon the long-term weights for each of the asset classes; (3) strategically altering the investment mix weights away from normal in an attempt to capture excess returns from short-term fluctuations in asset class prices (*market timing*); and (4) selecting individual securities within an asset class (*security selection*). The first two decisions are part of investment policy. The latter are the contribution of investment strategy. BH&B develop a method for measuring the performance of these activities that compose the investment management process - investment policy, market timing and security selection. The asset allocation decision - determining "*the allocation of an investor's portfolio across a number of major asset classes*" [SHARPE, 1992, p. 7] - was shown to have a major impact on performance. Indeed, BH&B find that asset allocation explains about 94% of the cross-sectional variations in returns.[31] Hence, this seminal study on performance attribution has stressed that the asset allocation policy decision is far more important than market timing or security selection: "(...) *the total return to a plan is dominated by investment policy decisions. Active management, while important, describes far less of a plan's returns than investment policy*" [BRINSON, HOOD and BEEBOWER, 1986, p. 43]. The performance attribution framework provided by BH&B to decompose total returns is quite simple. Returns on a passive benchmark portfolio representing the fund's long term asset classes are compared with the actual returns resulting from

[31] Similar results were obtained by SHARPE [1992].

the combination of investment policy plus investment strategy (market timing and security selection).

Once the asset allocation decision is made, performance should now be measured relatively to the benchmark that defines the assets. The critical variable is asset allocation, and the performance measure should be the returns relative to the benchmark.

In this way, by measuring the contribution of active management over a passive benchmark, this approach to performance evaluation does not suffer from some of the problems inherent to the traditional measures of performance, which we have discussed earlier. In particular, problems associated with the appropriateness of benchmarks, the suitability of beta as the risk measure and the validity of the asset pricing theory assumed are overcome with this type of analysis.[32]

The idea that the asset allocation decision is a major factor affecting the return of a portfolio led managers to concentrate on *asset-class investing*, that is, investing in categories of assets with similar characteristics and performance patterns rather than individual securities. As investment managers become increasingly specialized, they chose to concentrate not only on a specific asset class (for example stocks or bonds), but within those asset classes they focus efforts on certain types of securities. For examples, equity

[32] Subsequent studies that develop related methods for determining the impact of different risk-taking decisions on investment returns include, for example, HENSEL, EZRA and ILKIW [1991], ANKRIM [1992] and SINGER and KARNOSKY [1995].

market subsegments can be identified by such characteristics as size and value versus growth. These specializations have become to be known as investment styles.

The concept of style management has become increasingly important both to academics and professional managers. This trend is reflected in the emergence of academic studies devoted to this topic, in the number of style indices designed to measure subsegments of the market and in the appearance of style index funds as an alternative to active management.

Investment styles have significant implications for the evaluation of investment performance. Performance similarities are to be expected within a certain style, because these portfolios share similar characteristics and factor exposures that lead to a certain pattern of expected returns.[33] Hence evidence of differential performance may not reflect the manager's ability to select securities, but simply the category of securities in which the manager invests. Conversely, a fund manager may earn a superior return relative to other funds with the same style; but the style itself may perform worse in a particular period. In this case, if evaluated against a broad benchmark, this manager will be judged negatively. In addition, the use of general benchmarks is inappropriate for evaluating performance since they can be gamed by managers aware of CAPM anomalies, like the size or dividend-yield effects. For these

[33] Evidence that different style indices result in different performance patterns can be found for example in CHRISTOPHERSON and WILLIAMS [1995] and LUCK [1995].

reasons, evaluation of the manager's performance requires the consideration of his investment style.

One means of measuring the performance relative to the manager's respective style is by developing benchmark portfolios that specifically represent the manager's investment style. These benchmarks, known as "normal" or "custom" portfolios, are composed of the type of securities in which the manager invests in, weighted in a manner consistent with the manager's investment process.[34] In other words, "*it incorporates the prominent and persistent characteristics of a manager's portfolio in the absence of active management*" [BAILEY, 1992a, p. 10]. Style benchmarks are designed to capture the essence of the manager's style without reflecting the manager's value added. In this way, these benchmarks allow for the separation of the returns into those attributable to style (or asset class exposure of the manager) and the returns that reflect the manager's skill within a particular style.

The assumption that the manager's performance is to be compared against a valid benchmark raises the question of defining what is a "correct" benchmark. From what has been presented above, it is easily accepted that "*good benchmarks increase the proficiency of performance evaluation, highlighting the active management contributions of managers. Poor benchmarks obscure manager skills (...) and distort performance evaluation*"

[34] Detailed discussions on normal portfolios and their construction process appear in KRITZMAN [1987] and RENNIE and COWHEY [1990].

[BAILEY, 1992b, p. 34, 38]. The issue of benchmark quality[35] has been emphasized in a number of recent articles by DIVECHA and GRINOLD [1989], TIERNEY and WINSTON [1990, 1991], BAILEY [1992a, 1992b], BAILEY and TIERNEY [1993] and TIERNEY and BAILEY [1995]. In particular, BAILEY [1992a] identifies several properties a benchmark should possess. Specifically, a valid benchmark should be:

- unambiguous;
- investable;
- measurable;
- appropriate;
- reflective of current investment opinions; and
- specified in advance.

The importance of investment style as the basis for performance evaluation and compensation schemes raises the issue of mutual fund classification systems. Assignment by style assumes that the composition of the fund will not contradict its objectives. Both BROWN and GOETZMANN [1997], using a classification algorithm which groups managers into similar styles according to their realized performance, and DiBARTOLOMEO and WITKOWSKI [1997], using style analysis, find evidence that mutual funds often misclassify themselves, and suggest that such misclassification may be intentional. As these authors comment:

[35] Benchmark quality is defined as the "*ability to reflect accurately the prominent and persistent risk-return characteristics that a manager's portfolio would exhibit in the absence of active management*" [BAILEY, 1992b, p. 34].

"There is a great need for stylistic classifications that are objectively and empirically determinated, consistent across managers and related to the manager's strategy. The objectivity is important because of the moral hazard inherent in allowing managers to self-report their styles, without objective verification" [BROWN and GOETZMANN, 1996, p. 3].

In this context, two general approaches have been developed for determining a portfolio's style: the return-based approach and the portfolio-based approach. The latter, as described by CHRISTOPHERSON [1995] and CHRISTOPHERSON and TRITTIN [1995], is based on the characteristics of the securities within the portfolio. Portfolio characteristics such as *price-to-book* (P/B), *price-to-earnings* (P/E) and *dividend yield* are determined and managers are assigned to the style group accordingly. The main strength of this approach is that, as it carefully examines the assets, more information in the evaluation process may produce a more accurate determination of the manager's style. However, a potential weakness of this approach, besides being a very labourious process, is that it is not easy to determine whether the evaluation has been complete and unbiased (TRZCINKA [1995b]).

The return-based approach, also known as "effective mix" or correlation based" approach, was introduced by SHARPE [1988, 1992] and assigns style on the basis of return pattern analysis. This methodology proceeds from the assumption that observing portfolio returns is more important than observing portfolio holdings. The goal of style analysis is to determine the fund's exposure to changes in the returns of major asset classes through the use of an asset class factor model.

There are several critical decisions in a return-based analysis. Besides the decision about the appropriate time period, the most important choice is the asset classes selected to represent styles. The selected asset classes must be: 1) "mutually exclusive", 2) "exhaustive" and 3) have returns that differ.[36] The return on each asset class is represented by a market capitalization index of returns. Using this model, it is a simple matter to determine the combination of style indexes that gives the highest R^2 to the manager's returns.

Sharpe suggests a model using 12 indices representing factors that affect stock returns.[37] Four of these indices represent U.S. equity classes. In effect, equity market subsegments can be identified by such characteristics as size (small/large) and value/growth. Studies, for example FAMA and FRENCH [1992] and ROLL [1995] have shown that, in the U.S., returns on value and growth stocks can differ significantly and that value stocks have outperformed growth stocks over the long run. CAPAUL, ROWLEY and SHARPE [1993] addressed this question relative to other major markets[38] and also documented the existence of a significant value-growth factor in each country.

---

36 See SHARPE [1992, p.8].

37 ROLL [1995] suggests that no more than five factors are relevant for representing equity returns.

38 U.K., France, Germany, Switzerland, Japan and U.S.

In addition to determining the manager's style, which is the mix of indices' returns that match the portfolios' returns the closest, style analysis provides a natural method for constructing benchmarks, or "normal" portfolios. The return on the fund can be compared to the return of a predetermined asset mix with the same style. It is the excess return over the style benchmark that reflects the manager's ability to select stocks.

The simplicity and objectivity of this methodology are unanimously considered its main strength. There is no need to investigate the portfolio holdings to determine style: only the returns of the fund and of the set of indices are required. However, two objections were raised by CHRISTOPHERSON [1995] and CHRISTOPHERSON and TRITTIN [1995] to this methodology. The first is the susceptibility of returns to noisy data. These researchers argue that correlational analysis cannot distinguish between noise in the data and true factor exposures, which can induce errors in factor identification. The second point is that the returns-based approach cannot capture changing styles, that is, it "*tends to be blind to style dynamics*" [CHRISTOPHERSON, 1995, p. 37]. This results from the nature of the methodology itself, since the most recent time point is as important as the most remote time period, causing a delay in the recognition of a style shift. TRZCINKA [1995a] provides a comparison of two approaches and discusses Christopherson's critiques, commenting that "*the practical question (...) is whether these general arguments about models of portfolio behaviour imply that effective mix is* *useless. This is an empirical question, because it is always easy to show that economic data violate statistical assumptions of economic models, but it is common to find that economic models are still useful"* [TRZCINKA, 1995a, p.45].

Another methodology based on the analysis of covariance structures between manager returns and index returns was developed by TIERNEY and WINSTON [1991]. Their "style-point" method portrays the correlation of the manager's returns with its index returns in a two-dimensional space (small to large and value to growth), allowing for a intuitively appealing visualization of a portfolio's investment style.

A natural question at this point is how styles behave in light of the CAPM and the assumption of efficient markets. According to the CAPM, all stocks are driven by the market factor, and consequently their expected return will be dependent upon beta. In addition, efficiency of the market implies that all securities are correctly priced given all available information. In such a world, style management would not be more profitable than investing in any arbitrary subset of securities, and no security characteristic would lead to differential return. Thus, evidence that style management earns a higher than expected return is evidence than the security market is inefficient or that the risk adjustment model is not correct. The financial literature has found style dimensions such as size and value/growth to be associated with cross-sectional difference for equities (FAMA and FRENCH [1992]). SHARPE [1992] gives other style dimensions for both equity and fixed-income assets. Hence, a multifactor model with at least two risk premiums would, a priori, have more potential to explain style returns than a single risk premium model.

ROLL [1995] investigates the origin of historically differential returns and explores three possible explanations. First, return differentials across investment styles might simply be statistical aberrations (for example, problems of data mining and selection bias). This being so, they would not reflect differences in *expected* returns and thus would not likely be repeated. Secondly, return differentials are risk premiums. They reflect differences in expected returns, but this is compensation for bearing risk. Finally, return differentials represent market opportunities, which occur above and beyond any measurable risk. In this case, investing according to style could be expected to earn extra return without any additional exposure to loss. ROLL [1995] empirically examines the second and third possible explanations arguing that by assuming some risk/return model it is straightforward to determine whether style performance could be attributable to risk. Evidence shows that three style dimensions (size, earning/price and book-to-market) are statistically significant determinants of return. Yet, both the single-factor CAPM and the multifactor APT (with five factors) used did not fully explain differential performance among the styles. Hence Roll concluded that there is evidence that style is associated with extra-return risk, which is consistent with market inefficiency. However, one should be reminded of FAMA's [1991] observation that any test of market efficiency is both a test of the asset pricing model and market efficiency, so "*because of the joint-hypothesis problem, precise inferences about the degree of market efficiency are likely to remain impossible*" [FAMA, 1991, p. 1576].

2.8. PERSISTENCE OF PERFORMANCE

With few exceptions, the evidence presented earlier on mutual fund performance studies indicates, despite differences in the methodology and data used that, in general, managers have not been especially successful at active portfolio strategies, either in the form of security selection and/or market timing. In this way, these studies support the Efficient Market Hypothesis (EMH) as the accepted paradigm in the academic community. While most mutual fund studies test whether funds exhibit abnormal performance, until recently little research has been devoted to analysing how future performance relates to past performance. Notwithstanding, the practitioner literature has for long given considerable attention to performance rankings of funds (compiled by the financial press or the mutual fund industry itself), expressing the belief that past performance predicts future performance. Early tests of performance persistence have been conducted as a part of a larger study of mutual fund performance.

LEHMANN and MODEST [1987] show some evidence of persistence, although they suggest that this finding is quite sensitive to the risk-adjustment model used to compute the performance measures. GRINBLATT and TITMAN [1988] report statistical evidence of persistence in mutual fund risk-adjusted returns over five-year periods. LAKONISHOK, SCHLEIFER and VISHNY [1992] found, for some pension fund managers, evidence of persistence in rankings for two- to three-year periods. More recently, several studies published in the early 1990s have directly examined persistence in fund performance and claim to have detected a "hot hand" phenomenon.[39] For example HENDRICKS, PATEL and ZECKHAUSER [1993] and GOETZMANN and IBBOTSON [1994] argue that past mutual fund returns predict future returns. This type of evidence is not only inconsistent with efficient markets, which states that past performance is no guide to future performance, but has also practical value to investors, as it suggests that they may realize abnormal returns by purchasing recently good-performing funds.[40] Thus, the question of persistence of performance remains a controversial issue itself, with some studies providing evidence of consistency of performance and others attributing these results to biases present in the data, which would produce the appearance that performance is predictable even when it is not. We will next review the history of investigation into the

[39] This term was transposed to the finance literature by HENDRICKS, PATEL and ZECKHAUSER [1993], to characterize funds that deliver sustained superior performance.

[40] Furthermore, most advertising by the mutual fund industry is based precisely on their claims to persistent high rankings.

question of persistence and discuss a number of issues outstanding.

Studies which address the question of persistence can be divided into those that do not find persistence and those that do. In examining the consistency of mutual fund risk-adjusted performance, KLEMKOSKY [1977] finds no evidence that past performance is useful to predict future performance. KRITZMAN [1983] finds that there is no relation in relative rankings between the performance of fixed-income managers for two consecutive five-year periods. Similar results were found by DUNN and THEISEN [1983] on institutional portfolios and by ELTON, GRUBER and RENTZLER [1990] on commodity funds. Relatively to pension fund managers, CHRISTOPHERSON and TURNER [1991] found no relationship between α's over successive three-year periods. More recently HENDRICKS, PATEL and ZECKHAUSER [1993], GOETZMANN and IBBOTSON [1994], BROWN and GOETZMANN [1995], WERMERS [1997] and CARHART [1997] find evidence of short-term persistence in mutual funds. GRINBLATT and TITMAN [1992], ELTON, GRUBER and BLAKE [1996a] and VOLKMAN and WOHAR [1996] observe persistence over longer periods. SHUKLA and TRZCINKA [1994] and CARHART [1997] find that persistence is concentrated in the poorly performing funds.

Although many studies provide evidence on performance persistence, the results require careful interpretation. BROWN, GOETZMANN, IBBOTSON and ROSS (BGI&R) [1992] argue that persistent positive performance can result from survivorship bias in the data. In fact, commonly employed data sets of mutual funds typically show the past records of all funds in existence during the total sample period. Such a sampling scheme creates the possibility of survivorship bias since only the returns of surviving funds are selected, whereas those of disappearing funds are omitted. Thus, to study only funds that survive (superior performers) overstates the measured performance. BGI&R [1992] investigated the influence of survivorship bias in performance studies and showed that this effect generates the appearance of persistence in performance even in the absence of "true" persistence. One means of correcting for survivorship bias is by employing a sample that includes the records of *all* funds in existence during the evaluation period, including those who existed in any year but disappeared (MALKIEL [1995], BROWN and GOETZMANN [1995], CARHART [1995, 1997], ELTON, GRUBER and BLAKE [1996a], GRUBER [1996]). Other studies explicitly estimated and/or corrected for survivorship bias (GRINBLATT and TITMAN [1989a], HENDRICKS, PATEL and ZECKHAUSER [1993], CARHART [1995], ELTON, GRUBER and BLAKE [1996c]).

A related point concerns the existence of "cold-hands", that is, persistence of negative performance, documented in several studies, such as those of BGI&R [1992], SHUKLA and TRZCINKA [1994] GRUBER [1996] and CARHART [1997]. The finding that negative performance tends to be followed by future poor performance is a puzzling question. BGI&R [1992] suggest that negative performance can persist for institutional reasons such as immunity from periodic performance review (which allow funds with sustained negative performance to survive) and the difficulty to short sell shares of mutual funds. This issue is also discussed by GRUBER [1996], who questions why do any investors remain in funds that consistently perform poorly. According to Gruber, there is a possible explanation for this puzzle: the existence of two groups of investors, a sophisticated clientele and a disadvantaged clientele. The first group makes investing decisions based on fund performance. The disadvantaged group consists of three types: (1) unsophisticated investors, whose decisions are based on influences such as advertising and advice from brokers; (2) institutionally disadvantaged investors, mainly represented by pension accounts which have restrictions set by the plan they are part of and (3) tax disadvantaged investors, a group that has held funds for enough time so that capital gain taxes make it inefficient to move away from. GOETZMANN and PELES [1997] investigate the role of investor psychology in decision making. They find that investor memories exhibit a positive bias, that is, investor recollections of past performance are consistently biased above actual past performance.[41] This bias, consistent with the theory of cognitive dissonance, may be why investors justify remaining in funds that consistently perform poorly and are slow to respond to past poor performance. In consequence, "*although the market rewards the top*

[41] The results of this study suggest that even well-informed investors tend to bias their perceptions about past performance.

performers each year, it does little to discipline poor performers" [GOETZMANN and PELES, 1997, p. 156].

The methodologies for assessing persistence as well as the construction of mutual fund data samples differ considerably among the studies we have mentioned.

HENDRICKS, PATEL and ZECKHAUSER (HP&Z) [1993] suggest computing the returns on a self-financing portfolio strategy corresponding to the "residual"[42] performance of the fund. The performance measure of such a portfolio is a measure of persistence. In their study, HP&Z find statistical evidence of persistence ("hot-hands"), although it is mostly a short-run phenomenon, with the strongest evidence for a one-year evaluation horizon. As a consequence, investors could profit from a strategy of investing in the top-performing funds in the prior four quarters. The results are robust to survivorship bias[43] as well as to various scenarios such as assessment with a variety of benchmark portfolios, including a multiple portfolio benchmark that accommodates well known anomalies and investigations with alternative data sets of mutual funds. CARHART [1997] replicates the methodology of HP&Z [1993] by forming portfolios of mutual funds based on their previous year's return, and evaluates the performance of the resulting portfolios. Using a sample free of survivorship bias, CARHART [1997] finds evidence of short-term

[42] "Residual" performance is perceived as the difference of the performance measures from the mean performance across managers.

[43] The sample was constructed to avoid problems of survivorship bias. Anyway, to confirm their findings, they explore the sensitivity of persistence inferences to subsamples purposely selected to induce survivorship bias, and such a bias appears to be unimportant.

persistence. However, rather than attributing this phenomenon to stock-picking skill (where stock-picking skill is defined as abnormal return after accounting for common factors in stock returns) or "hot-hands", Carhart provides evidence that short-term persistence is almost completely explained by common factors[44] in stock returns as well as investment expenses. Overall, Carhart argues that the results are consistent with market efficiency and do not provide convincing evidence of mutual fund portfolio manager stock-picking ability.

One other method for assessing performance persistence is through regression analysis, in which future performance of a manager (Period t+1) is regressed against a measure of past performance (Period t):

$$\text{Performance }(t+1) = a + b * \text{Performance }(t) + \varepsilon \qquad [2.8]$$

where performance can be measured in terms of total returns, risk-adjusted returns or information ratios. A significant positive **t** statistic for the slope coefficient in this regression would reject the null hypothesis that past information is unrelated to future performance and is thus evidence of persistence. This methodology was employed by GRINBLATT and TITMAN [1992], whose cross-sectional regressions used risk-adjusted alphas resulting from the 'P8' multiple portfolio

[44] These factors are accommodated in CARHART's [1997] four-factor model described in section 2.3.2.

benchmark[45] as inputs for performance. Grinblatt and Titman find evidence of positive persistence in mutual fund performance for five-year periods, which they attribute to manager's superior information or stock selection abilities.[46] KAHN and RUDD [1995] define performance in terms of total returns as well as selection returns and information ratios. Selection returns were computed via style analysis[47] and are defined as the portfolio return over the style benchmark return. In this context the information ratio is the ratio of selection return mean to standard deviation. Persistence of mutual funds was investigated through regression analysis, but the results did not support persistence for equity funds.

Another common methodology for examining persistence involves ranking funds by performance and observing whether the rankings tend to be preserved over time. By ranking funds according to total returns, BOGLE [1992] followed the track of the top equity fund performers for at least a period of ten years, and found that the return in one year has no relationship to its ranking in the subsequent year. BAUMAN and MILLER [1994, 1995] found that the rankings of performance by total return quartiles were relatively consistent over time when based on investment styles. SHUKLA and TRZCINKA [1994]

---

[45] This benchmark is described in section 2.3.2.

[46] The sample used by GRINBLATT and TITMAN [1992] was not corrected for survivorship bias, since previously GRINBLATT and TITMAN [1989a] conclude that survivorship bias is negligible.

[47] Style analysis, originally developed by SHARPE [1988, 1992] was described in section 2.7.

argue that if funds perform well consistently, their rankings from one period to another should be highly correlated. Their results (for alternative risk-adjustment benchmarks) indicate that there is little persistence of performance among positive performers while, on the contrary, there is significant persistence in poorly performing funds. BAL and LEGER [1996] evaluate the funds' performance by the Sharpe measure and find persistence of fund rankings over time. However, in light of the benchmark problem discussed in section 2.3.2, these results do not appear robust and may be explained by the use of an inappropriate framework for risk adjustment - ELTON, GRUBER, DAS and HLAVKA's [1993] critiques to the IPPOLITO [1989] study seem to apply here as well. In examining persistence of performance, ELTON, GRUBER and BLAKE [1996a] separate funds into deciles on the basis of risk-adjusted return[48] and find these rankings to be highly correlated with future performance over three-year periods. Rather than being a short-term phenomenon (noted in the "hot-hands" literature), ELTON, GRUBER and BLAKE [1996a] suggest that there is a longer persistence in performance. In addition, they show portfolios of funds formed on the basis of past information can produce positive risk-adjusted excess returns. GRUBER [1996], using raw returns as well as risk-adjusted returns (from single and four-index models), also provides evidence of persistence in mutual fund performance over one to three years. In addition,

[48] Using the multi-index model described in section 2.3.2.

he finds that persistence is more significant when measured in terms of the four-factor index performance measure.

An alternative approach to assess persistence follows a nonparametric methodology based upon two-way contingency tables. Over successive periods, funds are categorized as winners and losers by ranking fund performance according to whether or not they were above or below median performance. Contingency tables show the frequency with which winners and losers repeat. If statistical evidence shows that winners in one period remain winners in the subsequent period, the case for persistence in performance is shown. GOETZMANN and IBBOTSON [1994] report contingency tables for total and risk-adjusted returns and for a variety of time periods and performance horizon intervals, and conclude in favour of short-term persistence. The regression analysis performed for the same scenarios confirm the results. BROWN and GOETZMANN [1995] also explored the phenomenon of persistence through the methodology based upon contingency tables, using a sample free of survivorship bias. Their results indicated that risk-adjusted performance of mutual funds persists, and that it is due to common investment strategies that are not captured by risk-adjustment procedures. Moreover, by desegregating the persistence tests on an annual basis, Brown and Goetzmann find this phenomenon is highly dependent upon the time period of analysis. Following the same methodology but using total returns, MALKIEL [1995] corroborates that conclusion on the time dependency of results. In particular, the strong persistence detected in the decade of 1970 disappeared during the 1980s. Moreover, using a sample free of survivorship bias, Malkiel measures the extent of this effect, and finds that survivorship bias is considerably more important than previous studies have suggested. Contingency tables were also constructed in the KAHN and RUDD [1995] study described above, which confirmed their conclusions obtained via regression analysis. More recently, and following the same methodology, PHELPS and DETZEL [1997] document persistence in mutual fund performance. However, this persistence disappears when returns are adjusted for risk. In a study of the performance of hedge funds, BROWN, GOETZMANN and IBBOTSON [1998], using cross-sectional regressions and contingency tables, find no evidence of performance persistence either in terms of raw returns or style-adjusted returns.

Table 2.1. summarizes the most relevant mutual fund studies on performance persistence with respect to methodology, data sets, performance measure and benchmarks employed, as well as main findings.

Table 2.1

Summary of the performance persistence studies

METHODOLOGY	DATA	PERFORMANCE MEASURE AND BENCHMARKS	MAIN FINDINGS
TIME-SERIES SELF-FINANCING PORTFOLIO APPROACH			
• HENDRICKS, PATEL and ZECKHAUSER [1993]	Quarterly returns for 165 mutual funds over the period 1974-1988. Sample not affected by survivorship bias.	Risk-adjusted returns from single and multiple portfolio benchmarks.	Short-term persistence
• CARHART [1997]	Monthly returns for all equity funds in existence since 1962 to 1993 (1892 funds). Data free from survivorship bias.	Risk-adjusted returns for single and multiple portfolio benchmarks.	Persistence is concentrated in the poor performing funds.
CROSS SECTIONAL REGRESSIONS			
• GRINBLATT and TITMAN [1992]	Monthly returns for 279 mutual funds from December 1994 to December 1984.	Risk-adjusted returns from multiple portfolio benchmark ('P8')	Positive persistence for five-year periods.
RANKS CATEGORIES			
• SHUKLA and TRZCINKA [1994]	Monthly returns for 1387 stock and bond funds over the period April 1979 to March 1989.	Risk-adjusted returns from single and multiple portfolio benchmarks.	Persistence of inferior performance

• ELTON, GRUBER and BLAKE [1996a]	Monthly returns for 188 equity funds over the period 1997 to 1993. Sample free of survivorship bias.	Risk-adjusted returns from multiple portfolio benchmark.	Long-term persistence of performance
WINNERS/LOSERS CONTINGENCY TABLES			
• GOETZMANN and IBBOTSON [1994]	Monthly returns of 728 mutual funds over the period 1976 to 1988.	Raw returns and risk-adjusted returns from the single-index model.	Short-term persistence of performance.
• BROWN and GOETZMANN [1995]	Annual returns for all equity funds since 1976 to 1988 (829 funds). Sample free of survivorship bias.	Raw returns and risk-adjusted returns from single and multiple portfolio benchmarks.	Persistence of performance; Time dependency of persistence.
• MALKIEL [1995]	Quarterly returns for all equity funds in existence from 1971 to 1991 (724 funds). Sample free of survivorship bias.	Raw returns	Persistence of performance; Time dependency of performance.
• KAHN and RUDD [1995]	Monthly returns for 300 equity funds and 195 fixed-income funds over the period 1983-1993 and 1986-1993, respectively.	Raw returns, information ratios and risk-adjusted returns from style benchmarks.	No persistence for equity funds.

The mixed evidence on performance persistence studies imply that their conclusions must be interpreted with caution, in light of various key issues. Besides the topic of survivorship bias, it is a pertinent question, for example, to examine the influence of expenses in persistence of performance. The abnormal returns of successful funds could be eliminated by increases in fees, resulting in no continuity of performance even when managers are able to yield superior performance. Studies that have explicitly investigated the influence of fund expenses in performance persistence include VOLKMAN and WOHAR [1995], CARHART [1997] and ELTON, GRUBER and BLAKE [1996a]. In particular, VOLKMAN and WOHAR [1995] find that funds with low management fees demonstrate significantly positive persistence. CARHART [1997] finds a strong negative relation between load fees and performance. However, ELTON, GRUBER and BLAKE [1996a] find that there is still predictability even after the major impact of expenses have been removed.

Furthermore, the consideration of investment style in studies of persistence is also a relevant matter. Otherwise, one would not be sure whether persistence of performance is the result of manager skill or instead of an incorrect adjustment to the risk assumed by the manager. Several studies on persistence of performance stress that a more appropriate way to adjust for risk is to control for different investment styles.[49] In light of

[49] Either via multi-factor models, for example, the multiple factor model employed by ELTON, GRUBER and BLAKE [1996a], which is designed to account for styles influences by including a size and a value/growth factor or via the "returns-based" approach developed by SHARPE [1988, 1992].

what we have discussed on investment styles, it is reasonable to question if the repeat performance phenomenon is driven by style effects. Do some style segments perform differently from the market and do so consistently? BROWN and GOETZMANN [1995] address this issue and conclude that the pattern of performance persistence is little affected by subtracting the style benchmark. However, it should be noted that the identification of funds according to their style was based on conventional stylistic classifications, which as we have already seen in section 2.7, can present several problems. KAHN and RUDD [1995] argue that studies that do not control for style effects might generate the appearance of persistence and suggest that is a possible explanation for why their results (obtained through style analysis) differ from those who find evidence of persistence. SHARPE [1996] also examines the persistence of a sample of the 100 largest mutual funds through *his* methodology of style analysis. By ranking funds according to selection returns, past performance is related to future performance. The results obtained suggest some evidence on short-term persistence: "*The evidence (on persistence) is far from conclusive, statistically or econometrically. Perhaps the only safe conclusion is that there is little support for the thesis that (...) past losers are due to and likely to outperform past winners*" [SHARPE, 1996]. This type of evidence suggests that persistence of performance is not the result of superior management skills but rather the result of persistence in specific asset classes. In this way, it is consistent with the idea that the benchmark would not have been properly controlled for the various dimensions of risk. In the terminology of PHELPS and DETZEL [1997], the positive performance found in the literature might be driven by *macropersistence* rather than *micropersistence*.

Finally, we are particularly concerned with the analysis of performance persistence in the context of small markets, where the number of funds and the time period of analysis is quite low compared to countries such as the U.S. and the U.K. The applicability of the methodologies we have reviewed to such small markets (e.g. the Portuguese fund market) is not straightforward due to the possibility of biased results in the presence of a small sample. Therefore, tests used to assess persistence require adjustments for eventual small sample bias. We are not aware of any other investigation of performance persistence of mutual funds (or any other type of funds) for the case of limited sample size.

In conclusion, although the majority of the literature provides evidence on persistence of performance, there is no general consensus on how strong this phenomenon appears to be nor to what sources it should be attributed. Furthermore, several issues must be taken into consideration:

- Survivorship bias is a problem and can significantly influence the results;

- Different studies have looked at different time periods, and as MALKIEL [1995] suggests, persistence may be stronger in some periods than in others;
- Expense fees can also have impact on persistence;
- Many studies on persistence, although risk-adjusted, do not account for style effects;
- Possible bias resulting from a limited sample size.

2.9. CONCLUSIONS

In this chapter we have presented major methodologies for evaluating managed funds, at the same time discussing their limitations and the main debated issues. In particular, we have emphasized one of the most controversial issues in performance evaluation, which is the identification of the appropriate benchmark portfolio and discussed the use of single versus multiple benchmarks. We also illustrated the contribution of investment style to performance evaluation, in particular its application relative to style benchmarks, and presented alternative style measurement approaches, from multi-factor models to the "style analysis" methodology. Finally, we reviewed the literature and presented the state-of-the-art on one of the most recent topics in portfolio performance literature, performance persistence. We have described various approaches for measuring persistence of performance and discussed the main issues currently in debate, namely benchmark criteria, survivorship bias, and small sample bias. In the context of a small market (e.g. the Portuguese fund market), the latter consideration has greatly conditioned the choice of the methodology adopted to assess performance persistence. In this sense, and for reasons exposed in the next chapter, among the different methodologies, we have chosen to apply the methodology of two-way contingency tables of winners and losers to assess the performance persistence of our sample of funds. This methodology will be presented in the next chapter.

CHAPTER 3

THE ANALYSIS OF PERFORMANCE PERSISTENCE IN THE CONTEXT OF CONTINGENCY TABLES

3.1. INTRODUCTION

As we have seen in the previous chapter, when reviewing and discussing the literature, one means of assessing the persistence of performance is through the construction of two-way contingency tables[50] of performance over successive periods. Other methodologies, such as those involving rank portfolios of quartiles (or octiles) or regressing past performance on future performance are clearly methodologies that are not appropriate for evaluating the performance persistence of funds in small markets (e.g. Portuguese funds), due to the limited sample of funds available. In this context, and to overcome this limitation, the methodology based on contingency tables is an attractive and practical alternative for analysis. For each period, funds are defined as winners if they are above the median of all funds with respect to some performance measure. For the next period, funds are categorized as winners or losers following the same procedure.[51] The basic idea underlying this type of approach is that statistical evidence that winners in one period remain winners in the subsequent period is not consistent with the hypothesis of no persistence of performance.

In this chapter we will present and discuss methods of analysis of performance persistence in the context of frequency

---

[50] This term is attributed to PEARSON [1904](referenced in MURTEIRA [1990]), who defined contingency for a r x c table as a measure of distance relative to a situation of independence.

[51] Obviously, the possible combinations of winners and losers for two subsequent periods are winner-winner (WW), winner-loser (WL), loser-winner (LW) and loser-loser (LL).

data occurring in the form of cross-classifications or contingency tables. The statistical model for such data is derived when a sample from some population is classified with respect to two or more qualitative variables.[52] This classification results in a matrix with r rows and c columns, known as two-way (r x c) contingency tables.

3.2. CRITERIA FOR EVALUATING PERFORMANCE PERSISTENCE

MALKIEL [1995] shows the "percentage of repeat winners" and also constructs a Z-test for repeat winners to test the hypothesis of no winning persistence.

If **p** is the probability that a winner in one period remains a winner in the subsequent period, **p** would be expected to be ½ , if there is no persistence. Since the random variable **Y** of the number of persistently winning funds follows a binomial distribution *b(n,p)*, we can determine if the probability **p** of consistent winning is greater than ½. When **n** is reasonably large, the random variable $Z = (Y - np) / \sqrt{np(1-p)}$ will be approximately distributed as normal with zero mean and standard deviation one.[53]

[52] Other contexts for treatment of contingency tables are described for example in ANDERSEN [1990].

[53] MALKIEL [1995] defines **n** as the number of winner-winners and winner-losers.

BROWN and GOETZMANN [1995] construct contingency tables and calculate the odds ratio (also referred to as cross-product ratio) and a Z - statistic (the log odds ratio divided by its standard error). KHAN and RUDD [1995] utilize the chi-square statistic to test performance persistence via contingency tables. We will next describe these nonparametric tests.

Let $\mathbf{p}_{i.}$ represent the probability, in the population, of an observation belonging to the i^{th} category of the row variable and $\mathbf{p}_{.j}$ the corresponding probability for the j^{th} category of the column variable. Then, from the multiplication law of probability, independence between two variables in the population implies that:

$$p_{ij} = p_{i.}p_{.j} \qquad [3.1]$$

which is the null hypothesis H_0.

In terms of the frequencies F_{ij} to be expected in the ij^{th} cell of the contingency table, independence implies that:

$$F_{ij} = Np_{i.}p_{.j} \qquad [3.2]$$

Since the population probability values are unknown, the maximum likelihood estimates of the marginal probabilities (when the hypothesis is true) are:

$$\hat{p}_{i.} = \frac{n_{i.}}{N} \quad \text{and} \quad \hat{p}_{.j} = \frac{n_{.j}}{N} \qquad [3.3]$$

Hence, the use of $p_{i.}$ and $p_{.j}$ given above allow inference about the distribution of the population over the cells of the contingency table, if the two variables are independent. From equations [3.3] this estimate, which is generally represented by E_{ij}, is given by:

$$E_{ij} = N\hat{p}_{i.}\hat{p}_{.j} = \frac{n_{i.}n_{.j}}{N} \qquad [3.4]$$

When the two variables are independent, the estimated frequencies and the observed frequencies differ by amounts attributable to chance factors only. Consequently, it is reasonable that a test of independence of two variables forming a contingency table should be based on the size of the differences between the estimated values of the frequencies to be expected assuming that H_0 is true (E_{ij}) and the observed frequencies (n_{ij}). Such a test, first developed by PEARSON [1904],[54] is defined as:

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(n_{ij} - E_{ij}\right)^2}{E_{ij}} \qquad [3.5]$$

As we have seen, the magnitude of this statistic depends on the values of the differences $\left(n_{ij} - E_{ij}\right)$. If H_0 is true, then

[54] PEARSON, K. [1904] "On the theory of contingency and its relation to association and normal correlation", referenced in EVERITT [1992].

$(n_{ij} - E_{ij})^2$ should be reasonably small. Consequently, χ^2 will be smaller when H_0 is true than when it is false.

Under the assumption that the observed frequencies follow a multinomial distribution, and further assuming large expected frequencies, χ^2 can be shown to have approximately a chi-square distribution. Test of the hypothesis of independence can be performed by comparing the statistic χ^2 with the values of the chi-square distribution.[55]

A contingency table with r = 2, c = 2 (2 x 2 table) can be represented as in the general format presented in Table 3.1.

Table 3.1

General 2 x 2 contingency table

		Variable A		
		Category 1	Category 2	
Variable B	Category 1	a	b	a+b
	Category 2	c	d	c+d
		a+c	b+d	

For the particular case of 2 x 2 table, expression [3.5] for computing the chi-square statistic reduces to the following simplified form:

[55] The degrees of freedom of the chi-square distribution which approximates the distribution of X^2 (when H_0 is true) is given by the number of independent terms in equation [3.5], given that the row and column marginal totals are fixed, that is d.f.= (r-1)(c-1).

$$\chi^2 = \frac{N(ad - bc)^2}{(a + b)(c + d)(a + c)(b + d)} \quad [3.6]$$

An alternative test procedure to examine the independence of two variables in a multinomial sample is based on the odds ratio. Independence of two variables implies that the odds of variable A being observed at category 1 is the same whether variable B belongs to category 1 or 2. The null hypothesis that two variables are independent implies that the odds ratio:

$$\frac{p_{11}p_{22}}{p_{12}p_{21}} \quad [3.7]$$

equals one. Hence H_0 in expression [3.1] can be equivalently represented as:

$$H_0: \frac{p_{11}p_{22}}{p_{12}p_{21}} = 1 \quad [3.8]$$

For large samples, the log of the estimated odds ratio is normally distributed with standard error:[56]

$$\sigma_{\log(oddsratio)} = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}} \quad [3.9]$$

[56] See CHRISTENSEN [1990, p.40].

3.3. ADJUSTMENTS FOR SMALL EXPECTED FREQUENCIES: YATES CONTINUITY CORRECTION AND FISHER'S EXACT TEST

The chi-square distribution is a continuous probability distribution, being used as an approximation to the discrete probability of observed frequencies in deriving the distribution of the X^2 statistic. When all expected frequencies are small,[57] that approximation may be a poor one. YATES [1934][58] suggests a correction for continuity to improve this approximation. This procedure involves substituting the original frequencies by corrected frequencies[59] in equation [3.6], resulting in:

$$\chi^2 = \frac{N(|ad - bc| - .5N)^2}{(a+b)(c+d)(a+c)(b+d)} \qquad [3.10]$$

which is the chi-square value corrected for discontinuity.

The approximation by the chi-square distribution of the test statistic given in [3.4] assumes that the expected frequencies should not be too small. In the case of 2 x 2 tables with small expected frequencies, Fisher's exact test for a 2 x 2 contingency

57 Typically, the term "small" has been interpreted in the literature in the sense that a satisfactory approximation is achieved when the expected frequencies are five or more.

58 YATES [1934] "Contingency tables involving small numbers and the chi-square test", referenced in EVERITT [1992].

59 The correction, as described in EVERITT [1992] involves subtracting 0.5 from the positive discrepancies, (observed-expected), and adding 0.5 to the negative discrepancies.

table is shown in EVERITT [1992] and MOOD, GRAYBILL and BOES [1974] to be appropriate.

This nonparametric technique does not require the use of the chi-square distribution, and thus is not dependent on any large sample approximations. Instead, it employs the exact probability distribution of the observed frequencies. For fixed marginal totals, the required distribution is easily shown to be that associated with sampling without replacement from a finite population, namely a hypergeometric distribution. If the two variables are independent, the probability of obtaining any particular arrangement of the frequencies a, b, c, and d is (conditioned by the fixed row and column totals):

$$P = \frac{(a + b)!(a + c)!(c + d)!(b + d)!}{a!b!c!d!N!} \qquad [3.10]$$

This expression calculates the exact probability of obtaining any particular arrangement of the frequencies, starting with the smallest of the frequencies and adding the possibly more extreme cell frequencies than those observed.[60] These probabilities are then compared with the chosen significance level. If it is greater than the significance level, the hypothesis of independence is not rejected; otherwise the case for no association of the variables is not shown.

[60] The probability of obtaining any lower frequency is calculated by subtracting 1 from a and from d (if a is the lowest frequency) and adding 1 to c and b, and so on, down to including zero.

Since the Fisher test indicates departure from the null hypothesis in a specific direction,[61] multiplying the Fisher probability by two results in approximately the same probability value as Yates corrected chi-square (in the cases where the sample sizes for each group variable are the same). Anyway, for large samples, both the chi-square, the Yates corrected chi-square and the Fisher test are equivalent.

3.4. CONCLUSIONS

In this chapter we have described the methodology of analysis of contingency tables, which will be used for investigating the performance persistence of mutual funds. In this context, we have presented alternative methods for testing the independence of two qualitative variables (in this case, fund performance over two succeeding periods): the Z-test Repeat Winners, the Odds Ratio log test and the Chi-square test for independence. We also gave special emphasis to the necessary adjustments required in the context of limited sample size, namely the Yates continuity correction and Fisher exact p-value. In the next chapter these tests will be applied to our sample in order to assess the performance persistence of Portuguese equity funds, i.e., a small market.

[61] In contrast with the chi-square test, which assesses departures from the hypothesis in either direction.

CHAPTER 4

PERFORMANCE PERSISTENCE

4.1. INTRODUCTION

As we have seen when revising the literature, a number of recent studies have presented evidence in favour of the performance persistence phenomenon, that is, funds with above median performance in one period will continue to be superior performers in the subsequent period. This type of evidence lends support to the idea that past performance contains information about future performance, which would have implications in several fields of finance.

In this chapter we intend to explore the phenomenon of performance persistence in Portuguese equity funds. After describing the database, and as a background, we first analyze the overall performance of funds. Secondly, and based on the analysis and comparison of a number of criteria, we provide empirical evidence on the performance persistence of our sample of funds based on the methodology of contingency tables (presented in Chapter 3), and see whether the results are consistent with the repeat winner hypothesis. It should be stressed that our approach to the assessment of performance persistence is a multifaceted one, in the sense that we focus on the comparison and discussion of the applicability of various performance persistence criteria. In particular, for all situations we emphasize the pertinent issue concerning the necessary adjustments of the test statistics to a small sample context.

In addition to examining the persistence of returns (unadjusted and adjusted for risk), we also investigate the performance persistence of risk-characteristics. The term structure of performance persistence, that is, whether the persistence measures are a factor of the performance intervals, is also discussed. Finally, we examine performance persistence in terms of each individual fund composing the sample.

4.2. DATA DESCRIPTION

The history of open-end investment companies in Portugal is a very recent one compared to the extensively studied major markets of the U.S. and U.K. Since the first equity fund was launched in 1986, the industry has grown considerably, both in terms of the number of funds and the value of the assets. At the end of 1993, the total value of the 109 existing funds was 1 647 668 million escudos. By March 1998, there were 217 funds with total assets of 4 338 602.6 million escudos. Yet, the equity segment of the market still represents a small part of the industry,[62] despite the high rate of growth it recently experienced. By December 1993 there were 13 national equity funds, representing around 2 percent of the total assets.

[62] The market is dominated by the bond fund and treasury fund sectors, which represent approximately 35 and 23 percent of the total assets invested in mutual funds, respectively (as of March 1998).

Currently there are 26 national equity funds (21 open-end funds and 5 closed-end funds), representing approximately 12 percent of the total number of funds and 10 percent of the total assets (as of March 1998). This growth reflects closely the behaviour of the stock market, in particular the bull characteristics of the most recent years.

It also should be noted that the Portuguese mutual fund industry is very highly concentrated. By March 1998, the five largest companies (out of a total of 19) represented approximately 80 percent of the total assets invested in the sector. Also, relatively to the equity funds, the first five companies account for more than 80 of the total value of assets invested in equity funds. Another characteristic of the market is that mutual fund companies have traditionally been dominated by the banking industry, which obviously conditions the investors to a close relationship with the bank/fund company group.

Our mutual fund database consists of total daily returns on all equity funds over the four year period April 1994 through March 1998. Since equity funds gained popularity only in recent years, only 12 funds with reasonable sample sizes are available for analysis.[63] This choice is a result from the trade off between the number of funds in the sample and the number of observations for each fund, as a larger time period

[63] Of the 13 national equity funds available at the beginning of the sample period, one has changed into an international fund and, therefore, was excluded.

would significantly reduce the number of funds available and vice-versa.

The sample of funds was chosen according to the following criteria:

- Only mutual funds categorized as "national equity" funds by the "*Associação Portuguesa das Sociedades Gestoras de Fundos de Investimento*"[64] (APFIN) were included in the sample.

- Mutual funds investing in foreign stocks were omitted from the sample, since these funds have different risk components, which would imply the inclusion of additional sources of systematic risk outside Portugal. The same type of reasoning also applies to balanced funds.

The sample is therefore composed of funds whose objective is oriented to general equity. Also, in general, all of these funds invest a small proportion in Government securities and in the money market. Table 4.1 reports summary statistics for all funds in the data sample. We assign a unique letter to each fund.

[64] The Portuguese Mutual Fund Association.

Table 4.1

Summary statistics for the mutual funds in the sample

This table shows the composition (in %) of the mutual funds in the sample, as well as their total net assets. Market share is the total net assets of each fund divided by the total net assets of all open-end equity funds. Quantities are expressed in millions of PTE. Data obtained from APFIN.

	Percentage of:				
	Liquidity + Short-term Investments	Public Debt + Bonds	Stocks	Total Net Assets (Mil. PTE)	Market Share %
March-95					
A	39.43		60.08	806.60	4.14
B	29.66	8.43	61.15	866.80	4.44
C	31.94		72.39	3100.40	15.90
D	42.58		59.27	4613.20	23.66
F	27.84		76.06	1182.70	6.06
G	16.98	17.83	64.96	1826.40	9.37
H	21.95	20.03	59.23	937.50	4.81
I	18.01	38.58	43.41	74.60	0.38
J	13.27	17.96	72.37	883.30	4.53
K	28.12	9.44	64.77	320.10	1.64
L	2.74	7.64	54.60	3085.60	15.82
M	22.90	18.04	60.36	896.30	4.60
March-96					
A	20.96	3.89	64.34	456.04	2.21
B	30.41		71.62	888.83	4.31
C	21.76	2.26	78.27	5019.94	24.33
D	13.39		78.44	3882.87	18.82
F	5.14		93.92	1100.54	5.33
G	15.85	20.87	67.38	1615.48	7.83
H	4.04		96.62	860.63	4.17
I	17.18	3.94	80.21	94.57	0.46
J	3.20	14.23	79.12	646.12	3.13
K	8.98	25.43	65.76	228.78	1.11
L	25.54	17.73	62.15	2671.31	12.95
M	10.32		89.56	648.29	3.14
March-97					
A	15.39	1.56	82.93	5987.64	5.75
B	27.11	0.50	75.71	1174.18	1.13
C	16.47		86.01	21864.60	20.99
D	9.34		88.46	11387.12	10.93
F	-1.35		100.18	1984.46	1.91
G	20.20	2.53	79.49	2399.60	2.30
H	12.66		85.39	5286.88	5.08
I	31.73	1.42	79.31	171.41	0.16
J	7.92	7.12	85.63	9087.71	8.72
K	4.32	3.13	85.95	2064.18	1.98
L	20.66	4.55	77.22	8343.32	8.01
M	16.49		85.74	660.28	0.63
March-98					
A	16.76		83.86	17124.42	5.86
B	11.93		90.95	8879.5	3.04
C	9.37		92.03	63785.83	21.82
D	8.41		89.03	51863.4	17.74
F	11.12		90.27	3054.76	1.05
G	4.55	2.08	90.24	5383.16	1.84
H	20.53		92.38	17003.6	5.82
I	11.49	5.35	84.56	406.46	0.14
J	5.13	5.60	84.92	19722.28	6.75
K	14.4	0.11	87.45	8683.76	2.97
L	5.99	2.84	88.75	17303.56	5.92
M	6.21		103.78	1027.49	0.35

Daily data on net asset values was obtained from APFIN and hand-collected from the mutual fund companies themselves. The data was double-checked and found to be accurate. Dividend information and ex-dividend dates were collected from the mutual fund companies.

We constructed total performance returns with fund dividends reinvested at the ex-dividend date net asset value,[65] as follows:

$$R_{p,t} = \ln\left[\left(NAV_{p,t} + D_{p,t}\right)/NAV_{p,t-1}\right] \qquad [4.1]$$

where:

$R_{p,t}$ = Total return for fund **p** in period **t**;

$NAV_{p,t}$ = Net asset value of fund **p** at the end of period **t**;

$D_{p,t}$ = Dividend per unit paid by fund **p** at time **t**.[66]

Because returns were calculated using net asset values and distributions, this measure of total return is net of operating expenses, but gross of any sales charge.[67] No funds were merged

---

[65] In the case of the two funds that are not capitalization funds.

[66] However, several authors use simplified versions of this approach, as in BROWN and GOETZMANN [1995, p.681]:

$$R_t = \frac{\Delta NAV_t}{NAV_{t-1}} + \frac{D_t}{NAV_{t-1}}$$

[67] These returns may not correspond to the "product" delivered to investors, since the possible existence of a load fee is not accounted for in the calculation of net asset values. For our sample of funds, it is a common procedure to charge a redemption fee (which ranges from 1 to 2%) in case the funds are held for short periods of time.

or liquidated during this period, so there is no survivorship bias.

Market returns were also calculated according to expression [4.1], using publicly available indices representing the Portuguese stock market:

- BVL-Geral

  This is a capitalization weighted index of all stocks issued at the "Bolsa de Valores de Lisboa" (BVL), computed since 5 January 1988.

- BVL-30

  This index consists of a capitalization weighted index of 30 stocks issued at the BVL, considered the most representative in terms of market capitalization and liquidity. It has been computed since 4 January 1993.

- PSI-20

  This is a capitalization weighted index consisting of 20 of the largest and most liquid stocks issued at the BVL. Computed since 31 December 1992, it has been designed by the "Bolsa de Derivados do Porto" (BDP) not only to represent the Portuguese stock market, but also to support the derivative contracts.[68]

[68] Unlike the BVL indices, the PSI-20 is not adjusted for dividends, being a price appreciation index rather than a performance index. Therefore, for purposes of this study, we adjusted the PSI-20 for dividends. For reasons of simplicity and to avoid any misunderstanding, from this point forward, wherever we refer to the PSI-20 index, we allude to the PSI-20 adjusted for dividends.

Data on indices was obtained directly from the stock and derivatives exchanges, BVL and BDP, respectively. Of course, these three indices are highly correlated (over 99 percent). Also, both the BVL-30 and the PSI-20 represent approximately 80 percent of the total market capitalization.

The riskfree rate was proxied by the three-month LISBOR (the Lisbon Interbank Offered rate) and collected from the BDP.[69]

4.3. OVERALL PERFORMANCE EVALUATION

As a background, we show the overall performance (initially ignoring risk) of our sample of funds according to different performance criteria and relatively to bear and bull markets.

The mean quarterly returns of each fund are shown in Figure 4.1 and the cumulative returns in Figure 4.2. The returns to the PSI-20 are reported for purposes of comparison.[70]

---

[69] Preliminary results using the 3 month LISBOR or Treasury Bills as proxy for the riskfree rate showed no differences in fund performance.

[70] We give emphasis to the PSI-20 not only because it is the most recently computed index, not yet utilized in performance evaluation studies, but also because its 20 underlying shares constitute around 80 percent of the continuously quoted shares on the market. Also, the Portuguese mutual funds tend to concentrate their holdings on these types of stocks - the largest and most liquid stocks (MACHADO-SANTOS and ARMADA [1996]).

Figure 4.1

Quarterly fund mean returns

Average daily compounded rate of return for each fund and the index over the calendar quarter.

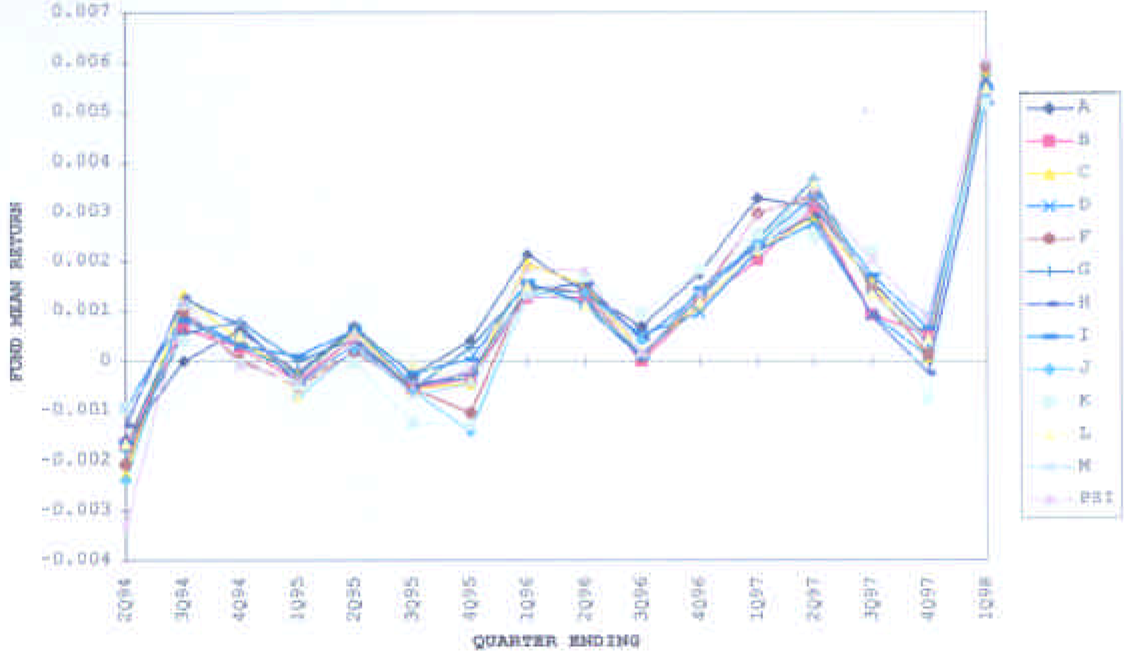
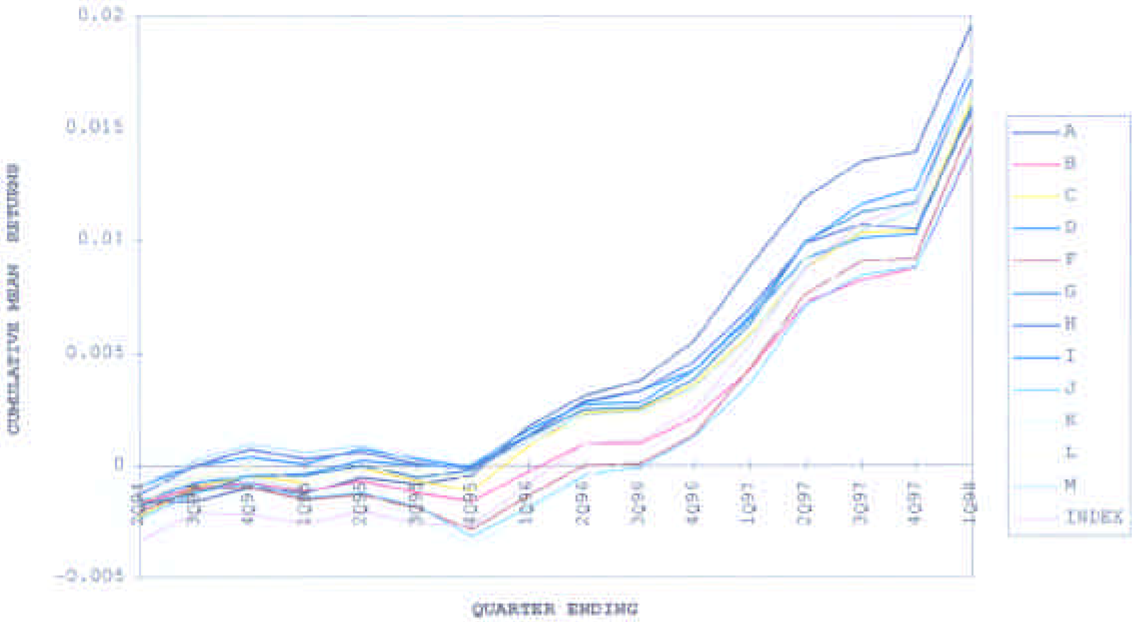


Figure 4.2

Fund cumulative returns (April 1994-March 1998)

Cumulative daily average of the daily continuously compounded rate of return for each fund and the index over the calendar quarter.



During the sample period, only funds A and I outperform the index. The mean overall LISBOR was 8.23% over this entire period, corresponding to 0.000226 per day. Obviously fund performances are more or less correlated, with slightly bear markets in the 2nd quarter 1994, 4th quarter 1994 through the 4th quarter of 1995, and 4th quarter 1997.

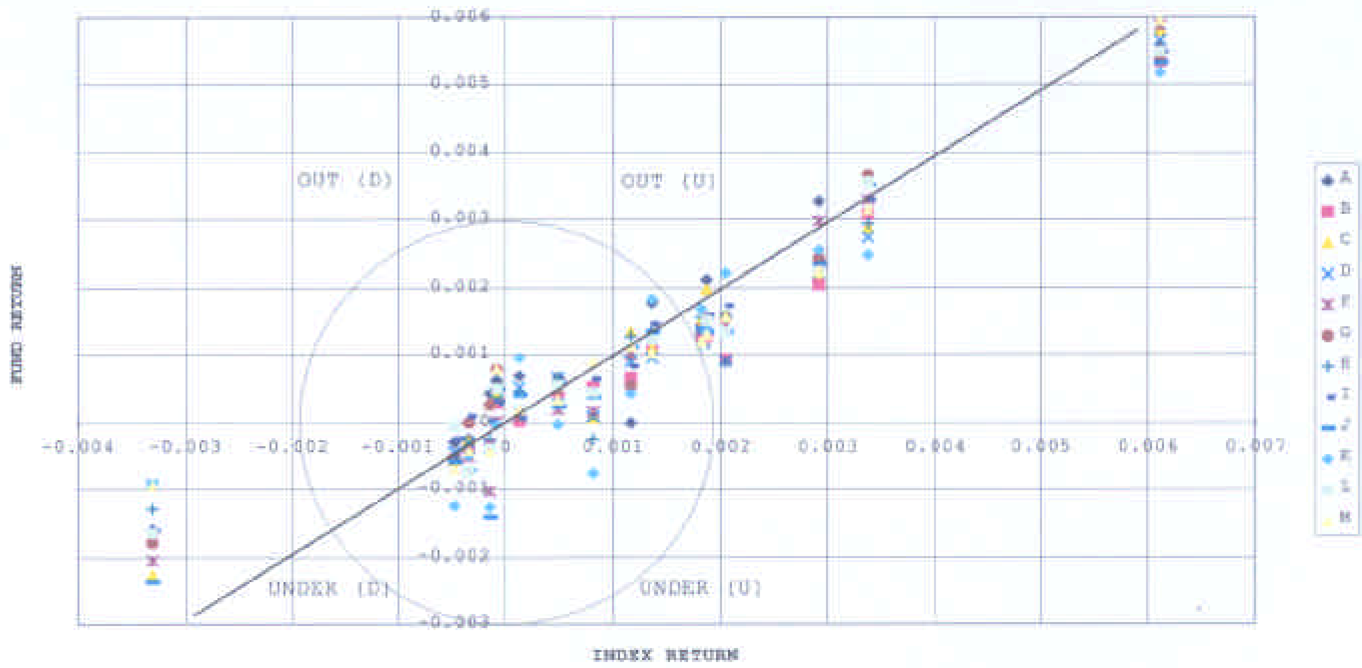
Do equity fund managers outperform the index in bear and/or bull markets?[71] In Figure 4.3, the index quarterly returns have been sorted in ascending order of five bear quarters continuing through eleven bull quarters. Then the fund returns, sorted according to the index return order, have been plotted against the index returns.

[71] We define bull and bear markets with respect to positive or negative index returns, respectively.

Figure 4.3

Fund performance circle

Y axis is average daily rate of return for each fund over the calendar quarter sorted according to the index return order on the X axis. OUT(U) is the quadrant for funds with returns > index return in "up-markets"; UNDER (U) is the quadrant for funds with returns < index return in "up-markets"; OUT (D) is the quadrant for funds with returns < index return in "down markets"; UNDER (D) is the quadrant for funds with return < index returns in "down-markets".



No particular pattern appears to dominate in this picture. Generally there are several quarterly fund returns in excess of index returns in down markets (shown as OUT (D), the approximate Northwest quadrant), which is consistent with the fact that the index is 100% invested in equities and this sample of equity funds includes funds which are sometimes invested in bonds and/or short-term investments (as shown in Table 4.1), obviously during or prior to bear markets, particularly in the 2nd quarter 1994. Also, at high or very high positive index returns, most funds underperform the index. Then, there is a cluster of fund returns around an inner circle during the slightly down market periods.

Thus, the performance circle analysis complements the "traditional" performance analysis in the sense that it considers the extent to which each fund underperforms or overperforms the index returns, as well as the bear and bull characteristics of the measurement period.

We have just seen that, over the sample period, some funds have generated (raw) returns that approximate (or exceed) the index. We now further expand the analysis and estimate fund performance relative to several different measures:

- Excess Returns relative to the riskfree rate;
- Excess returns relative to the market;
- Excess returns relative to a single-index model.

In formal terms, we measure performance as:

$$R_{p,t} - R_{f,t} \qquad [4.2]$$

$$R_{p,t} - R_{m,t} \qquad [4.3]$$

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p\left(R_{m,t} - R_{f,t}\right) + \varepsilon_{p,t} \qquad [4.4]$$

where:

$R_{p,t}$ = Return on fund **p** over period **t**;

$R_{f,t}$ = Return on the riskfree rate in period **t**;

$R_{m,t}$ = Return on the market index over period **t**;

β_p = Sensitivity of the excess return of fund **p** to the market (systematic risk);

α_p = Risk-adjusted excess return (Jensen's alpha).

Table 4.2. presents the overall daily performance for the period April 1994 to March 1998.

Table 4.2

Overall daily performance

This table shows the daily returns in excess of the riskfree rate and the market, as well as estimates of the coefficients obtained from the time series regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon_{p,t}$ for the period April 1994 to March 1998, using the **PSI-20** index as the benchmark portfolio.

Fund	$R_{p,t}-R_{f,t}$	$R_{p,t}-R_{m,t}$	α_p	β_p	R^2
A	0.00099	0.00011	0.000450	0.615	0.744
B	0.00065	-0.00023	0.000088	0.649	0.795
C	0.00079	-0.00010	0.000050	0.852	0.646
D	0.00077	-0.00011	0.000244	0.596	0.736
F	0.00072	-0.00016	0.000054	0.763	0.653
G	0.00084	-0.00004	0.000240	0.675	0.699
H	0.00075	-0.00013	0.000230	0.594	0.608
I	0.00088	0.00000	0.000257	0.715	0.808
J	0.00066	-0.00022	-0.000017	0.771	0.804
K	0.00064	-0.00024	0.000075	0.646	0.682
L	0.00081	-0.00008	0.000181	0.719	0.827
M	0.00085	-0.00003	0.000312	0.616	0.398
Mean	0.00078	-0.00010	0.000180	0.684	0.847

Simply looking at unadjusted returns, we can observe that, on average, during the sample period, mutual funds underperformed the market by 3.65 percent per year.[72] However, when we judge performance in terms of risk adjusted returns, on average mutual funds outperform the market by 6.57 percent per year. The average R^2 was 0.847, indicating that the model, in general, fits the data quite closely.[73]

[72] Although the analysis was conducted on a daily basis, and all tables report the results on a daily basis, we often present results in the text by multiplying the daily data by 365 (approximation to an annual basis).

[73] In fact, when we subdivide the sample period into annual subperiods, we can observe that the R^2 is particularly high (approximately 90 % for most funds) in the last 2 years. In Appendix 4.1 we present the regression estimates for individual years.

In order to test the sensitivity of fund performance to different indices, we estimated risk-adjusted returns relative to the BVL-30 and BVL-Geral. In addition, we addressed the question regarding the statistical significance of the estimated performance measures.

Table 4.3 presents estimates and statistical significance of the risk-adjusted returns for alternative market indices. Although almost all funds exhibit positive alphas, only for five/three funds is it statistically different from zero (at least at the 5 percent level),[74] using the PSI-20 and the BVL indices, respectively. Only one fund shows very strong evidence of superior performance (fund A), being statistically significant at the 1 percent level[75] (for either benchmark considered). It is also surprising to observe that we do not find any fund with statistically significant negative performance for the overall period.

---

[74] The critical t-statistic, $t_{critical\ (981,0.025)} \approx 1.962$.

[75] The critical t-statistic, $t_{critical\ (981,0.005)} \approx 2.581$.

Table 4.3.

Jensen's alpha for alternative market indices

This tables shows estimates of the coefficients obtained from the time series regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon_{p,t}$ computed for the four year period April 1994 to March 1998.

	Index	α_p	t-stat	β_p	t-stat	R^2
A	PSI-20	0.000450	4.690 **	0.615	53.369	0.744
	BVL-30	0.000406	4.363 **	0.648	55.574	0.759
	BVL-G	0.000397	4.172 **	0.702	53.930	0.748
B	PSI-20	0.000088	1.000	0.649	61.674	0.795
	BVL-30	0.000042	0.502	0.683	64.457	0.809
	BVL-G	0.000036	0.404	0.737	60.870	0.791
C	PSI-20	0.000050	0.300	0.852	42.302	0.646
	BVL-30	-0.000016	-0.101	0.905	44.567	0.669
	BVL-G	-0.000032	-0.199	0.984	44.188	0.666
D	PSI-20	0.000244	2.573 *	0.596	52.275	0.736
	BVL-30	0.000197	2.194 *	0.633	56.348	0.764
	BVL-G	0.000186	2.049 *	0.689	55.597	0.759
F	PSI-20	0.000054	0.363	0.763	42.966	0.653
	BVL-30	0.000006	0.043	0.797	43.059	0.654
	BVL-G	-0.000002	-0.015	0.861	41.813	0.641
G	PSI-20	0.000240	2.038 *	0.675	47.687	0.699
	BVL-30	0.000191	1.669	0.713	49.639	0.715
	BVL-G	0.000163	1.503	0.794	53.586	0.745
H	PSI-20	0.000230	1.819	0.594	38.995	0.608
	BVL-30	0.000185	1.498	0.630	40.633	0.627
	BVL-G	0.000172	1.394	0.687	40.618	0.627
I	PSI-20	0.000257	2.768 **	0.715	64.151	0.808
	BVL-30	0.000207	2.317 *	0.752	67.152	0.821
	BVL-G	0.000183	2.176 *	0.831	72.218	0.842
J	PSI-20	-0.000017	-0.171	0.771	63.491	0.804
	BVL-30	-0.000070	-0.711	0.810	65.924	0.816
	BVL-G	-0.000077	-0.752	0.873	62.056	0.797
K	PSI-20	0.000075	0.639	0.646	45.878	0.682
	BVL-30	0.000024	0.216	0.686	48.626	0.707
	BVL-G	0.000004	0.035	0.757	50.370	0.721
L	PSI-20	0.000181	2.075 *	0.719	68.580	0.827
	BVL-30	0.000134	1.569	0.754	70.528	0.835
	BVL-G	0.000126	1.402	0.814	66.276	0.817
M	PSI-20	0.000312	1.550	0.616	25.491	0.398
	BVL-30	0.000261	1.317	0.657	26.445	0.416
	BVL-G	0.000250	1.258	0.714	26.266	0.413

two-tail test for alpha:

* Statistically significant at the 5 percent level

* * Statistically significant at the 1 percent level

Although the risk-adjusted returns are slightly changed according to which index is used, the choice of the index does not seem to affect the relative performance of the funds. In fact, Spearman's Rank Correlation coefficient[76] shows clearly that the fund rankings by performance are not affected by the use of alternative market indices, as reported in Table 4.4.

Table 4.4

Correlation among indices

This table reports Spearman's Rank Correlation Coefficient between the rankings of fund performance according to alternative market indices.

	PSI	BVL-30	BVL-Geral
PSI			
BVL-30	1 *		
BVL-Geral	0.986 *	0.986 *	

* Statistically significant at the 1 percent level

The results, however, require careful interpretation. One of the assumptions which provides the basis for deriving the sampling distributions for the coefficient estimators (which, in turn, are used for making statistical inferences) is that the residuals are homoscedastic. Violation of this assumption has potentially serious implications on inferences based on the results of the least squares [GREENE, p. 391].

[76] This statistic is defined in Chapter 5, section 5.2.

In fact, the application of ordinary least squares (OLS) ignoring the possibility of a non-constant disturbance variance would lead to estimators which are unbiased and consistent but no longer efficient. Other estimation methods producing unbiased estimators will have smaller sampling variances. In this context, hypothesis tests based on (inefficient) OLS standard errors are not valid. The possibility of serial correlation of the disturbances poses an identical problem, as it also causes loss of efficiency of the least squares estimators.

As only rarely one can be certain that the residuals are not heteroscedastic and autocorrelated, we first perform tests conceived to detect these problems, rather than just assuming they are present.[77]

To test for heteroscedasticity we used the statistical test developed by WHITE [1980].[78] The major advantage of this test is that it does not require knowledge about the nature of the heteroscedasticity. To test for the possibility of autocorrelated disturbances, the Durbin-Watson test is, by far, one of the most widely used tests. However it has two inherent problems: besides the difficulty resulting of the inconclusive region, it is not appropriate if the process is

---

77 This is the approach suggested by GREENE [1993]. Nevertheless, in many performance studies the results are corrected for heteroscedasticity and autocorrelation without actually testing whether or not they are present (e.g. VOLKMAN AND WOHAR [1995]).

78 See Appendix 4.2 for a description of White's methodology.

not first-order autoregressive AR(1).[79] A less restrictive approach to test for serial correlation is the BREUSCH [1978] - GODFREY [1978] Lagrange multiplier test, applicable for AR(p) processes, where p can be specified for any positive order.[80]

The results for the White and Breusch-Godfrey tests for heteroscedasticity and autocorrelation of the regression disturbances are summarized in Table 4.5.

Table 4.5

Summary of the heteroscedasticity and autocorrelation tests

This tables shows the results of the White [1980] test for heteroscedasticy and the Breush [1978] - Godfrey [1978] Lagrange multiplier test for autocorrelation. Whenever the initials H and A appear, it represents a rejection of the null hypothesis of homoscedasticity and independence of the regression disturbances, at the respective significance level.

	1%			5%			10%		
Funds	PSI	BVL-30	BVL-G	PSI	BVL-30	BVL-G	PSI	BVL-30	BVL-G
A	H,A	H,A	H,A	H,A	H,A	H,A	H,A	H,A	H,A
B	H,A	H	H	H,A	H	H	H,A	H	H
C	H,A	H,A	H,A	H,A	H,A	H,A	H,A	H,A	H,A
D	H,A	H,A	H	H,A	H,A	H,A	H,A	H,A	H,A
F	A	A	A	A	A	A	A	A	A
G	H	H	H	H	H	H	H	H	H
H	H,A	H,A	H,A	H,A	H,A	H,A	H,A	H,A	H,A
I	H	H,A	H,A	H,A	H,A	H,A	H,A	H,A	H,A
J	H	H	H	H	H	H	H A	H	H
K	H,A	H,A	H,A	H,A	H,A	H,A	H,A	H,A	H,A
L	H	H	H	H	H	H,A	H	H,A	H,A
M	A	A	A	H,A	A	A	H,A	H,A	A

[79] I.e., of the form:

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t ,$$

where $E[u_t] = 0$, $E[u_t^2] = \sigma_u^2$, and $cov[u_t, u_s] = 0$ for every $t \neq s$. For a more detailed discussion of the limitations of the Durbin-Watson statistic see JOHNSTON [1987] and GREENE [1993].

[80] See Appendix 4.3 for a description of the Breusch-Godfrey Lagrange multiplier test.

From the analysis of the table it is evident that, for the majority of funds, we can reject both the hypothesis of homoscedasticity and independence of residuals. In such a case, inference based on least squares can be affected by inefficient standard errors; so, a mechanism for correction is required.

To correct for inefficiencies in the standard errors, we use an alternative method for determining appropriate estimators of the variance of least squares estimators. Such a method has been developed by WHITE [1980], who derived a heteroscedasticity-consistent variance co-variance matrix for calculating standard errors (and t-statistics), briefly described next.[81]

In the presence of heteroscedasticity, the conventional estimated variance matrix for the least squares estimator, $\sigma^2(\mathbf{X'X})^{-1}$ is inappropriate;[82] the appropriate matrix of **b** is:

$$\left(\mathbf{X'}\,\mathbf{X}^{-1}\right)\left[\mathbf{X'}\left(\sigma^2\Omega\right)\mathbf{X}\right](\mathbf{X'}\,\mathbf{X})^{-1} \qquad [4.5]$$

where Ω is a matrix with $\mathrm{var}(\varepsilon_i) = \sigma_i^2$ in the diagonal and zeros elsewhere [GREENE, 1993].

[81] BREEN, JAGANNATHAN and OFER [1986] emphasize the need to correct for heteroscedasticity and document the adequacy of White's correction.

[82] Recall the general form of the regression model in matrix notation is:

$$\mathbf{y} = \mathbf{X}\,b + e$$

where X is the (n observations * k variables) matrix: $\begin{bmatrix} X_1 \\ \cdot \\ \cdot \\ \cdot \\ X_n \end{bmatrix}$.

Although estimates of $\sigma^2\Omega$ are unavailable, it is possible to estimate the covariance matrix through :

$$\Sigma = \frac{\sigma^2 \mathbf{X}' \Omega \mathbf{X}}{n} = \frac{1}{n}\sum_{i} \sigma_i^2 \mathbf{x}_i \mathbf{x}_i' \qquad [4.6]$$

where **n** is the number of observations.

WHITE [1980] has shown that under general conditions, the matrix

$$\mathbf{S}_0 = \frac{1}{n}\sum_{I} e_i^2 \mathbf{x}_i \mathbf{x}_i' \qquad [4.7]$$

where e_i is the ith least squares residual, is a consistent estimator of **Σ.** Therefore, the White covariance matrix:

$$\text{Var}[\mathbf{b}] = n(\mathbf{X}'\mathbf{X})^{-1}\mathbf{S}_0(\mathbf{X}'\mathbf{X})^{-1} \qquad [4.8]$$

can be used as an estimate of the true variance of the least squares estimator.

This result is of extreme importance, as it allows for appropriate inferences based on the least squares coefficients even without actually specifying the type of heteroscedasticity (which is probably unknown in most cases). It also constitutes an advantage over alternative estimation techniques which require detailed formulation of **Ω**.

We also address the problem of autocorrelated disturbances. Similarly to heteroscedasticity, autocorrelation also questions the validity of inferences based on the least squares. As before, the presence of autocorrelation does not affect the unbiasedness and consistency of OLS,[83] but causes loss of efficiency. As usual, the conventional estimated variance matrix is an inappropriate estimator of $\sigma^2(\mathbf{X'X})^{-1}(\mathbf{X\Omega X})(\mathbf{X'X})^{-1}$. In this context, an appropriate estimator of the variance matrix of the least squares is necessary. In the case where the form of the autocorrelation is unknown (estimates of the parameter Ω unavailable for direct computation), there is an approach which parallels the use of the White estimator for heteroscedasticity.[84] Following White's suggestion for heteroscedasticity, NEWEY and WEST [1987] proposed an estimator for autocorrelated disturbances with an unspecified structure:

$$S_* = \mathbf{S}_0 + \frac{1}{T}\sum_{j=1}^{L}\sum_{t=j+1}^{T} w_j e_t e_{t-j}\left(\mathbf{x}_t\mathbf{x}'_{t-j} + \mathbf{x}_{t-j}\mathbf{x}'_t\right) \qquad [4.9]$$

where $\mathbf{S}_0$ is the White estimator in [4.7] and

$$w_j = 1 - \frac{j}{L+1} \qquad [4.10]$$

[83] Unless one of the regressors is a lagged variable.

[84] Which presumes that the residuals of the estimated regression are serially uncorrelated.

L, the truncation lag, is a parameter representing the number of autocorrelations used to approximate the dynamics of the residuals e_t. Although this test is considered reasonably powerful [GREENE, 1993], it presents the difficulty related to the choice of L, the truncation lag. Following the suggestion of NEWEY and WEST [1987], it has been set as $4(T/100)^{2/9}$.[85]

The procedure above described allows for correct standard errors in the presence of serial correlation and heteroscedasticity by modifying the traditional OLS variance covariance matrix with a weighting matrix such that the weights are a function of the moving average terms.

In light of the above exposed, we will evaluate (and reinterpret) the significance of the estimates based on a positive-definite consistent estimate of the variance-covariance matrix as devised by WHITE [1980] (correction for heteroscedasticity) and NEWEY and WEST [1987] (correction for heteroscedasticity and serial correlation).

The statistical significance of the estimates for the OLS, the White and the Newey and West regressions are presented in Table 4.6.[86]

[85] In terms of practical application, we used L = 3,4,5,and 6 for the cases of four-year, two-year one-year and quarterly periods, respectively.

[86] The standard errors of $\hat{\alpha}_p$ and p-values, as well as statistical significance of the b coefficient are shown in Appendix 4.4.

Table 4.6

Summary of the OLS estimates and statistical significance tests corrected for heteroscedasticity (WHITE [1980]) and autocorrelation and heteroscedasticity (NEWEY and WEST [1987])

This table the shows estimates of alpha (risk-adjusted returns) from the regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p\ (R_{m,t} - R_{f,t}) + \varepsilon_{p,t}$ computed for the four-year performance period and for different indices. OLS t-statistics, reported in parenthesis below the estimates of Jensen's alpha, are inefficient because of heteroscedasticity and autocorrelation. To correct for heteroscedastic and both autocorrelated and heteroscedastic error terms, we compute t-statistics obtained by using the WHITE [1980] and NEWEY and WEST procedure [1987].

		Index		
Fund		PSI	BVL-30	BVL-Geral
A	Regression alpha	0.000450	0.000406	0.000397
	OLS t-stat	(4.69) **	(4.36) **	(4.17) **
	White t-stat	(4.79) **	(4.56) **	(4.40) **
	Newey-West t-stat	(3.89) **	(3.80) **	(3.75) **
B	Regression alpha	0.000088	0.000042	0.000036
	OLS t-stat	(1.00)	(0.50)	(0.40)
	White t-stat	(1.00)	(0.51)	(0.42)
	Newey-West t-stat	(0.90)	(0.47)	(0.39)
C	Regression alpha	0.000050	-0.000016	-0.000032
	OLS t-stat	(0.30)	(-0.10)	(-0.20)
	White t-stat	(0.26)	(-0.09)	(-0.17)
	Newey-West t-stat	(0.30)	(-0.10)	(-0.20)
D	Regression alpha	0.000244	0.000197	0.000186
	OLS t-stat	(2.57) *	(2.19) *	(2.05) *
	White t-stat	(2.21) *	(1.90)	(1.84)
	Newey-West t-stat	(2.00) *	(1.75)	(1.74)
F	Regression alpha	0.000054	0.000006	-0.000002
	OLS t-stat	(0.36)	(0.04)	(-0.01)
	White t-stat	(0.35)	(0.04)	(-0.01)
	Newey-West t-stat	(0.48)	(0.06)	(-0.02)
G	Regression alpha	0.000240	0.000191	0.000163
	OLS t-stat	(2.04) *	(1.67)	(1.50)
	White t-stat	(2.01) *	(1.67)	(1.54)
	Newey-West t-stat	(2.03) *	(1.75)	(1.64)
H	Regression alpha	0.000230	0.000185	0.000172
	OLS t-stat	(1.82)	(1.50)	(1.39)
	White t-stat	(1.73)	(1.44)	(1.37)
	Newey-West t-stat	(1.67)	(1.44)	(1.40)
I	Regression alpha	0.000257	0.000207	0.000183
	OLS t-stat	(2.77) **	(2.32) *	(2.18) *
	White t-stat	(2.61) **	(2.22) *	(2.11) *
	Newey-West t-stat	(2.61) **	(2.32) *	(2.20) *
J	Regression alpha	-0.000017	-0.000070	-0.000077
	OLS t-stat	(-0.17)	(-0.71)	(-0.75)
	White t-stat	(-0.17)	(-0.72)	(-0.76)
	Newey-West t-stat	(-0.16)	(-0.68)	(-0.72)
K	Regression alpha	0.000075	0.000024	0.000004
	OLS t-stat	(0.64)	(0.22)	(0.04)
	White t-stat	(0.61)	(0.21)	(0.03)
	Newey-West t-stat	(0.51)	(0.17)	(0.03)
L	Regression alpha	0.000181	0.000134	0.000126
	OLS t-stat	(2.08) *	(1.57)	(1.40)
	White t-stat	(2.06) *	(1.58)	(1.46)
	Newey-West t-stat	(2.08) *	(1.69)	(1.57)
M	Regression alpha	0.000312	0.000261	0.000250
	OLS t-stat	(1.55)	(1.32)	(1.26)
	White t-stat	(1.42)	(1.20)	(1.16)
	Newey-West t-stat	(2.07) *	(1.85)	(1.82)
Average	Regression alpha	0.000180	0.000131	0.000117
	OLS t-stat	(2.34) *	(1.82)	(1.61)
	White t-stat	(2.15) *	(1.70)	(1.55)
	Newey-West t-stat	(1.95)	(1.60)	(1.51)

* Statistically significant at the 5 percent level
** Statistically significant at the 1 percent level

We can verify from the previous table that the significance results are not substantially altered when we correct for heteroscedasticity and autocorrelation. In fact, this is not surprising since for most cases the $\hat{\alpha}_p$'s are not statistically different from zero. However, in general, the t-statistics are smaller when these corrections are performed, as expected.

A closer analysis reveals that of those funds with conventional statistically significant coefficients, only fund D is no longer significant after the correction (for indices BVL-30 and BVL-Geral). Funds A and I are "immune" to correction for heteroscedasticity and autocorrelation, in the sense that their risk-adjusted returns continue to be statistically different from zero. These corrections also show that, on average, funds do not outperform the risk-adjusted benchmark, once more confirming our concern as to having incorrectly rejected the null hypothesis of $\hat{\alpha}_p=0$.

In general, most funds do not exhibit statistically significant alphas during the sample period. The only consistent evidence of superior performance is observed when the PSI-20 index is used as benchmark. However, we must note two cautions. First, the literature suggests that these findings are likely to be influenced by survivorship bias. Yet, to our best knowledge, and due to the emerging characteristics of the Portuguese equity fund industry, no fund disappeared or merged during the sample period. Second, the CAPM framework may not have adjusted appropriately for risk. In fact, and although the explanatory power of CAPM is in general high (particularly in the recent years), multi-index models have been shown to be more sensitive benchmarks (as discussed in the review of the literature). Still, the unavailability of style indices for the Portuguese stock market and for the entire sample period has raised difficulties to the application of multifactor models. In fact, and until very recently, we cannot consider that the Portuguese capital market has had a dimension that could justify its division into distinct style subsegments (such as value versus growth).[87] Also, we have implicitly adjusted for style by considering funds that are oriented towards the same objectives (general equity), and probably will fall into the same style category.

4.4. PERFORMANCE PERSISTENCE: EMPIRICAL EVIDENCE

4.4.1. Quarterly contingency tables of excess returns

The conclusions about the performance of funds during the period 1994-1998 are more consistent with the original

[87] The only style indices found are value and growth indices, computed since 31 December 1996 and a small capitalization index, computed since 1 December 1997, both made available by Morgan Stanley & Co. International.

findings of JENSEN [1968] than those of IPPOLITO [1989][88] in the sense that, on average, the positive alphas are small and insignificant. Nevertheless, it is possible that some funds are consistently superior performers. Recent literature (reviewed and discussed in Chapter 2) has presented extensive evidence in favour of this "hot hands" phenomenon, alleging that if a fund manager possesses superior information or management skills, the performance measure relatively to that fund should continue to be superior. As Gruber states: *"The surprising thing about persistence is not that it exists, but rather how strong it appears to be."* [GRUBER, 1996, p.793].

In this section, we will test the hypothesis of no performance persistence of our sample of Portuguese mutual funds. We examine this claim applying the methodology based on the construction of two-way contingency tables (as in MALKIEL [1995] and BROWN and GOETZMANN [1995]).

Table 4.7 reports the quarterly mean excess returns for our sample of equity funds.

[88] IPPOLITO [1989] found that mutual funds earn gross returns sufficient to cover their expenses. Such a result would be consistent with the view of GROSSMAN and STIGLITZ [1980] on market efficiency (in the sense that it includes some compensation for the costs of information gathering analysis).

Table 4.7

Frequency of repeat performers: excess returns

WW = funds with average excess returns > median, repeated subsequent quarter; LL = funds with average excess returns < median, repeated subsequent quarter; LW and WL correspond to funds with relative return reversals. Excess returns are measured as the quarterly average of daily return less daily riskfree rate.

	A	B	C	D	F	G	H	I	J	K	L	M	MEDIAN
2Q94	-0.00195	-0.00202	-0.00262	-0.00127	-0.00240	-0.00214	-0.00164	-0.00196	-0.00272	-0.00129	-0.00200	-0.00129	-0.00198
3Q94	-0.00033	0.00033	0.00102	0.00059	0.00069	0.00023	0.00095	0.00051	0.00081	0.00010	0.00080	0.00080	0.00064
4Q94	0.00033	-0.00001	0.00013	0.00006	-0.00026	0.00051	0.00044	0.00003	0.00022	-0.00033	0.00022	0.00050	0.00018
1Q95	-0.00055	-0.00068	-0.00057	-0.00059	-0.00081	-0.00031	-0.00075	-0.00021	-0.00101	-0.00079	-0.00102	-0.00071	-0.00070
2Q95	0.00039	0.00017	0.00032	0.00039	-0.00009	0.00018	0.00006	0.00032	-0.00005	-0.00033	0.00026	0.00005	0.00017
3Q95	-0.00055	-0.00082	-0.00081	-0.00069	-0.00080	-0.00079	-0.00076	-0.00050	-0.00085	-0.00150	-0.00032	-0.00096	-0.00080
4Q95	0.00017	-0.00061	-0.00068	-0.00061	-0.00127	0.00001	-0.00049	-0.00023	-0.00167	-0.00151	-0.00063	-0.00070	-0.00062
1Q96	0.00190	0.00105	0.00175	0.00122	0.00128	0.00130	0.00122	0.00137	0.00112	0.00113	0.00130	0.00108	0.00125
2Q96	0.00117	0.00108	0.00130	0.00118	0.00118	0.00098	0.00135	0.00092	0.00119	0.00147	0.00092	0.00099	0.00118
3Q96	0.00049	-0.00018	-0.00001	0.00033	-0.00012	-0.00009	0.00024	-0.00014	0.00021	0.00076	0.00000	-0.00002	-0.00001
4Q96	0.00159	0.00097	0.00092	0.00078	0.00107	0.00107	0.00116	0.00128	0.00116	0.00164	0.00103	0.00084	0.00107
1Q97	0.00311	0.00188	0.00204	0.00205	0.00281	0.00227	0.00212	0.00219	0.00217	0.00238	0.00207	0.00208	0.00215
2Q97	0.00293	0.00292	0.00273	0.00259	0.00314	0.00350	0.00277	0.00314	0.00335	0.00231	0.00340	0.00300	0.00297
3Q97	0.00144	0.00078	0.00144	0.00077	0.00130	0.00118	0.00074	0.00157	0.00119	0.00207	0.00118	0.00140	0.00124
4Q97	0.00031	0.00039	-0.00008	-0.00005	0.00001	0.00030	-0.00039	0.00051	0.00024	-0.00090	0.00031	0.00077	0.00027
1Q98	0.00545	0.00522	0.00568	0.00553	0.00575	0.00528	0.00506	0.00536	0.00520	0.00505	0.00537	0.00583	0.00536
TOTAL MEAN	0.00099	0.00065	0.00079	0.00077	0.00072	0.00084	0.00075	0.00088	0.00066	0.00064	0.00081	0.00085	0.00078

REPEATS		A	B	C	D	F	G	H	I	J	K	L	M	SUM	TOTAL
1994 3Q	WW	0	0	0	0	0	0	1	0	0	0	0	1	2	
	LW	0	0	1	0	1	0	0	0	1	0	1	0	4	
	WL	1	0	0	1	0	0	0	1	0	1	0	0	4	
	LL	0	1	0	0	0	1	0	0	0	0	0	0	2	12
1994 4Q	WW	0	0	0	0	0	0	1	0	1	0	1	1	4	
	LW	1	0	0	0	0	1	0	0	0	0	0	0	2	
	WL	0	0	1	0	1	0	0	0	0	0	0	0	2	
	LL	0	1	0	1	0	0	0	1	0	1	0	0	4	12
1995 1Q	WW	1	0	0	0	0	1	0	0	0	0	0	0	2	
	LW	0	1	1	1	0	0	0	1	0	0	0	0	4	
	WL	0	0	0	0	0	0	1	0	1	0	1	1	4	
	LL	0	0	0	0	1	0	0	0	0	1	0	0	2	12
1995 2Q	WW	1	0	1	1	0	1	0	1	0	0	0	0	5	
	LW	0	0	0	0	0	0	0	0	0	0	1	0	1	
	WL	0	1	0	0	0	0	0	0	0	0	0	0	1	
	LL	0	0	0	0	1	0	1	0	1	1	0	1	5	12
1995 3Q	WW	1	0	0	1	0	1	0	1	0	0	1	0	5	
	LW	0	0	0	0	0	0	1	0	0	0	0	0	1	
	WL	0	0	1	0	0	0	0	0	0	0	0	0	1	
	LL	0	1	0	0	1	0	0	0	1	1	0	1	5	12
1995 4Q	WW	1	0	0	1	0	1	1	1	0	0	0	0	5	
	LW	0	1	0	0	0	0	0	0	0	0	0	0	1	
	WL	0	0	0	0	0	0	0	0	0	0	1	0	1	
	LL	0	0	1	0	1	0	0	0	1	1	0	1	5	12
1996 1Q	WW	1	0	0	0	0	1	0	1	0	0	0	0	3	
	LW	0	0	1	0	1	0	0	0	0	0	1	0	3	
	WL	0	1	0	1	0	0	1	0	0	0	0	0	3	
	LL	0	0	0	0	0	0	0	0	1	1	0	1	3	12
1996 2Q	WW	0	0	1	0	1	0	0	0	0	0	0	0	2	
	LW	0	0	0	1	0	0	1	0	1	1	0	0	4	
	WL	1	0	0	0	0	1	0	1	0	0	1	0	4	
	LL	0	1	0	0	0	0	0	0	0	0	0	1	2	12
1996 3Q	WW	0	0	0	1	0	0	1	0	1	1	0	0	4	
	LW	1	0	0	0	0	0	0	0	0	0	1	0	2	
	WL	0	0	1	0	1	0	0	0	0	0	0	0	2	
	LL	0	1	0	0	0	1	0	1	0	0	0	1	4	12
1996 4Q	WW	1	0	0	0	0	0	1	0	1	1	0	0	4	
	LW	0	0	0	0	0	1	0	1	0	0	0	0	2	
	WL	0	0	0	1	0	0	0	0	0	0	1	0	2	
	LL	0	1	1	0	1	0	0	0	0	0	0	1	4	12
1997 1Q	WW	1	0	0	0	0	1	0	1	1	1	0	0	5	
	LW	0	0	0	0	1	0	0	0	0	0	0	0	1	
	WL	0	0	0	0	0	0	1	0	0	0	0	0	1	
	LL	0	1	1	1	0	0	0	0	0	0	1	1	5	12
1997 2Q	WW	0	0	0	0	1	1	0	1	1	0	0	0	4	
	LW	0	0	0	0	0	0	0	0	0	0	1	1	2	
	WL	1	0	0	0	0	0	0	0	0	1	0	0	2	
	LL	0	1	1	1	0	0	1	0	0	0	0	0	4	12
1997 3Q	WW	0	0	0	0	1	0	0	1	0	0	0	1	3	
	LW	1	0	1	0	0	0	0	0	0	1	0	0	3	
	WL	0	0	0	0	0	1	0	0	1	0	1	0	3	
	LL	0	1	0	1	0	0	1	0	0	0	0	0	3	12
1997 4Q	WW	1	0	0	0	0	0	0	1	0	0	0	1	3	
	LW	0	1	0	0	0	1	0	0	0	0	1	0	3	
	WL	0	0	1	0	1	0	0	0	0	1	0	0	3	
	LL	0	0	0	1	0	0	1	0	1	0	0	0	3	12
1998 1Q	WW	1	0	0	0	0	0	0	0	0	0	1	1	3	
	LW	0	0	1	1	1	0	0	0	0	0	0	0	3	
	WL	0	1	0	0	0	1	0	1	0	0	0	0	3	
	LL	0	0	0	0	0	0	1	0	1	1	0	0	3	12
CONSISTENCY OF INDIVIDUAL FUND PERFORMANCE															
1994-1997	WW	9	0	2	4	3	7	5	8	5	3	3	5	54	
	LW	3	3	5	3	4	3	2	2	2	2	6	1	36	
	WL	3	3	4	3	3	3	3	3	2	3	5	1	36	
	LL	0	9	4	5	5	2	5	2	6	7	1	8	54	180

These quarterly mean returns have been reformatted for the purpose of constructing persistence tables. The winners (defined as having excess returns greater than the median) in the initial quarter are divided into winners/losers for the subsequent quarter. The winners in one quarter that remain winners in the following quarter are denoted WW, etc., and the sum of repeats is in the last column. In the "Consistency of individual fund performance" the total repeats in each category are shown for each fund. Clearly fund A is the most consistent repeat winner, and fund B the most consistent repeat loser.

In table 4.8 we report the contingency table of winners and losers for each quarter. The null hypothesis that there is no relationship in performance across periods was tested following the criteria of MALKIEL [1995] (*z*-test for repeat winners), BROWN and GOETZMANN [1995] (test statistic for the Odds Ratio) and KHAN and RUDD [1995] (chi-square test of independence).

Table 4.8

Contingency table of excess returns

WW = funds with average excess returns > median, repeated subsequent quarter; LL = funds with average excess returns < median, repeated subsequent quarter; LW and WL correspond to funds with relative return reversals. Excess returns are measured as the quarterly average of daily return less daily riskfree rate. We report performance persistence statistics according to various criteria:
Percent Repeat Winners = WW / (WW+WL);

$$\text{Z-test Repeat Winners} = \left(WW - 0.5 * (WW + WL)\right) / \sqrt{(WW + WL) * 0.5 * 0.5};$$

$$\text{Odds Ratio} = (WW*LL) / (LW*WL);$$

$$\text{Z-statistic} = \left(\log(\text{OddsRatio})\right)/\sigma_{\log(\text{OddsRatio})};$$

$$\sigma_{\log(\text{OddsRatio})} = \sqrt{\frac{1}{WW} + \frac{1}{LW} + \frac{1}{WL} + \frac{1}{LL}};$$

$$\text{Chi-square} = \sum (WW - N/4)^2 \text{ for all WW, LW, WL LL;}$$

P-values are based on two-tail tests.

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	PERCENT REPEAT L	MALKIEL Z-TEST L	p-value
3Q94	2	4	4	2	0.33	-0.82	0.414	0.33	-0.82	0.41
4Q94	4	2	2	4	0.67	0.82	0.414	0.67	0.82	0.41
1Q95	2	4	4	2	0.33	-0.82	0.414	0.33	-0.82	0.41
2Q95	5	1	1	5	0.83	1.63	0.102	0.83	1.63	0.10
3Q95	5	1	1	5	0.83	1.63	0.102	0.83	1.63	0.10
4Q95	5	1	1	5	0.83	1.63	0.102	0.83	1.63	0.10
1Q96	3	3	3	3	0.50	0.00	1.000	0.50	0.00	1.00
2Q96	2	4	4	2	0.33	-0.82	0.414	0.33	-0.82	0.41
3Q96	4	2	2	4	0.67	0.82	0.414	0.67	0.82	0.41
4Q96	4	2	2	4	0.67	0.82	0.414	0.67	0.82	0.41
1Q97	5	1	1	5	0.83	1.63	0.102	0.83	1.63	0.10
2Q97	4	2	2	4	0.67	0.82	0.414	0.67	0.82	0.41
3Q97	3	3	3	3	0.50	0.00	1.000	0.50	0.00	1.00
4Q97	3	3	3	3	0.50	0.00	1.000	0.50	0.00	1.00
1Q98	3	3	3	3	0.50	0.00	1.000	0.50	0.00	1.00
TOTAL	54	36	36	54	0.60	1.90	0.058	0.60	1.90	0.06

	WW	LW	WL	LL	ODDS RATIO	B&G Z-STAT	σ (LOG)	p-value	CHI-SQ	p-value
3Q94	2	4	4	2	0.25	-1.13	1.22	0.258	1.333	0.248
4Q94	4	2	2	4	4.00	1.13	1.22	0.258	1.333	0.248
1Q95	2	4	4	2	0.25	-1.13	1.22	0.258	1.333	0.248
2Q95	5	1	1	5	25.00	2.08 *	1.55	0.038	5.333 *	0.021
3Q95	5	1	1	5	25.00	2.08 *	1.55	0.038	5.333 *	0.021
4Q95	5	1	1	5	25.00	2.08 *	1.55	0.038	5.333 *	0.021
1Q96	3	3	3	3	1.00	0.00	1.15	1.000	0.000	1.000
2Q96	2	4	4	2	0.25	-1.13	1.22	0.258	1.333	0.248
3Q96	4	2	2	4	4.00	1.13	1.22	0.258	1.333	0.248
4Q96	4	2	2	4	4.00	1.13	1.22	0.258	1.333	0.248
1Q97	5	1	1	5	25.00	2.08 *	1.55	0.038	5.333 *	0.021
2Q97	4	2	2	4	4.00	1.13	1.22	0.258	1.333	0.248
3Q97	3	3	3	3	1.00	0.00	1.15	1.000	0.000	1.000
4Q97	3	3	3	3	1.00	0.00	1.15	1.000	0.000	1.000
1Q98	3	3	3	3	1.00	0.00	1.15	1.000	0.000	1.000
TOTAL	54	36	36	54	2.25	2.66 **	0.30	0.008	7.200 **	0.007

* Statistically significant at the 5 percent level
** Statistically significant at the 1 percent level

The contingency table supports the contention that there is some persistence in fund performance, as indicated by the total positive Z-test repeat winners, Z-statistic of the Odds Ratio (greater than one) and Chi-square test. However, according to Malkiel's criteria, the persistence is not significant, as we cannot reject the null hypothesis of no performance persistence at the 5 percent significance level (although only marginally):

$$Z_{RW\ sample} \approx 1.90 < Z_{critical\ (0.025)} \approx 1.960$$

$$\text{p-value} \approx 0.058$$ [89]

Strong evidence of performance persistence is obtained when we apply the criteria based on the remaining tests, and even at a 1 percent significance level. In fact, in relation to the Odds Ratio Z-statistic:

$$Z_{OR\ sample} \approx 2.66 > Z_{critical\ (0.005)} \approx 2.576$$

$$\text{p-value} \approx 0.008$$

and relatively to the Chi-square test:

$$\chi^2_{sample} = 7.2 > \chi^2_{1,0.01} \approx 6.635$$

$$\text{p-value} \approx 0.007$$

[89] In general, for any hypothesis being tested, the probability value, or p-value, is the probability that the test statistic would be at least as large as the value actually observed if the null hypothesis is true. In this sense, the p-value reflects how much agreement there is between the data and the null hypothesis. If it is small, then given the value of the statistic actually observed, it is unlikely that the null hypothesis is true.

However, there is less evidence of persistence for most quarters considered separately. While there is slightly greater percentage of repeat winners than would be expected by chance, we cannot reject the null hypothesis of no persistence according to Malkiel's *z*-test. Relatively to the remaining two criteria, there are only four quarters (2Q95, 3Q95, 4Q95 and 1Q97) which indicate possible persistence of performance, at a 5 percent level (the p-values are 0.038 and 0.021 for the Odds Ratio and Chi-square tests, respectively). We can also observe a reversal pattern in three of the quarters (3Q94, 1Q95 and 2Q96), although these are not statistically significant.

Figure 4.4

Frequency of repeat winning and losing: excess returns

The bars represent the number of winning and losing funds each quarter that were winners or losers in the following quarter. WW, LL, LW and WL defined previously.

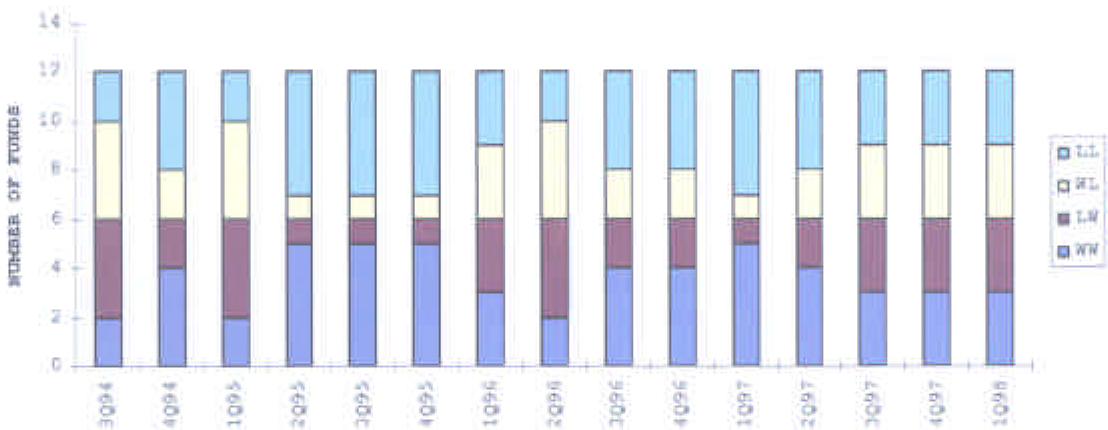


Figure 4.4 above shows the frequency of repeat losers and winners from quarter to quarter. Of course there are the same number of winners-winners as losers-losers, but the dual combination appears to be perhaps randomly distributed from quarter to quarter. One would not know beforehand in such a small sample whether the Percentage Repeat Winners or, alternatively, the Odds Ratio and Chi-Square tests would be a sufficient criteria.

4.4.2. Adjustment for small sample bias

Although indicating some evidence of performance persistence, these results require careful interpretation. The previous tests, valid asymptotically, do not adjust for possible small sample bias. In this context, we have followed the correction suggested by YATES [1934] to improve the approximation to the Chi-Square distribution and, alternatively, Fisher's nonparametric test, which does not use the chi-square distribution at all. Instead, the exact probability distribution of the observed frequencies is employed.[90]

Both the Yates' continuity correction and Fisher's exact p-values have been calculated for all of the quarters, as shown.

[90] These procedures are described in Chapter 3.

Table 4.9

Contingency table of excess returns:

adjustment for small sample bias

WW, WL, LW and LL defined as previously. Yates continuity correction and Fisher exact p-value as described in Chapter 3. P-values are based on two-tail tests.

	WW	LW	WL	LL	YATES CONT.CORR.	p-value	FISHER EXACT P-VALUE
3Q94	2	4	4	2	0.333	0.564	0.567
4Q94	4	2	2	4	0.333	0.564	0.567
1Q95	2	4	4	2	0.333	0.564	0.567
2Q95	5	1	1	5	3	0.083	0.080
3Q95	5	1	1	5	3	0.083	0.080
4Q95	5	1	1	5	3	0.083	0.080
1Q96	3	3	3	3	0.333	0.564	1.000
2Q96	2	4	4	2	0.333	0.564	0.567
3Q96	4	2	2	4	0.333	0.564	0.567
4Q96	4	2	2	4	0.333	0.564	0.567
1Q97	5	1	1	5	3	0.083	0.080
2Q97	4	2	2	4	0.333	0.564	0.567
3Q97	3	3	3	3	0.333	0.564	1.000
4Q97	3	3	3	3	0.333	0.564	1.000
1Q98	3	3	3	3	0.333	0.564	1.000
TOTAL	54	36	36	54	6.422 *	0.011	

* Statistically significant at the 5 percent level

Analysis of the results allow us to conclude that, at the 5 percent level, the small sample bias is indeed significant, since for 2Q95, 3Q95, 4Q95 and 1Q97 the p-value is increased from 0.021 to 0.083 (Yates) and 0.080 (Fisher). Also, these comparisons show that the Yates' continuity correction is a good approximation for the Fisher exact test. The Yates and Fisher p-values are roughly the same for the individual quarters, indicating there is no significant persistence of performance for these quarters.[91]

Since the Chi-square statistics for the total quarters have reasonable observed frequencies, there is no required Yates or Fisher adjustment, and the conclusion is consistent with the Odds Ratio Z-statistic noted above.

We also investigated whether the nonparametric test of market timing proposed by HENRIKSSON and MERTON (HM) [1981] could be generalized to a context of performance persistence. Similarly to PESARAN and TIMMERMANN [1992a], who have shown that this test is better interpreted as an exact test of independence within a 2 * 2 contingency table, we found that the HM test is identical to Fisher's nonparametric test.[92] PESARAN and TIMMERMANN [1992b] developed an alternative nonparametric test of predictive performance based on the proportion of times that the direction of change in the variable under consideration is correctly

---

[91] The previous table also uncovers a problem with the Yates statistic. In the cases where the cell frequencies are equal (1Q96, 3Q97, 4Q97 and 1Q98), the Yates continuity correction presents problems in terms of practical application. In fact, since this procedure involves subtracting and adding 0.5 from each cell frequency, what occurred was that the corrected frequencies turned out to be $\begin{matrix}3.5 & 3.5\\ 3.5 & 3.5\end{matrix}$, not only in this case but also in the quarters of the $\begin{matrix}4 & 2\\ 2 & 4\end{matrix}$ and $\begin{matrix}2 & 4\\ 4 & 2\end{matrix}$ type. In all cases, the corrected chi-square was 0.333, with a corresponding p-value of 0.564. For the $\begin{matrix}3 & 3\\ 3 & 3\end{matrix}$ quarters, this value contrasts with Fisher's exact p-value (=1). For this reason, from this point forward and whenever it is possible, we will use Fisher's exact p-value instead of the Yates continuity correction.

[92] This is demonstrated in Appendix 4.5.

predicted. This application measures the degree of association between realized and expected changes in some variable.

Let x_t be the predictor of y_t and suppose there are n observations on (y_t, x_t). Define:

$$Y_t = 1 \text{ if } y_t > 0$$
$$= 0 \text{ otherwise,}$$
$$X_t = 1 \text{ if } x_t > 0$$
$$= 0 \text{ otherwise,}$$

and

$$Z_t = 1 \text{ if } y_t x_t > 0$$
$$= 0 \text{ otherwise.}$$

Let $P_y = P\ (y_t > 0)$, $P_x = P\ (x_t > 0)$, and denote by $\hat{P}$ the proportion of times that the sign of y_t is predicted correctly; then,

$$\hat{P} = n^{-1} \sum_{t=1}^{n} Z_t = \overline{Z} \qquad [4.11]$$

For the 2 x 2 case, and on the null hypothesis that y_t and x_t are independently distributed, the test statistic can be based on the standardized statistic:

$$S_n = \frac{\hat{P} - \hat{P}_*}{\left\{ Var\left(\hat{P}\right) - Var\left(\hat{P}_*\right) \right\}^{1/2}} \sim N(0,1) \qquad [4.12]$$

where:

$$var\left(\hat{P}\right) = n^{-1}\,\hat{P}_*\left(1 - \hat{P}_*\right) \qquad [4.13]$$

and

$$var\left(\hat{P}_*\right) = n^{-1}\left(2\,\hat{P}_Y - 1\right)^2 \hat{P}_X\left(1 - \hat{P}_X\right) + n^{-1}\left(2\,\hat{P}_X - 1\right)^2 \hat{P}_Y\left(1 - \hat{P}_Y\right) + 4n^{-2}\,\hat{P}_Y\,\hat{P}_X\left(1 - \hat{P}_Y\right)\left(1 - \hat{P}_X\right) \qquad [4.14]$$

The last term for var $\left(\hat{P}_*\right)$ is asymptotically negligible.[93]

---

[93] In terms of contingency tables the variables can be defined as:

	y > 0	y > 0	
X > 0	n_{11}	n_{12}	n_{10}
X < 0	n_{21}	n_{22}	n_{20}
	n_{o1}	n_{o2}	n

$$\hat{p} = \sum_{i=1}^{2} p_{ii} = \hat{p}_{11} + \hat{p}_{12} = \frac{n_{11}}{n} + \frac{n_{12}}{n}$$

$$\boxed{= \frac{WW + LL}{n}} = \text{Proportion of times there is persistence;}$$

$$\hat{P}_* = \sum_{i=1}^{2} \hat{p}_{io}\,\hat{p}_{oi} = \hat{p}_{1o}\,\hat{p}_{o1} + \hat{p}_{2o}\,\hat{p}_{o2} = \frac{n_{1o}}{n} * \frac{n_{o1}}{n} + \frac{n_{2o}}{n} * \frac{n_{o2}}{n}$$

$$\boxed{= \frac{WW + LW}{n} * \frac{WW + LW}{n} + \frac{LW + LL}{n} * \frac{WL + LL}{n}}$$

$$\hat{P}_X = \frac{\sum_{t=1}^{2} X_t}{n} = \frac{n_{11} + n_{12}}{n} \boxed{= \frac{WW + WL}{n}}$$

$$\hat{P}_Y = \frac{\sum_{t=1}^{2} Y_t}{n} = \frac{n_{11} + n_{21}}{n} = \boxed{\frac{WW + LW}{n}}$$

$$1 - \hat{P}_Y \boxed{= \frac{WL + LL}{n}}$$

The results of applying the PESARAN and TIMMERMANN [1992b] test to the quarterly mean excess returns are summarized in Table 4.10, along with the results of the Chi-square tests. They reveal strong evidence of persistence of performance for 2Q95, 3Q95, 4Q95, 1Q97 as well as for the overall period. Interestingly, this test yields conclusions similar to those obtained through the Chi-square test. Also, we can verify that the square of S_n is directly comparable to the Chi-square results, thus confirming Pesaran and Timmermann's assertion.[94] After this analysis we can conclude that this methodology does not seem an appropriate criteria for assessing the performance persistence of our sample of funds, as it does not adjust for small sample bias, which as we have seen, affects the results.

[94] In fact, PESARAN and TIMMERMANN [1992b] show that S^2_n and the Chi-square test are asymptotically equivalent for the 2 X 2 case.

Table 4.10

Contingency table of excess returns: application of the PESARAN and TIMMERMANN [1992b] nonparametric test

WW, WL, LW and LL defined as previously. S_n^2 as described in the text. P-values are based on two-tail tests.

	WW	LW	WL	LL	S_n	S_n^2	p-value
3Q94	2	4	4	2	-1.206	1.455	0.228
4Q94	4	2	2	4	1.206	1.455	0.228
1Q95	2	4	4	2	-1.206	1.455	0.228
2Q95	5	1	1	5	2.412	5.818	0.016 *
3Q95	5	1	1	5	2.412	5.818	0.016 *
4Q95	5	1	1	5	2.412	5.818	0.016 *
1Q96	3	3	3	3	0	0	1
2Q96	2	4	4	2	-1.206	1.455	0.228
3Q96	4	2	2	4	1.206	1.455	0.228
4Q96	4	2	2	4	1.206	1.455	0.228
1Q97	5	1	1	5	2.412	5.818	0.016 *
2Q97	4	2	2	4	1.206	1.455	0.228
3Q97	3	3	3	3	0	0	1
4Q97	3	3	3	3	0	0	1
1Q98	3	3	3	3	0	0	1
TOTAL	54	36	36	54	2.691	7.240	0.007 **

** Statistically significant at the 1 percent level

In light of the above analysis, and because of the possibility of a bias due to small sample size, it is also desirable to complement the previous analysis and explore the robustness of the significance levels through bootstrap simulations. Bootstrap resampling is a relatively recent computer-intensive method of hypothesis testing introduced by EFRON [1979]. The basic idea underlying this technique is the similarity of the sample to the population: "*when the sample contains all of the information, why not proceed as if the sample is the population for purposes of estimating the sampling distribution of the test statistic?*" [NOREEN, 1989, p.65].

This nonparametric method constitutes an attractive alternative to conventional parametric methods for two reasons. Firstly, because it is less restrictive, as knowledge of the sampling distribution of the test statistic is not required. Secondly, bootstrap p-values are more accurate in small samples as they do not rely on the asymptotic distribution of the tests.

The implementation of the bootstrap involves the following steps:

1) From the original sample simulate daily return series through randomization (without replacement)[95] over time.

2) Calculate the Odds Ratio and the Chi-square test statistics for the resulting contingency table of winners and losers.

These procedures should be repeated a large number of times (*n* iterations) to generate simulated distributions that correspond to the null hypothesis of no performance persistence. These replicates contain information that will be used to make inferences from our data.

3) Finally, compute the ratio:[96]

$$\frac{nge + 1}{ns + 1} \qquad [4.15]$$

[95] This replicates the procedure used by BROWN and GOETZMANN [1995], and involves shuffling the original series.

[96] See NOREEN [1989, pp.17].

where:

nge is the number of times the pseudostatistic of the shuffled data is greater than or equal to the actual statistic for the original data;

ns is the number of iterations.

This ratio corresponds to the statistical significance level of the test and allows for rejection of the null hypothesis whenever it is less than the specified significance level for the test (α).

The results from bootstrap simulations of the Odds Ratio and Chi-square statistics after 1000 iterations are shown in Table 4.11.

Table 4.11

Comparison of the conventional and bootstrap p-values

This table reports the persistence statistics and p-values found previously, as well as the p-values obtained from bootstrap simulations.

	ODDS RATIO			CHI-SQUARE			FISHER EXACT
	Z-STAT	*p-value*	*Bootstrap p-value*	*CHI-SQ*	*p-value*	*Bootstrap p-value*	*p-value*
3Q94	-1.13	0.258	0.55	1.33	0.248	0.55	0.57
4Q94	1.13	0.258	0.59	1.33	0.248	0.59	0.57
1Q95	-1.13	0.258	0.54	1.33	0.248	0.55	0.57
2Q95	2.08	0.038	0.07	5.33	0.021	0.07	0.08
3Q95	2.08	0.038	0.08	5.33	0.021	0.08	0.08
4Q95	2.08	0.038	0.08	5.33	0.021	0.08	0.08
1Q96	0.00	1	1	0.00	1.000	1	1
2Q96	-1.13	0.258	0.57	1.33	0.248	0.57	0.57
3Q96	1.13	0.258	0.58	1.33	0.248	0.58	0.57
4Q96	1.13	0.258	0.59	1.33	0.248	0.59	0.57
1Q97	2.08	0.038	0.08	5.33	0.021	0.09	0.08
2Q97	1.13	0.258	0.58	1.33	0.248	0.58	0.57
3Q97	0.00	1	1	0.00	1	1	1
4Q97	0.00	1	1	0.00	1	1	1
1Q98	0.00	1	1	0.00	1	1	1
TOTAL	2.66	0.008	0.02	7.20	0.007	0.018	

We can observe that bootstrap p-values are more conservative than the asymptotic significance levels in the sense that the bootstrap p-values are, for every quarter and overall, greater than or equal to the p-values obtained from the conventional tests.[97] In particular for 2Q95, 3Q95, 4Q95 and 1Q97, the p-values have increased from 0.038/0.021 to approximately 0.07/0.08, no longer being statistically significant, at least at the 5 percent level. Only at the 10 percent level we might admit the hypothesis of performance persistence for these quarters.

We also verify from analysis of the table that the bootstrap p-values of the Z-statistic and Chi-square statistic are practically identical between themselves and in relation to Fisher's exact p-value obtained previously. This confirms our expectations as to the appropriateness of this methodology in a context of limited sample sizes. Hence, since the bootstrap methodologies are redundant, from now on we will use Fisher's exact p-value (or, if it is not possible, the Yates continuity correction) for adjusting the statistics to small sample bias.

[97] Contrary to BROWN AND GOETZMANN [1995], whose bootstrapped p-values generally agree with the theoretical distributions of the test statistic.

4.4.3. Performance persistence of risk-adjusted returns

To the extent that, at least in theory, excess returns in an efficient market are obtainable only be assuming excess risk, one would expect that risk-adjusted returns would exhibit lower performance for winners and lower performance persistence.[98] In order to examine whether consistency in performance reflects management skill or simply the differential returns between high risk and low risk funds, we employ two commonly used measures of risk-adjusted performance.

Table 4.12 shows the contingency table and test statistics results when funds are defined as winners/losers according to the SHARPE [1966] measure of performance.[99]

---

[98] However, VOLKMAN and WOHAR [1995] concluded that high-risk (as measured by stated goals) maximum capital gain funds demonstrated strong positive persistence in abnormal returns.

[99] The table containing the frequency of repeat performers according to SHARPE's [1966] ratio is presented in Appendix 4.6.

Table 4.12

Contingency table of Sharpe Returns

WW = Funds with Sharpe returns > median, repeated subsequent quarter; LL = Funds with Sharpe returns < median, repeated subsequent quarter; LW and WL correspond to funds with performance reversals. Sharpe measure = Average quarterly excess returns/Standard deviation. Statistical tests as described in Table 4.8. P-values are based on two-tail tests.

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	ODDS RATIO	B&G Z-STAT	σ (LOG)	p-value
3Q94	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13	1.22	0.258
4Q94	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
1Q95	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13	1.22	0.258
2Q95	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
3Q95	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
4Q95	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
1Q96	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
2Q96	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
3Q96	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
4Q96	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
1Q97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
2Q97	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
3Q97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
4Q97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
1Q98	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
TOTAL	52	38	38	52	0.58	1.48	0.140	1.87	2.08 *	0.30	0.038

	WW	LW	WL	LL	CHI-SQ	p-value	FISHER EXACT P-VALUE
3Q94	2	4	4	2	1.333	0.248	0.567
4Q94	4	2	2	4	1.333	0.248	0.567
1Q95	2	4	4	2	1.333	0.248	0.567
2Q95	3	3	3	3	0.000	1.000	1.000
3Q95	4	2	2	4	1.333	0.248	0.567
4Q95	4	2	2	4	1.333	0.248	0.567
1Q96	3	3	3	3	0.000	1.000	1.000
2Q96	4	2	2	4	1.333	0.248	0.567
3Q96	4	2	2	4	1.333	0.248	0.567
4Q96	3	3	3	3	0.000	1.000	1.000
1Q97	4	2	2	4	1.333	0.248	0.567
2Q97	3	3	3	3	0.000	1.000	1.000
3Q97	4	2	2	4	1.333	0.248	0.567
4Q97	4	2	2	4	1.333	0.248	0.567
1Q98	4	2	2	4	1.333	0.248	0.567
TOTAL	52	38	38	52	4.356 *	0.037	

* Statistically significant at the 5 percent level

As can be seen, the overall distribution of WW, LW, WL and LL successions is different from the previous table. The test statistics indicate percentage Repeat Winners greater than 50 percent, and Odds Ratio greater than one, and a high Chi-square statistic, but in all cases there is a reduced significance of persistence, as expected. In fact, both the Odds Ratio Z-statistic and the Chi-square statistic indicate possible persistence of performance, although it is no longer significant at the 1 percent level. For the Odds Ratio Z-test:

$$Z_{sample} \approx 2.08 > Z_{critical\ (0.025)} \approx 1.960$$

$$Z_{sample} \approx 2.08 > Z_{critical(0.025)} \approx 1.960 \quad (eq.)$$

$$\text{p-value} \approx 0.038$$

and for the Chi-square test:

$$\chi^2_{sample} \approx 4.356 > \chi^2_{1,0.05} \approx 3.841$$

$$\text{p-value} \approx 0.037$$

When we test the statistical significance of the individual quarterly results, these lead to a failure to reject the null hypothesis of no performance persistence for any quarter (the p-values are all higher than 0.05).[100]

Analysis of the four criteria allow us to conclude that risk adjustment by total risk reduces significantly the evidence of consistency observed previously in terms of excess returns. Hence, a superior performance in one quarter is as likely to be followed by an inferior performance in the next quarter as it is by superior performance. This brings into question why these fund managers should be rewarded according to their raw return performance over any quarter.

[100] The p-values from Fisher's exact test are in this case redundant, and will further reinforce this conclusion.

We also investigated if persistence is present in terms of alpha risk-adjusted returns (JENSEN's [1968] measure of performance). The results are displayed in table 4.13.[101]

Table 4.13

Contingency table of alpha returns

WW = Funds with alphas > median, repeated subsequent quarter; LL = Funds with alphas < median, repeated subsequent quarter; LW and WL correspond to funds with performance reversals. Jensen's alpha is the alpha from CAPM's empirical analogue: $R_{p,t}-R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon_{p,t}$. Statistical tests as described in Table 4.8. P-values are based on two-tail tests.

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	ODDS RATIO	B&G Z-STAT	σ (LOG)	p-value
3Q94	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13	1.22	0.258
4Q94	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
1Q95	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13	1.22	0.258
2Q95	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
3Q95	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
4Q95	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
1Q96	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13	1.22	0.258
2Q96	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
3Q96	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
4Q96	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
1Q97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
2Q97	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
3Q97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
4Q97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
1Q98	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
TOTAL	52	38	38	52	0.58	1.48	0.140	1.87	2.08 *	0.30	0.038

	WW	LW	WL	LL	CHI-SQ	p-value	FISHER EXACT P-VALUE
3Q94	2	4	4	2	1.333	0.248	0.567
4Q94	4	2	2	4	1.333	0.248	0.567
1Q95	2	4	4	2	1.333	0.248	0.567
2Q95	4	2	2	4	1.333	0.248	0.567
3Q95	5	1	1	5	5.333 *	0.021	0.080
4Q95	3	3	3	3	0.000	1.000	1.000
1Q96	2	4	4	2	1.333	0.248	0.567
2Q96	4	2	2	4	1.333	0.248	0.567
3Q96	5	1	1	5	5.333 *	0.021	0.080
4Q96	3	3	3	3	0.000	1.000	1.000
1Q97	4	2	2	4	1.333	0.248	0.567
2Q97	3	3	3	3	0.000	1.000	1.000
3Q97	4	2	2	4	1.333	0.248	0.567
4Q97	4	2	2	4	1.333	0.248	0.567
1Q98	3	3	3	3	0.000	1.000	1.000
TOTAL	52	38	38	52	4.356 *	0.037	

* Statistically significant at the 5 percent level

[101] Quarterly estimates of the regressions are presented in Appendix 4.7. The table containing the frequency of repeat performances according to Jensen's alpha is presented in Appendix 4.8.

The above contingency table shows that risk-adjustment by systematic risk also affects the pattern of persistence. Interestingly, the overall distribution of the frequency of winner and loser repeats is identical to that observed using Sharpe's performance measure.[102] Hence, once again there is little evidence of performance persistence for the combined results: only for the Odds Ratio Z-statistic and Chi-square statistic can we reject the null hypothesis of no performance persistence, at the 5 percent level.

There are yet slight differences in the pattern observed for the individual quarters.[103] The Odds Ratio Z-statistic and Chi-square statistic indicate evidence of persistence for two quarters (3Q95 and 3Q96) at the 5 percent level (the p-values are 0.038 and 0.021, respectively). However, the significance is eliminated when we adjust for small sample bias (Fisher's exact p-value = 0.083).

In general, our results are not consistent with the conclusions of most of the studies. For example, GOETZMANN and IBBOTSON [1994] and BROWN and GOETZMANN [1995][104] find that adjustment for risk further supports the repeat winner phenomenon. An additional issue to emphasize concerns the similarity of the contingency table tests resulting from

---

[102] And also, the test statistics for the overall period are identical.

[103] With the exception of Malkiel's Z-test, which has not yet indicated any evidence (at least at the 5 percent level) of performance persistence for any quarter.

[104] MALKIEL [1995] only assesses the persistence of performance in terms of total returns.

the use of systematic and total risk. While one might argue that investors are unconcerned with total risk because it can be diversified, the results suggest that these portfolios might not be sufficiently diversified.

In light of the above, we will next examine the pattern of risk characteristics of our sample of funds.

4.4.4. Performance Persistence of Risk Characteristics

Risk measurement is a debatable issue among performance evaluators. We use two measures of risk (standard deviation and beta) to address the question of the persistence of risk among fund managers for the four year sample period.

Table 4.14 shows that there is a greater persistence in risk (as measure by the standard deviation) than in returns.[105]

[105] The table containing the frequency of repeats for risk "winners and losers" is presented in Appendix 4.9.

Table 4.14

Contingency table of risk persistence: standard deviation

WW = Funds with risk > median, repeated subsequent quarter; LL = Funds with risk < median, repeated subsequent quarter; LW and WL correspond to funds with performance reversals. Risk measure = average quarterly standard deviation. Statistical tests as described in Table 4.8. P-values are based on two-tail tests.

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	ODDS RATIO	B&G Z-STAT	σ (LOG)	p-value
3Q94	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
4Q94	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
1Q95	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
2Q95	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
3Q95	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
4Q95	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
1Q96	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
2Q96	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
3Q96	5.5	0.5	0.5	5.5	0.92	2.04 *	0.041	121.00	2.30 *	2.09	0.022
4Q96	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
1Q97	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
2Q97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
3Q97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
4Q97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
1Q98	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
TOTAL	62	28	28	62	0.69	3.58 **	0.000	4.90	4.94 **	0.32	0.000

	WW	LW	WL	LL	CHI-SQ	p-value	FISHER EXACT P-VALUE
3Q94	4	2	2	4	1.333	0.248	0.567
4Q94	4	2	2	4	1.333	0.248	0.567
1Q95	4	2	2	4	1.333	0.248	0.567
2Q95	4	2	2	4	1.333	0.248	0.567
3Q95	5	1	1	5	5.333 *	0.021	0.080
4Q95	3	3	3	3	0.000	1.000	1.000
1Q96	3	3	3	3	0.000	1.000	1.000
2Q96	5	1	1	5	5.333 *	0.021	0.080
3Q96	6	0	0	6	12.000 **	0.001	0.002 **
4Q96	4	2	2	4	1.333	0.248	0.567
1Q97	5	1	1	5	5.333 *	0.021	0.080
2Q97	4	2	2	4	1.333	0.248	0.567
3Q97	4	2	2	4	1.333	0.248	0.567
4Q97	4	2	2	4	1.333	0.248	0.567
1Q98	3	3	3	3	0.000	1.000	1.000
TOTAL	62	28	28	62	25.689 **	0.000	

* Statistically significant at the 5 percent level
** Statistically significant at the 1 percent level

In this case, a risk "winner" has a standard deviation in excess of the median, and a risk "loser" has standard deviation less than the median.

It immediately stands out from the table that the overall statistics are approximately over twice than those observed for excess returns, allowing for a clear rejection of the null hypothesis of no "risk persistence", at a significance level of 1 percent. In fact, the resulting statistics (Z-test Repeat Winners, Odds Ratio Z-statistic and Chi-square statistic) are, respectively:

$$Z_{RW\ sample} \approx 3.58 > Z_{critical\ (0.005)} \approx 2.576$$

p-value ≈ 0 [106]

$$Z_{OR\ sample} \approx 4.90 > Z_{critical\ (0.005)} \approx 2.576$$

p-value ≈ 0

and

$$\chi^2_{sample} \approx 25.689 > \chi^2_{1,0.01} \approx 6.635$$

There is also evidence of risk persistence for four quarters (3Q95, 2Q96, 3Q96, and 1Q97),[107] as indicated by the corresponding p-values (<0.05). While these should be adjusted for small sample bias, the 3Q96 p-value is not significantly modified (from 0.001 to 0.002), remaining significant at the 1 percent level.[108]

---

[106] Curiously, until now it is the first time that Malkiel's Z-test Repeat Winners criteria allows for rejection of the null hypothesis.

[107] In fitting log-linear models to contingency table data for small samples, EVERITT [1992] suggests increasing cell frequencies by the addition of a small constant (for example 0.5) to remove "sampling zeros". Zeros occur in the quarterly contingency table only in 3Q96, which has been amended by adding 0.5 to LW and WL, and subtracting 0.5 from WW and LL for consistency [EVERITT, 1992, p.136].

[108] Interestingly, this is the only quarter for which the Repeat Winner Z-test indicates evidence of persistence.

The persistence of risk is further intensified when measured in terms of systematic risk (beta), as shown in Table 4.15 below.[109]

Table 4.15

Contingency table of risk persistence: Beta

WW = Funds with risk > median, repeated subsequent quarter; LL = Funds with risk < median, repeated subsequent quarter; LW and WL correspond to funds with performance reversals. Risk measure = average quarterly standard deviation. Statistical tests as described in Table 4.8. P-values are based on two-tail tests.

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	ODDS RATIO	B&G Z-STAT	σ (LOG)	p-value
3Q94	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
4Q94	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
1Q95	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
2Q95	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
3Q95	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
4Q95	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
1Q96	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
2Q96	5.5	0.5	0.5	5.5	0.92	2.04 *	0.041	121.00	2.30 *	2.09	0.022
3Q96	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
4Q96	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
1Q97	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
2Q97	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
3Q97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
4Q97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
1Q98	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
TOTAL	66	24	24	66	0.73	4.43 **	0.000	7.56	6.00 **	0.34	0.000

	WW	LW	WL	LL	CHI-SQ	p-value	FISHER EXACT P-VALUE
3Q94	5	1	1	5	5.333 *	0.021	0.080
4Q94	5	1	1	5	5.333 *	0.021	0.080
1Q95	4	2	2	4	1.333	0.248	0.567
2Q95	4	2	2	4	1.333	0.248	0.567
3Q95	4	2	2	4	1.333	0.248	0.567
4Q95	4	2	2	4	1.333	0.248	0.567
1Q96	5	1	1	5	5.333 *	0.021	0.080
2Q96	6	0	0	5	12.000 **	0.001	0.002 **
3Q96	5	1	1	5	5.333 *	0.021	0.080
4Q96	3	3	3	3	0.000	1.000	1.000
1Q97	5	1	1	5	5.333 *	0.021	0.080
2Q97	5	1	1	5	5.333 *	0.021	0.080
3Q97	4	2	2	4	1.333	0.248	0.567
4Q97	4	2	2	4	1.333	0.248	0.567
1Q98	3	3	3	3	0.000	1.000	1.000
TOTAL	66	24	24	66	39.200 **	0.000	

* Statistically significant at the 5 percent level
** Statistically significant at the 1 percent level

[109] The table of frequency of repeat winning and losing funds (in terms of systematic risk) is presented in Appendix 4.10.

As can be observed, the overall results indicate strong evidence of risk persistence, allowing for rejection of the null hypothesis of no risk persistence at the 1 percent level, as attested by:

$$Z_{RW\ sample} \approx 4.43 > Z_{critical\ (0.005)} \approx 2.576$$

$$\text{p-value} = 0$$

$$Z_{OR\ sample} \approx 6 > Z_{critical\ (0.005)} \approx 2.576$$

$$\text{p-value} = 0$$

and

$$\chi^2_{sample} \approx 39.2 > \chi^2_{1,0.01} \approx 6.635$$

for the Repeat Winner Z-test, Odds Ratio Z-test, and Chi-square test, respectively.

There is also strong evidence of risk persistence for approximately half of the quarters (p-values < 0.05). Once again, correction for small sample size bias reduces the significance of persistence for all quarters except 2Q96, which continues to be statistically significant at the 1 percent level.[110]

We can conclude from both tables 4.14 and 4.15 that there seems to be a greater persistence of risk-characteristics than of returns (risk-adjusted or not). This implies that the risky funds are consistently risk takers over time and vice-versa, i. e., the less risky funds are consistently risk-averters. Perhaps the

[110] As before, this conclusion is consistent with the corresponding Repeat Winner Z-statistic for this quarter.

relevant question at this stage is whether the excess return funds are also the excess risk funds.

Figure 4.5

Fund risk (standard deviation) - return

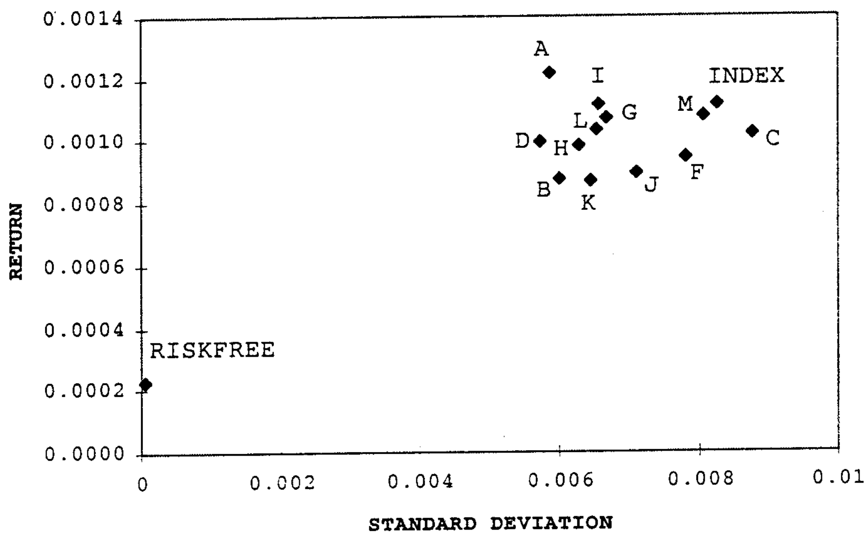


Figure 4.6

Fund risk (Beta) - return

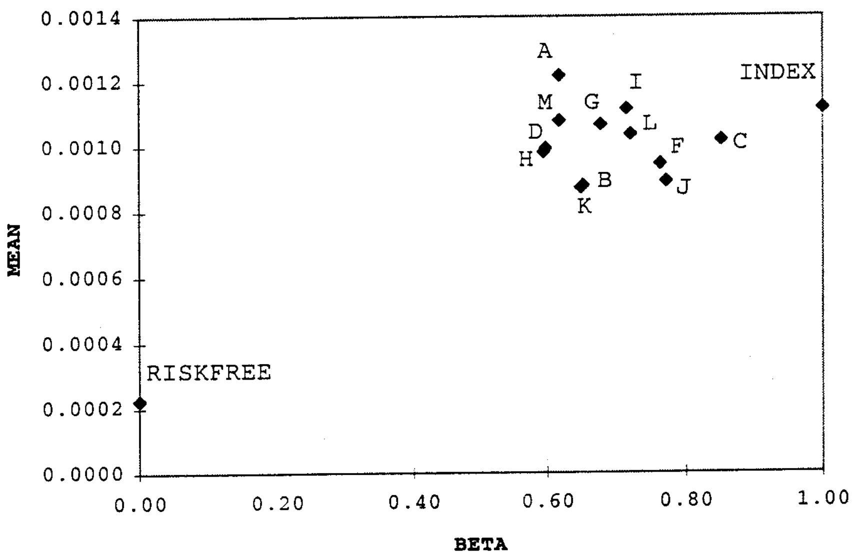


Figure 4.5 shows that some fund managers delivering superior returns are also undertaking substantial risk, but others (such as A) are Sharpe efficient (that is, high excess returns per unit risk). Others, such as C and F are Sharpe inefficient (that is, low returns per unit risk). The same type of conclusions can be drawn from analysis of Figure 4.6, since the distribution of the funds in the scatterplot area is broadly similar. It is particularly interesting to observe fund A which, as we have seen in the previous sections, is the most persistent winning fund (measured either in terms of risk-adjusted or unadjusted returns). In terms of risk persistence, it is now the persistent losing fund, a fact which is reflected in Figures 4.5 and 4.6.

4.4.5. Term structure of Performance Persistence

So far, our conclusions on performance persistence have been based on the results obtained from quarterly contingency tables of returns. The question we will now raise is: to what extent do the performance persistent results depend on the interval of return measurement, for the same overall sample period? This is, indirectly, a question regarding the term structure of performance persistence. Is the pattern of overall performance persistence within the sample and for individual funds the same for shorter and longer periods of measurement? This issue is addressed by considering performance persistence measurements for two years, one year, half year and monthly, compared to the previous quarterly intervals.

Perhaps surprisingly almost all of the other individual period performance persistence measures show that there is very little persistence. The contingency table of excess returns for two years (Table 4.16), for one year (Table 4.17), for the individual half years (Table 4.18)[111] and for every individual month (Table 4.19)[112] are presented next.

Table 4.16

Contingency table of excess returns: two year periods

WW, LL, WL and LW defined as in Table 4.8. Excess returns are measured as the two-year average of daily return less daily riskfree rate. Statistical tests as described in Table 4.8. P-values are based on two-tail tests.

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	ODDS RATIO	B&G Z-STAT	σ (LOG)	p-value
Apr.96-Mar.98	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000

	WW	LW	WL	LL	CHI-SQ	p-value	FISHER EXACT P-VALUE
Apr.96-Mar.98	3	3	3	3	0.000	1.000	1.000

As we can observe from Table 4.16 above, when performance is measured in terms of two year periods, there is no evidence of performance persistence. The null hypothesis of no performance persistence cannot be rejected at the 5 percent level according to any of the criteria used:

111 The table containing the frequency of the repeat performances relatively to Tables 4.16 and 4.17 is presented in Appendix 4.11. Appendix 4.12 presents the frequencies corresponding to Table 4.18.

112 Table 4.19 is a summary of the complete contingency table, which is available in Appendix 4.13.

$$Z_{RW\ sample} \approx 0.5 < Z_{critical\ (0.025)} \approx 1.960$$

p-value = 1

$$Z_{OR\ sample} \approx 0 < Z_{critical\ (0.025)} \approx 1.960$$

p-value = 1

$$\chi^2_{sample} \approx 0 < \chi^2_{1,0.05} \approx 3.841$$

p-value = 1

and also Fisher's exact p-value = 1.

Table 4.17 below displays the contingency table of excess returns for annual measurements.

Table 4.17

Contingency table of excess returns: annual periods

WW, LL, WL and LW defined as in Table 4.8. Excess returns are measured as the annual average of daily return less daily riskfree rate. Statistical tests as described in Table 4.8. P-values are based on two-tail tests.

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	ODDS RATIO	B&G Z-STAT	σ (LOG)	p-value
2nd Year	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
3rd Year	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13	1.22	0.258
4th Year	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13	1.22	0.258
TOTAL	8	10	10	8	0.44	-0.47	0.637	0.64	-0.67	0.67	0.506

	WW	LW	WL	LL	CHI-SQ	p-value	FISHER EXACT P-VALUE
2nd Year	4	2	2	4	1.333	0.414	0.567
3rd Year	2	4	4	2	1.333	0.414	0.567
4th Year	2	4	4	2	1.333	0.414	0.567
TOTAL	8	10	10	8	0.444	0.505	

Again, these results are substantially different from those obtained when performance is measured relatively to quarterly periods. There is no evidence whatsoever of any performance persistence for any year or for the combined results. This conclusion is consistent with all of the tests performed: neither the Percentage Repeat Winner Z-test, Odds Ratio Z-statistic nor the Chi-square criteria allow for rejection of the null hypothesis, as can be seen by:

$$Z_{RW\ sample} \approx -0.47 < Z_{critical\ (0.025)} \approx 1.960$$

$$\text{p-value} \approx 0.637$$

$$Z_{OR\ sample} \approx -0.67 < Z_{critical\ (0.025)} \approx 1.960$$

$$\text{p-value} \approx 1$$

$$\chi^2_{sample} \approx 0.444 < \chi^2_{1,0.05} \approx 3.841$$

$$\text{p-value} \approx 0.505$$

Furthermore, if any evidence exists, it is more consistent with the existence of reversal patterns (as indicated by the negative values of Malkiel's Z-test and Brown and Goetzmann Z-statistic, although not significant).

Table 4.18

Contingency table of excess returns: half year periods

WW, LL, WL and LW defined as in Table 4.8. Excess returns are measured as the half-year (October-March and April-September) average of daily return less daily riskfree rate. Statistical tests as described in Table 4.8. P-values are based on two-tail tests.

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	ODDS RATIO	B&G Z-STAT	σ (LOG)	p-value
O94-M95	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
A95-S95	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
O95-M96	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
A96-S96	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
O96-M97	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
A97-S97	5	1	1	5	0.83	1.63	0.102	25.00	2.08 *	1.55	0.038
O97-M98	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
TOTAL	26	16	16	26	0.62	1.54	0.123	2.64	2.16 *	0.45	0.031

	WW	LW	WL	LL	CHI-SQ	p-value	FISHER EXACT P-VALUE
O94-M95	3	3	3	3	0.000	1.000	1.000
A95-S95	4	2	2	4	1.333	0.248	0.567
O95-M96	5	1	1	5	5.333 *	0.021	0.080
A96-S96	3	3	3	3	0.000	1.000	1.000
O96-M97	3	3	3	3	0.000	1.000	1.000
A97-S97	5	1	1	5	5.333 *	0.021	0.080
O97-M98	3	3	3	3	0.000	1.000	1.000
TOTAL	26	16	16	26	4.762 *	0.029	

* Statistically significant at the 5 percent level

Although not completely eliminated, the evidence of persistence is slightly reduced (relatively to quarterly measurements) when we construct a half year contingency table of excess returns. This is suggested by the statistics for the overall period:

$$Z_{RW\ sample} \approx 0.94 < Z_{critical\ (0.025)} \approx 1.960$$

$$\text{p-value} \approx 0.123$$

$$Z_{OR\ sample} \approx 2.16 > Z_{critical\ (0.025)} \approx 1.960$$

$$\text{p-value} \approx 0.031$$

$$\chi^2_{sample} \approx 4.762 > \chi^2_{1,0.05} \approx 3.841$$

$$\text{p-value} \approx 0.029$$

Also, two of the half-years (Oct.95-Mar.96 and Apr.97-Sept.97) present evidence of persistence according to the Odds Ratio and Chi-square statistics (p-values < 0.05). Not surprisingly though, this evidence disappears when adjustments for small sample bias are executed (Fisher's exact p-value ≈ 0.083).

Table 4.19

Summary contingency table of excess returns: monthly periods

WW, LL, WL and LW defined as in Table 4.8. Excess returns are measured as the monthly average of daily return less daily riskfree rate. Statistical tests as described in Table 4.8. P-values are based on two-tail tests.

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	ODDS RATIO	B&G Z-STAT	σ (LOG)	p-value
TOTAL	147	135	135	147	0.52	0.71	0.475	1.19	1.01	0.17	0.312

	WW	LW	WL	LL	CHI-SQ	p-value	YATES CONT CORR.	p-value
TOTAL	147	135	135	147	1.021	0.312	0.858	0.3543

The results of the monthly contingency table of excess returns confirm that there is practically no evidence of performance persistence. For the overall period, and according to either criteria, we cannot reject the null hypothesis of independence of performance in subsequent periods:

$$Z_{RW\ sample} \approx 0.71 < Z_{critical\ (0.025)} \approx 1.960$$

$$\text{p-value} \approx 0.475$$

$$Z_{OR\ sample} \approx 1.01 < Z_{critical\ (0.025)} \approx 1.960$$

$$\text{p-value} \approx 0.312$$

$$\chi^2_{sample} \approx 1.021 < \chi^2_{1,0.05} \approx 3.841$$

$$\text{p-value} \approx 0.312$$

The monthly intervals show almost exactly random distributions of winners and losers. In fact, for the individual months, we can only reject the null hypothesis for five months (out of 47), which exhibit p-values lower than 0.05. Even so, this conclusion is not supported by Fisher's exact p-value (≈ 0.080).

In short, when we consider alternative periods for performance measurement, especially monthly and annual returns, there is little (if any) evidence of performance persistence (not noted by BROWN and GOETZMANN [1995] or MALKIEL [1995]), as summarized in Table 4.20.

Table 4.20

Summary of persistence statistics for different measurement periods (Excess Returns)

Interval	Repeat Winner	Z-test	Odds Ratio	Z-stat	Chi-square	Yates / Fisher
Month	0.52	0.71	1.19	1.01	1.02	0.858
		(0.475)		(0.312)	(0.312)	(0.354)
Quarter	0.60	1.90	2.25	2.66	7.20	6.42
		(0.058)		(0.008)	(0.007)	(0.011)
Half-Year	0.62	1.54	2.64	2.16	4.76	3.86
		(0.123)		(0.031)	(0.029)	(0.050)
Annual	0.44	-0.47	0.64	-0.67	0.44	0.11
		(0.637)		(0.506)	(0.505)	(0.739)
Bi-annual	0.50	0.00	0.64	0.00	0.00	0.33
		(1.000)		(1.000)	(1.000)	(1.000)

p-values in parenthesis

This effect is reinforced when performance is measured in terms of risk-adjusted returns, as we will analyze in the following tables (Table 4.21, 4.22 and 4.23).[113]

Table 4.21

Contingency table of fund alphas: two year periods

WW, LL, WL and LW defined as in Table 4.13. Statistical tests as described in Table 4.8. P-values are based on two-tail tests.

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	ODDS RATIO	B&G Z-STAT	σ (LOG)	p-value
Apr.96-Mar.98	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000

	WW	LW	WL	LL	CHI-SQ	p-value	FISHER EXACT P-VALUE
Apr.96-Mar.98	3	3	3	3	0.000	1.000	1.000

[113] The complete tables of frequency of repeat performances corresponding to Tables 4.21, 4.22 and 4.23 are presented in Appendices 4.14, 4.15 and 4.17. The half-year regression estimates of alpha are presented in Appendix 4.16.

When returns are risk-adjusted and measured for two year periods, the results are consistent with the null hypothesis of no performance persistence. Also, the distribution of WW, WL, LW, and LL successions matches that already observed for the case of excess returns (Table 4.17).

Table 4.22

Contingency table of alpha returns: one year periods

WW, LL, WL and LW defined as in Table 4.13. Statistical tests as described in Table 4.8. P-values are based on two-tail tests.

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	ODDS RATIO	B&G Z-STAT	σ (LOG)	p-value
2nd Year	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
3rd Year	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13	1.22	0.258
4th Year	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13	1.22	0.258
TOTAL	8	10	10	8	0.44	-0.47	0.637	0.64	-0.67	0.67	0.506

	WW	LW	WL	LL	CHI-SQ	p-value	FISHER EXACT P-VALUE
2nd Year	4	2	2	4	1.333	0.414	0.567
3rd Year	2	4	4	2	1.333	0.414	0.567
4th Year	2	4	4	2	1.333	0.414	0.567
TOTAL	8	10	10	8	0.444	0.505	

Curiously (and similarly to the two-year period results), the distribution of WW, WL, LW and LL is identical to the pattern resulting from excess returns (Table 4.17), and analyzed previously. So, once again, we find no evidence of performance persistence.[114]

[114] If any, it is more consistent with reversals of performance, although not at a statistically significant level.

Table 4.23

Contingency table of alpha returns: half year periods

WW, LL, WL and LW defined as in Table 4.8. Statistical tests as described in Table 4.8. P-values are based on two-tail tests.

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	ODDS RATIO	B&G Z-STAT	σ (LOG)	p-value
O94-M95	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13	1.22	0.258
A95-S95	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
O95-M96	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
A96-S96	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
O96-M97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
A97-S97	4	2	2	4	0.67	0.82	0.414	4.00	1.13	1.22	0.258
O97-M98	3	3	3	3	0.50	0.00	1.000	1.00	0.00	1.15	1.000
TOTAL	24	18	18	24	0.57	0.93	0.355	1.78	1.30	0.44	0.192

	WW	LW	WL	LL	CHI-SQ	p-value	FISHER EXACT P-VALUE
O94-M95	2	4	4	2	1.333	0.248	0.567
A95-S95	4	2	2	4	1.333	0.248	0.567
O95-M96	4	2	2	4	1.333	0.248	0.567
A96-S96	3	3	3	3	0.000	1.000	1.000
O96-M97	4	2	2	4	1.333	0.248	0.567
A97-S97	4	2	2	4	1.333	0.248	0.567
O97-M98	3	3	3	3	0.000	1.000	1.000
TOTAL	24	18	18	24	1.714	0.190	

Not surprisingly, any evidence of persistence in performance practically disappears when we consider half-year periods. This is due not only to the fact that we are considering risk-adjusted returns,[115] but also to the effect of choosing a different measurement period.

Table 4.24 below is a summary of the persistence statistics for different measurement periods, when performance is evaluated according to alpha returns.

[115] Which, as we have already seen in section 4.4.3, has reduced the intensity of the persistence phenomenon.

Table 4.24

Summary of persistence statistics for different measurement periods (Alpha returns)

Interval	Repeat Winner	Z-test	Odds Ratio	Z-stat	Chi-square	Yates / Fisher
Quarter	0.58	1.48	1.87	2.08	4.356	3.755
		(0.140)		(0.038)	(0.037)	(0.053)
Half-Year	0.57	0.93	1.78	1.30	1.71	1.19
		(0.355)		(0.192)	(0.190)	(0.275)
Annual	0.44	-0.47	0.64	-0.67	0.44	1.11
		(0.637)		(0.506)	(0.505)	(0.739)
Bi-annual	0.50	0.00	1.00	0.00	0.00	
		(1.000)		(1.000)	(1.000)	(1.000)

p-values in parenthesis

Analysis of the tables presented in this section seem to confirm that the issue of the selection of intervals may influence the performance persistence results. For our sample of Portuguese equity funds there appears to be some evidence of performance persistence, as measured by quarterly periods, that is not observed over shorter (monthly) or longer (year or two-year) periods.

4.4.6. Individual fund performance persistence

So far, we have seen that, except in the context of quarterly contingency tables and returns unadjusted for risk, there is little evidence of performance persistence for the

Portuguese equity fund industry. However, individually, some funds seem to exhibit characteristics of superior persistent (Fund A) or inferior persistent (Fund B) performance over both short and long intervals and for both adjusted and unadjusted returns.

We will next present and analyze contingency tables of performance persistence for individual funds. We only report tables for quarterly and monthly periods of measurement, because it is improbable that the persistence for any fund will be statistically significant over the half year repeats, because of the very limited frequencies.

The persistence assessment criteria based on the Odds Ratio Z-statistic and the significance levels corrected for small sample bias are not computed in this section, since they are inappropriate for testing the performance persistence of individual funds.[116]

[116] In the cases of superior/inferior persistence performing funds, one of the quadrants (Northwest/Southeast) of the contingency table will be excessively weighted (relatively to the inverse quadrant). The Fisher exact p-value can only be doubled to give the equivalent of a two-tailed test in the case where the sample sizes in each group variable are the same (EVERITT [1992]). Besides the Fisher test, the application of the Odds Ratio Z-statistic and the Yates p-values in this type of context lead to incongruous results. This can be illustrated for a hypothetical extreme example of a persistent winning fund:

	W	L
W	15	0
L	0	0

The Repeat Winners z-test (≈ 3.87; p-value ≈ 0.0001) and the Chi-square test (≈ 45; p-value ≈ 0.000) indicate very strong evidence of persistence. Yet the Odds Ratio and Yates statistic are not computable, whereas Fisher exact p-value = 1.

Anyhow, not correcting for small sample bias in this particular context will probably not have a major effect on the conclusions, since we have a total of 15 and 47 frequencies for each fund, for quarterly and monthly intervals, respectively.

Table 4.25

Performance persistence for individual funds
(Quarterly excess returns)

Panel A. Consistency of individual fund performance

		A	B	C	D	F	G	H	I	J	K	L	M
1994-1998	WW	9	0	2	4	3	7	5	8	5	3	3	5
	LW	3	3	5	3	4	3	2	2	2	2	6	1
	WL	3	3	4	3	3	3	3	3	2	3	5	1
	LL	0	9	4	5	5	2	5	2	6	7	1	8

Panel B: Performance Persistence statistics

Fund	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	CHI-SQ	p-value
A	9	3	3	0	0.75	1.73	0.083	11.400 **	0.001
B	0	3	3	9	0.00	-1.73	0.083	11.400 **	0.001
C	2	5	4	4	0.33	-0.82	0.414	1.267	0.260
D	4	3	3	5	0.57	0.38	0.705	0.733	0.392
F	3	4	3	5	0.50	0.00	1.000	0.733	0.392
G	7	3	3	2	0.70	1.26	0.206	3.933 *	0.047
H	5	2	3	5	0.63	0.71	0.479	1.800	0.180
I	8	2	3	2	0.73	1.51	0.132	6.600 *	0.010
J	5	2	2	6	0.71	1.13	0.257	3.400	0.065
K	3	2	3	7	0.50	0.00	1.000	3.933 *	0.047
L	3	6	5	1	0.38	-0.71	0.479	3.933 *	0.047
M	5	1	1	8	0.83	1.63	0.102	9.267 **	0.002
TOTAL	54	36	36	54					

* Statistically significant at the 5 percent level
** Statistically significant at the 1 percent level

We can verify from the analysis of Table 4.25 that funds A, G and I are the most persistent winning funds. This evidence is statistically significant at least at the 5 percent level, according to the Odds Ratio Z-statistic and the corresponding p-values (less than 0.05). Based on the same test, the quarterly loser persistence of funds B, K and M is also significant, although only marginally for K.[117] No fund exhibits any type of statistical significance when the Repeat Winner Z-test is

[117] The p-value is 0.047, very close to the 0.05 limit. The same is appliable to fund G, although on a winning perspective.

considered, but this is not surprising, since this test has consistently given p-values which are more conservative than those resulting from the remaining tests.

Table 4.26

Performance persistence for individual funds

(Monthly excess returns)

Panel A. Consistency of individual fund performance

		A	B	C	D	F	G	H	I	J	K	L	M
1994-1998	WW	21	6	16	12	9	11	13	10	9	14	13	13
	LW	10	10	10	13	13	14	8	13	13	8	13	10
	WL	11	10	9	12	12	14	9	13	14	9	12	10
	LL	5	21	12	10	13	8	17	11	11	16	9	14

Panel B: Performance Persistence statistics

Fund	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	CHI-SQ	p-value
A	21	10	11	5	0.66	1.77	0.077	11.468 **	0.001
B	6	10	10	21	0.38	-1.00	0.317	10.617 **	0.001
C	16	10	9	12	0.64	1.40	0.162	2.447	0.118
D	12	13	12	10	0.50	0.00	1.000	0.404	0.525
F	9	13	12	13	0.43	-0.65	0.513	0.915	0.339
G	11	14	14	8	0.44	-0.60	0.549	2.106	0.147
H	13	8	9	17	0.59	0.85	0.394	4.319	0.038
I	10	13	13	11	0.43	-0.63	0.532	0.574	0.448
J	9	13	14	11	0.39	-1.04	0.297	1.255	0.263
K	14	8	9	16	0.61	1.04	0.297	3.809	0.051
L	13	13	12	9	0.52	0.20	0.841	0.915	0.339
M	13	10	10	14	0.57	0.63	0.532	1.085	0.298
TOTAL	147	135	135	147					

* Statistically significant at the 5 percent level
** Statistically significant at the 1 percent level

In Table 4.26, shorter interval measurements show similar results for the very top performing fund A and the very poor performing fund B. However, the results are different for funds G, K and M. The performance persistence on a monthly basis for these funds is no longer statistically significant under any of the measures. Fund H now shows significant loser persistence in terms of the Chi-square statistic test (although only marginally).

Table 4.27

Performance persistence for individual funds (Quarterly alphas)

Panel A. Consistency of individual fund performance

		A	B	C	D	F	G	H	I	J	K	L	M
1994-1998	WW	13	0	4	3	0	2	7	8	2	4	5	4
	LW	1	4	3	5	6	4	2	2	1	3	3	4
	WL	0	5	3	4	6	4	2	2	2	3	4	3
	LL	1	6	5	3	3	5	4	3	10	5	3	4

Panel B: Performance Persistence statistics

Fund	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	CHI-SQ	p-value
A	13	1	0	1	1.00	3.61 **	0.000	30.600 **	0.000
B	0	4	5	6	0.00	-2.24 *	0.025	5.533 *	0.019
C	4	3	3	5	0.57	0.38	0.705	0.733	0.392
D	3	5	4	3	0.43	-0.38	0.705	0.733	0.392
F	0	6	6	3	0.00	-2.45	0.014	6.600 **	0.010
G	2	4	4	5	0.33	-0.82	0.414	1.267	0.260
H	7	2	2	4	0.78	1.67	0.096	4.467 *	0.035
I	8	2	2	3	0.80	1.90	0.058	6.600 **	0.010
J	2	1	2	10	0.50	0.00	1.000	14.067 **	0.000
K	4	3	3	5	0.57	0.38	0.705	0.733	0.392
L	5	3	4	3	0.56	0.33	0.739	0.733	0.392
M	4	4	3	4	0.57	0.38	0.705	0.200	0.655
TOTAL	52	38	38	52					

* Statistically significant at the 5 percent level
** Statistically significant at the 1 percent level

Once again, the most persistent winner (A) and loser (B) funds continue to be statistically significant, but in this case the evidence is stronger. This is attested by the respective p-values of the Chi-square tests (0 and 0.019, respectively). Furthermore, it is important to note that this evidence is also consistent with the results of the Percentage Repeat winners p-values (0 and 0.025, respectively). When the returns are risk-adjusted, we can also reject the null hypothesis of no performance persistence for funds H and I (winning persistence) and F and J (losing persistence).[118]

Table 4.27 below summarizes the results of the performance persistence tests for individual funds:

Table 4.28

Summary of the performance persistence tests for individual funds

Whenever the initials W and L appear, it is evidence of a consistently winning (W) or losing (F) fund. If only one initial appears, it represents a rejection of the null hypothesis of no performance persistence according to the chi-square test. Two initials indicate rejection according to both the Chi-square and the Repeat Winner Z-test.

Fund	Quarterly Excess Returns	Monthly Excess Returns	Quarterly Alpha
A	W**	W**	W**,W**
B	L**	L**	L**,L**
C			
D			
F			L*,L**
G	W*		
H		L*	W*
I	W**		W**
J			L**
K	L*		
L	L*		
M	L**		

* Statistically significant at the 5 percent level
** Statistically significant at the 1 percent level

In summary, although the evidence of performance persistence for the overall sample of funds is only supported in a limited context (quarterly excess returns), we have found evidence that some funds are consistently winning (fund A) and losing (fund B) funds. We have also verified that this conclusion is robust both to the time period and to the performance measure used.

[118] The evidence for fund F is ambiguous because it is not a very clear Winner or Loser.

4.5. CONCLUSIONS

In this chapter we have examined in some detail the performance persistence of our sample of funds. Our concluding comments can be divided into those concerning the statistical tests and those concerning the performance persistence of funds.

Several criteria (Repeat Winners test, Odds Ratio test, Chi-square test, Yates Continuity Correction, Fisher's exact p-value and the bootstrap) have been applied and compared. There are several practical and conceptual problems in attempting to evaluate the performance persistence of equity fund managers relatively to these measures.

First, when only small samples are available, significant adjustments must be made to the test statistics. These corrections, which will complement the inferences based on the asymptotic tests, are accomplished through the Yates continuity and Fisher exact p-value, which give similar probability values that will generally turn out to be more conservative.[119] We also provide results from bootstrap simulations, in order to verify the significance of the persistence coefficients, and find that the p-values obtained are identical to Fisher's exact p-value. Therefore, the efficacy of these tests is clearly demonstrated.

[119] In fact, they have precisely been criticized for being apparently too conservative (for example, by CONOVER [1980]).

Secondly, these performance persistence criteria will not always be appropriate for all circumstances. The application of all these tests in the context of a small size sample lead to a number of general points we should note. The Odds Ratio test and Chi-square will generally lead to the same type of conclusions relative to fund persistence. The latter, though, has the disadvantage of not being able to detect reversals in performance, since it is always positive (the former will reflect reversals with a negative sign). Despite this weakness, the Chi-square test is more appropriate for testing the performance persistence of individual funds, whereas in this context the Odds Ratio, the Yates and the Fisher test have not proved feasible. Also, the Odds Ratio test also presents problems in terms of practical application, in the case where winners-losers and losers-winners are zero (which is not an uncommon situation in small sample sizes). Finally, both the Yates and Fisher tests, appropriate for small samples, might present problems: the former has presented problems in the case of equally weighted cells; the latter becomes computationally tedious in the case that where the smallest cell value is even moderately large. The bootstrap p-value is a potentially valid alternative in these cases.

In terms of fund persistence, the relative performance of quarterly excess returns for the Portuguese equity funds is consistent with the hypothesis of overall persistence, but not for quarters separately. Evidence of persistence is significantly reduced when fund returns are adjusted for risk (either total or systematic risk). Consideration of the persistence of risk characteristics separately is relevant, as we found strong evidence of repeat "risk-takers" and "risk averters". We also find that the persistence analysis should be conducted over different time intervals, both more and less frequently, as the term structure of performance persistence is considered.

Lastly, although in general there is not strong evidence of persistence (except on a quarterly basis), there is evidence that some funds, individually, are persistently winning or losing funds.

Of course, this methodology does not consider the degree of winning/losing of the funds. Hence the next logical step is to see whether a strategy of investing in the top-performing funds can produce returns statistically different from those obtained from the bottom-performing funds. This will be discussed in the next chapter.

APPENDICES

Appendix 4.1

Annual estimates of the regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon_{p,t}$ relative to the PSI-20 index

T-statistics reported in parenthesis below the α_p estimates, computed according to NEWEY and WEST [1987].

Fund		A	B	C	D	F	G	H	I	J	K	L	M
Apr-95-Mar.95	α_p	-0.00034	-0.00023	-0.00005	-0.00009	-0.00010	-0.00008	-0.00001	0.00001	-0.00012	-0.00020	0.00001	0.00006
		(-1.57)	(-1.17)	(-0.21)	(-0.61)	(-0.77)	(-0.36)	(-0.04)	(0.06)	(-0.45)	(-1.05)	(0.09)	(0.28)
	β_p	0.293	0.334	0.406	0.177	0.588	0.318	0.199	0.381	0.578	0.389	0.448	0.176
	R^2	0.417	0.552	0.397	0.177	0.856	0.350	0.244	0.549	0.574	0.558	0.637	0.254
Apr.95-Mar.96	α_p	0.00038	-0.00016	0.00006	-0.00003	-0.00035	0.00007	-0.00007	0.00009	-0.00045	-0.00062	0.00000	-0.00022
		(2.33)	(-0.99)	(0.36)	(-0.32)	(-2.18)	(0.31)	(-0.32)	(0.77)	(-2.44)	(-2.42)	(0.04)	(-0.82)
	β_p	0.543	0.569	0.652	0.558	0.704	0.581	0.409	0.668	0.657	0.390	0.653	0.379
	R^2	0.611	0.559	0.514	0.682	0.646	0.321	0.147	0.624	0.574	0.201	0.684	0.091
Apr.96-Mar.97	α_p	0.00070	0.00005	0.00010	0.00033	0.00015	0.00007	0.00031	0.00013	0.00013	0.00072	0.00010	0.00010
		(4.49)	(0.54)	(0.82)	(2.62)	(1.16)	(0.43)	(2.61)	(0.99)	(0.86)	(3.74)	(0.73)	(0.84)
	β_p	0.674	0.665	0.701	0.554	0.828	0.732	0.689	0.734	0.799	0.630	0.702	0.659
	R^2	0.648	0.829	0.682	0.669	0.813	0.742	0.750	0.799	0.785	0.537	0.779	0.717
Apr.97-Mar.98	α_p	0.00031	-0.00004	-0.00086	-0.00013	0.00007	0.00005	-0.00031	0.00003	-0.00011	-0.00020	0.00000	0.00025
		(1.67)	(-0.24)	(-1.58)	(-0.71)	(0.19)	(0.30)	(-1.44)	(0.25)	(-0.62)	(-0.75)	(0.02)	(0.65)
	β_p	0.751	0.796	1.119	0.796	0.834	0.837	0.783	0.868	0.866	0.796	0.854	0.836
	R^2	0.931	0.947	0.784	0.946	0.606	0.933	0.889	0.962	0.945	0.885	0.950	0.528

Appendix 4.2

WHITE's [1980] test for heteroscedasticity

The basic idea underlying the detection of heteroscedasticity involves examination of OLS residuals for evidence of the relation between σ_i^2 and the characteristics of the observations.

In order to test the null hypothesis of the form:

$$H_0 : \sigma_i^2 = \sigma \quad \text{for all } i$$

$$H_1 : \text{Not } H_0$$

WHITE [1980] has developed the following methodology:

1) Estimate the main regression for each fund, compute the respective residuals ($\hat{\varepsilon}_{p,t}$) and square them.

2) Regress $\hat{\varepsilon}^2_{p,t}$ against a constant, all the explanatory variables in the main regression as well as their squares and cross-products. White's test is based on the R^2 of this auxiliary regression.

3) The statistic nR^2 is asymptotically distributed as chi-squared with P-1 degrees of freedom, where P is the number of regressors (not including the constant) and n is the number of observations.

Appendix 4.3

The BREUSCH [1978] - GODFREY [1978] test for serial correlation

The Breush-Godfrey test is a Lagrange multiplier test of the form:

H_0: No autocorrelation

H_1: $\varepsilon_t = AR(p)$ or $\varepsilon_t = MA(p)$

and is carried out as follows:

1) Estimate the main regression and compute the respective residuals.

2) Regress the residuals against a constant, all explanatory variables and on the lagged residuals $e_{t-1},\ldots,e_{t-p,}$ and obtain R^2. The Breusch-Godfrey test is based on the R^2 of this auxiliary regression.

3) The statistic nR^2 is asymptotically chi-square distributed with P degrees of freedom, where P is the order of the process thought to be determining the disturbances and n is the number of observations.

This test is a joint test of the first P autocorrelations of the residuals, not just the first.

Appendix 4.4

OLS estimates and statistical significance tests corrected for heteroscedasticity (WHITE [1980]) and autocorrelation and heteroscedasticity (NEWEY and WEST [1987])

Fund	PSI α_p	PSI β_p	BVL-30 α_p	BVL-30 β_p	BVL-Geral α_p	BVL-Geral β_p
A OLS						
Regression estimates	0.000450	0.615	0.000406	0.648	0.000397	0.702
Se	0.000096	0.011521	0.000093	0.011661	0.000095	0.013013
t-stat	4.69	53.37	4.36	55.57	4.17	53.93
p-value	0.000003	0.000000	0.000014	0.000000	0.000033	0.000000
WHITE						
Se	0.000094	0.022363	0.000089	0.019928	0.000090	0.022049
t-stat	4.79	27.49	4.56	32.52	4.40	31.83
p-value	0.000002	0.000000	0.000006	0.000000	0.000012	0.000000
NEWEY and WEST						
Se	0.000115	0.033553	0.000107	0.029123	0.000106	0.029011
t-stat	3.89	18.32	3.80	22.25	3.75	24.19
p-value	0.000105	0.000000	0.000155	0.000000	0.000185	0.000000
B OLS						
Regression estimates	0.000088	0.649	0.000042	0.683	0.000036	0.737
Se	0.000088	0.010525	0.000085	0.010600	0.000089	0.012107
t-stat	1.00	61.67	0.50	64.46	0.40	60.87
p-value	0.317345	0.000000	0.615623	0.000000	0.686214	0.000000
WHITE						
Se	0.000088	0.026643	0.000084	0.023908	0.000086	0.026136
t-stat	1.00	24.36	0.51	28.58	0.42	28.20
p-value	0.317144	0.000000	0.611099	0.000000	0.677726	0.000000
NEWEY and WEST						
Se	0.000097	0.041028	0.000090	0.036230	0.000092	0.036853
t-stat	0.90	15.82	0.47	18.86	0.39	20.00
p-value	0.366079	0.000000	0.638125	0.000000	0.695846	0.000000
C OLS						
Regression estimates	0.000050	0.852	-0.000016	0.905	-0.000032	0.984
Se	0.000168	0.020136	0.000162	0.020301	0.000163	0.022276
t-stat	0.30	42.30	-0.10	44.57	-0.20	44.19
p-value	0.764296	0.000000	0.919209	0.000000	0.842019	0.000000
WHITE						
Se	0.000196	0.097326	0.000191	0.097374	0.000189	0.105008
t-stat	0.26	8.75	-0.09	9.29	-0.17	9.37
p-value	0.797512	0.000000	0.931490	0.000000	0.863835	0.000000
NEWEY and WEST						
Se	0.000165	0.137246	0.000162	0.137832	0.000161	0.148369
t-stat	0.30	6.21	-0.10	6.56	-0.20	6.63
p-value	0.761143	0.000000	0.919396	0.000000	0.840185	0.000000
D OLS						
Regression estimates	0.000244	0.596	0.000197	0.633	0.000186	0.689
Se	0.000095	0.011394	0.000090	0.011238	0.000091	0.012387
t-stat	2.57	52.27	2.19	56.35	2.05	55.60
p-value	0.010230	0.000000	0.028471	0.000000	0.040691	0.000000
WHITE						
Se	0.000110	0.043131	0.000104	0.039360	0.000101	0.038842
t-stat	2.21	13.81	1.90	16.09	1.84	17.73
p-value	0.027396	0.000000	0.058018	0.000000	0.065745	0.000000
NEWEY and WEST						
Se	0.000122	0.060627	0.000113	0.055351	0.000107	0.054931
t-stat	2.00	9.82	1.75	11.44	1.74	12.54
p-value	0.045396	0.000000	0.080948	0.000000	0.082594	0.000000

Appendix 4.4 (Continued)

F	Regression estimates	0.000054	0.763	0.000006	0.797	-0.000002	0.861
	Se	0.000148	0.017764	0.000148	0.018509	0.000151	0.020582
	t-stat	0.36	42.97	0.04	43.06	-0.01	41.81
	p-value	0.716716	0.000000	0.965434	0.000000	0.988341	0.000000
	WHITE						
	Se	0.000155	0.034416	0.000155	0.033018	0.000158	0.036687
	t-stat	0.35	22.18	0.04	24.14	-0.01	23.46
	p-value	0.729138	0.000000	0.967071	0.000000	0.988854	0.000000
	NEWEY and WEST						
	Se	0.000111	0.046316	0.000106	0.044348	0.000104	0.048000
	t-stat	0.48	16.48	0.06	17.97	-0.02	17.93
	p-value	0.628677	0.000000	0.951679	0.000000	0.983161	0.000000
G	OLS						
	Regression estimates	0.000240	0.675	0.000191	0.713	0.000163	0.794
	Se	0.000118	0.014159	0.000115	0.014360	0.000108	0.014816
	t-stat	2.04	47.69	1.67	49.64	1.50	53.59
	p-value	0.041778	0.000000	0.095492	0.000000	0.133266	0.000000
	WHITE						
	Se	0.000119	0.030429	0.000114	0.026690	0.000106	0.027999
	t-stat	2.01	22.19	1.67	26.71	1.54	28.36
	p-value	0.044396	0.000000	0.095158	0.000000	0.123076	0.000000
	NEWEY and WEST						
	Se	0.000118	0.045492	0.000109	0.039830	0.000100	0.038668
	t-stat	2.03	14.84	1.75	17.90	1.64	20.53
	p-value	0.042681	0.000000	0.079782	0.000000	0.101962	0.000000
H	OLS						
	Regression estimates	0.000230	0.594	0.000185	0.630	0.000172	0.687
	Se	0.000127	0.015233	0.000124	0.015495	0.000124	0.016909
	t-stat	1.82	38.99	1.50	40.63	1.39	40.62
	p-value	0.069275	0.000000	0.134541	0.000000	0.163691	0.000000
	WHITE						
	Se	0.000133	0.039676	0.000128	0.036457	0.000126	0.036042
	t-stat	1.73	14.97	1.44	17.27	1.37	19.06
	p-value	0.084049	0.000000	0.149710	0.000000	0.171305	0.000000
	NEWEY and WEST						
	Se	0.000138	0.058031	0.000128	0.053114	0.000123	0.052744
	t-stat	1.67	10.24	1.44	11.85	1.40	13.02
	p-value	0.095241	0.000000	0.148870	0.000000	0.162914	0.000000
I	OLS						
	Regression estimates	0.000257	0.715	0.000207	0.752	0.000183	0.831
	Se	0.000093	0.011142	0.000089	0.011201	0.000084	0.011504
	t-stat	2.77	64.15	2.32	67.15	2.18	72.22
	p-value	0.005748	0.000000	0.020726	0.000000	0.029807	0.000000
	WHITE						
	Se	0.000098	0.032829	0.000093	0.028100	0.000087	0.025884
	t-stat	2.61	21.77	2.22	26.77	2.11	32.10
	p-value	0.009202	0.000000	0.026429	0.000000	0.034878	0.000000
	NEWEY and WEST						
	Se	0.000098	0.045128	0.000089	0.039351	0.000083	0.036600
	t-stat	2.61	15.84	2.32	19.11	2.20	22.70
	p-value	0.009201	0.000000	0.020427	0.000000	0.027829	0.000000
J	OLS						
	Regression estimates	-0.000017	0.771	-0.000070	0.810	-0.000077	0.873
	Se	0.000101	0.012136	0.000098	0.012283	0.000103	0.014071
	t-stat	-0.17	63.49	-0.71	65.92	-0.75	62.06
	p-value	0.863974	0.000000	0.477291	0.000000	0.452159	0.000000
	WHITE						
	Se	0.000100	0.022633	0.000097	0.020563	0.000102	0.026189
	t-stat	-0.17	34.04	-0.72	39.38	-0.76	33.34
	p-value	0.862744	0.000000	0.473065	0.000000	0.449660	0.000000
	NEWEY and WEST						
	Se	0.000108	0.027803	0.000103	0.025178	0.000108	0.030077
	t-stat	-0.16	27.71	-0.68	32.16	-0.72	29.03
	p-value	0.872585	0.000000	0.499153	0.000000	0.473608	0.000000

Appendix 4.4 (Continued)

K	OLS						
	Regression estimates	0.000075	0.646	0.000024	0.686	0.000004	0.757
	Se	0.000117	0.014087	0.000113	0.014116	0.000110	0.015019
	t-stat	0.64	45.88	0.22	48.63	0.04	50.37
	p-value	0.523066	0.000000	0.829034	0.000000	0.971902	0.000000
	WHITE						
	Se	0.000123	0.030548	0.000118	0.028357	0.000116	0.029325
	t-stat	0.61	21.16	0.21	24.21	0.03	25.80
	p-value	0.542903	0.000000	0.836669	0.000000	0.973288	0.000000
	NEWEY and WEST						
	Se	0.000147	0.043838	0.000140	0.040414	0.000135	0.040914
	t-stat	0.51	14.74	0.17	16.98	0.03	18.49
	p-value	0.609692	0.000000	0.861680	0.000000	0.977055	0.000000
L	OLS						
	Regression estimates	0.000181	0.719	0.000134	0.754	0.000126	0.814
	Se	0.000087	0.010489	0.000085	0.010693	0.000090	0.012281
	t-stat	2.08	68.58	1.57	70.53	1.40	66.28
	p-value	0.038233	0.000000	0.117051	0.000000	0.161123	0.000000
	WHITE						
	Se	0.000088	0.026453	0.000085	0.022623	0.000086	0.024131
	t-stat	2.06	27.19	1.58	33.33	1.46	33.73
	p-value	0.039498	0.000000	0.114108	0.000000	0.144914	0.000000
	NEWEY and WEST						
	Se	0.000087	0.036604	0.000079	0.031612	0.000080	0.032429
	t-stat	2.08	19.65	1.69	23.86	1.57	25.10
	p-value	0.037704	0.000000	0.090395	0.000000	0.116071	0.000000
M	OLS						
	Regression estimates	0.000312	0.616	0.000261	0.657	0.000250	0.714
	Se	0.000201	0.024169	0.000198	0.024843	0.000199	0.027182
	t-stat	1.55	25.49	1.32	26.45	1.26	26.27
	p-value	0.121388	0.000000	0.188042	0.000000	0.208667	0.000000
	WHITE						
	Se	0.000220	0.048339	0.000217	0.046552	0.000216	0.046930
	t-stat	1.42	12.75	1.20	14.11	1.16	15.21
	p-value	0.156222	0.000000	0.229493	0.000000	0.246462	0.000000
	NEWEY and WEST						
	Se	0.000150	0.059959	0.000141	0.055613	0.000137	0.054608
	t-stat	2.07	10.28	1.85	11.81	1.82	13.07
	p-value	0.038382	0.000000	0.064989	0.000000	0.068771	0.000000
MEAN	OLS						
	Regression estimates	0.000180	0.684	0.000131	0.722	0.000117	0.787
	Se	0.000077	0.009268	0.000072	0.009011	0.000073	0.009939
	t-stat	2.34	73.83	1.82	80.16	1.61	79.16
	p-value	0.019600	0.000000	0.069500	0.000000	0.107700	0.000000
	WHITE						
	Se	0.000084	0.032076	0.000077	0.028150	0.000075	0.028283
	t-stat	2.15	21.33	1.70	25.66	1.55	27.82
	p-value	0.032000	0.000000	0.089500	0.000000	0.120500	0.000000
	NEWEY and WEST						
	Se	0.000093	0.047658	0.000082	0.042523	0.000078	0.042397
	t-stat	1.95	14.36	1.60	16.99	1.51	18.56
	p-value	0.051900	0.000000	0.109200	0.000000	0.131800	0.000000

Appendix 4.5

The non-parametric HENRIKSSON and MERTON [1981] test

HENRIKSSON and MERTON [1981] (HM) develop a nonparametric test for analysing the statistical significance of the correlation between a market timer's forecast and the actual values of excess returns.

The test examines the null hypothesis that the market timer has no forecasting ability, that is:

$$H_0: p_1(t)+p_2(t)=1$$

where the conditional probabilities are defined as:

$$p_1 = \text{prob [forecast } (R_{m,t} \leq R_{f,t}) \mid R_{m,t} \leq R_{f,t})]$$

$$p_1 = \text{prob [forecast } (R_{m,t} > R_{f,t}) \mid R_{m,t} > R_{f,t})]$$

To determine the probability that a given outcome came from a population that satisfies the null hypothesis, HM define:

N_1 = number of observations where $R_{m,t} \leq R_{f,t}$;

N_2 = number of observations where $R_{m,t} > R_{f,t}$;

$N = N_1 + N_2$ = total number of observations;

n_1 = number of correct forecasts, given $R_{m,t} \leq R_{f,t}$;

n_2 = number of correct forecasts, given $R_{m,t} > R_{f,t}$;

which represents a two-dimensional contingency table depending on the signs of the predicted and realized values of the excess return:

	$R_{m,t} \leq R_{f,t}$	$R_{m,t} > R_{f,t}$	
$R_{m,t} \leq R_{f,t}$	n_1	n_2	n
$R_{m,t} > R_{f,t}$			N
	N_1	N_2	

HM show that under H_0:

$$P\left(n_1 \mid N_1, N, n\right) = \frac{\binom{N_1}{n}\binom{N_2}{n - n_1}}{\binom{N}{n}}$$

The probability distribution as given by the previous equation may be recognized as the probability of the hypergeometric distribution. Rearranging in terms of an $\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}$ table, the previous equation can be rewritten as:

$$\frac{\binom{a+c}{a}\binom{b+d}{b}}{\binom{N}{a+b}}$$

$$= \frac{\dfrac{(a+c)!}{a!\,((a+c)-a)!}\,\dfrac{(b+d)!}{b!\,((b+d)-b)!}}{\dfrac{N!}{(a+b)!\,(N-(a+b))!}}$$

$$= \frac{(a+c)!\,(b+d)\,(a \div b)!\,(c+d)!}{N!\,a!\,b!\,c!\,d!}$$

Hence, the nonparametric test of market timing proposed by HM is equal to Fisher's exact test in a 2 x 2 table.

Appendix 4.6

Frequency of repeat winners and losers: Sharpe returns

	A	B	C	D	F	G	H	I	J	K	L	M	MEDIAN
2Q94	-0.36728	-0.34773	-0.35348	-0.45126	-0.32779	-0.37517	-0.48387	-0.32103	-0.27234	-0.19057	-0.29325	-0.50329	-0.35061
3Q94	-0.11486	0.12639	0.19061	0.23499	0.14925	0.08856	0.22214	0.14299	0.17725	0.05782	0.19704	0.21321	0.16325
4Q94	0.13745	-0.00543	0.04409	0.03696	-0.10390	0.11094	0.23493	0.01247	0.07197	-0.16896	0.07222	0.29816	0.05803
1Q95	-0.15206	-0.19039	-0.13697	-0.22747	-0.15473	-0.07413	-0.29548	-0.05126	-0.19717	-0.16447	-0.28315	-0.29083	-0.17743
2Q95	0.14794	0.05297	0.00259	0.15872	-0.02811	0.04774	0.01796	0.13059	-0.01501	-0.10696	0.07635	0.01317	0.03285
3Q95	-0.21096	-0.31581	-0.25706	-0.30944	-0.28187	-0.24189	-0.28398	-0.23990	-0.26550	-0.37360	-0.09486	-0.35801	-0.27369
4Q95	0.05559	-0.19621	-0.18048	-0.20376	-0.28075	0.00214	-0.06242	-0.04399	-0.39305	-0.31080	-0.17212	-0.06884	-0.17630
1Q96	0.47255	0.21776	0.34233	0.28676	0.25395	0.29465	0.25725	0.27900	0.23067	0.30405	0.30976	0.26716	0.28288
2Q96	0.46320	0.34943	0.38772	0.47794	0.36203	0.30898	0.40131	0.26587	0.35829	0.41436	0.33568	0.30881	0.36016
3Q96	0.23364	-0.07681	-0.00411	0.14873	-0.04166	-0.04036	0.09589	-0.04892	0.07171	0.24649	-0.00165	-0.01032	-0.00288
4Q96	0.40769	0.31462	0.31393	0.25560	0.27917	0.29864	0.34977	0.33117	0.29345	0.45413	0.30116	0.26990	0.30754
1Q97	0.43772	0.32237	0.29137	0.37153	0.36946	0.31823	0.33576	0.36227	0.29524	0.34397	0.31024	0.32947	0.33261
2Q97	0.40195	0.37876	0.31666	0.36151	0.38264	0.42926	0.35038	0.39854	0.37992	0.37980	0.39107	0.35953	0.37986
3Q97	0.15513	0.09182	0.12728	0.08968	0.07680	0.12280	0.08546	0.15162	0.12039	0.17996	0.12773	0.06768	0.12159
4Q97	0.02774	0.02921	-0.00312	-0.00413	0.00072	0.02167	-0.02924	0.03544	0.01710	-0.06964	0.02215	0.05910	0.01939
1Q98	0.59201	0.58248	0.50392	0.56841	0.59918	0.57707	0.54555	0.59354	0.58145	0.60095	0.57565	0.60742	0.58196
GLOBAL PERIOD	0.16877	0.10856	0.09058	0.13389	0.09259	0.12628	0.12211	0.13514	0.09419	0.10018	0.12364	0.10597	0.11434

REPEATS		A	B	C	D	F	G	H	I	J	K	L	M	SUM	TOTAL
1994 3Q	WW	0	0	0	0	0	0	0	0	1	0	1	0	2	
	LW	0	0	1	1	0	0	1	0	0	0	0	1	4	
	WL	0	1	0	0	1	0	0	1	0	1	0	0	4	
	LL	1	0	0	0	0	1	0	0	0	0	0	0	2	12
1994 4Q	WW	0	0	0	0	0	0	1	0	1	0	1	1	4	
	LW	1	0	0	0	0	1	0	0	0	0	0	0	2	
	WL	0	0	1	1	0	0	0	0	0	0	0	0	2	
	LL	0	1	0	0	1	0	0	1	0	1	0	0	4	12
1995 1Q	WW	1	0	0	0	0	1	0	0	0	0	0	0	2	
	LW	0	0	1	0	1	0	0	1	0	1	0	0	4	
	WL	0	0	0	0	0	0	1	0	1	0	1	1	4	
	LL	0	1	0	1	0	0	0	0	0	0	0	0	2	12
1995 2Q	WW	1	0	0	0	0	1	0	1	0	0	0	0	3	
	LW	0	1	0	1	0	0	0	0	0	0	1	0	3	
	WL	0	0	1	0	1	0	0	0	0	1	0	0	3	
	LL	0	0	0	0	0	0	1	0	1	0	0	1	3	12
1995 3Q	WW	1	0	0	0	0	1	0	1	0	0	1	0	4	
	LW	0	0	1	0	0	0	0	0	1	0	0	0	2	
	WL	0	1	0	1	0	0	0	0	0	0	0	0	2	
	LL	0	0	0	0	1	0	1	0	0	1	0	1	4	12
1995 4Q	WW	1	0	0	0	0	1	0	1	0	0	1	0	4	
	LW	0	0	0	0	0	0	1	0	0	0	0	1	2	
	WL	0	0	1	0	0	0	0	0	1	0	0	0	2	
	LL	0	1	0	1	1	0	0	0	0	1	0	0	4	12
1996 1Q	WW	1	0	0	0	0	1	0	0	0	0	1	0	3	
	LW	0	0	1	1	0	0	0	0	0	1	0	0	3	
	WL	0	0	0	0	0	0	1	1	0	0	0	1	3	
	LL	0	1	0	0	1	0	0	0	1	0	0	0	3	12
1996 2Q	WW	1	0	1	1	0	0	0	0	0	1	0	0	4	
	LW	0	0	0	0	1	0	1	0	0	0	0	0	2	
	WL	0	0	0	0	0	1	0	0	0	0	1	0	2	
	LL	0	1	0	0	0	0	0	1	1	0	0	1	4	12
1996 3Q	WW	1	0	0	1	0	0	1	0	0	1	0	0	4	
	LW	0	0	0	0	0	0	0	0	1	0	1	0	2	
	WL	0	0	1	0	1	0	0	0	0	0	0	0	2	
	LL	0	1	0	0	0	1	0	1	0	0	0	1	4	12
1996 4Q	WW	1	0	0	0	0	0	1	0	0	1	0	0	3	
	LW	0	1	1	0	0	0	0	1	0	0	0	0	3	
	WL	0	0	0	1	0	0	0	0	1	0	1	0	3	
	LL	0	0	0	0	1	1	0	0	0	0	0	1	3	12
1997 1Q	WW	1	0	0	0	0	0	1	1	0	1	0	0	4	
	LW	0	0	0	1	1	0	0	0	0	0	0	0	2	
	WL	0	1	1	0	0	0	0	0	0	0	0	0	2	
	LL	0	0	0	0	0	1	0	0	1	0	1	1	4	12
1997 2Q	WW	1	0	0	0	1	0	0	1	0	0	0	0	3	
	LW	0	0	0	0	0	1	0	0	1	0	1	0	3	
	WL	0	0	0	1	0	0	1	0	0	1	0	0	3	
	LL	0	1	1	0	0	0	0	0	0	0	0	1	3	12
1997 3Q	WW	1	0	0	0	0	1	0	1	0	0	1	0	4	
	LW	0	0	1	0	0	0	0	0	0	1	0	0	2	
	WL	0	0	0	0	1	0	0	0	1	0	0	0	2	
	LL	0	1	0	1	0	0	1	0	0	0	0	1	4	12
1997 4Q	WW	1	0	0	0	0	1	0	1	0	0	1	0	4	
	LW	0	1	0	0	0	0	0	0	0	0	0	1	2	
	WL	0	0	1	0	0	0	0	0	0	1	0	0	2	
	LL	0	0	0	1	1	0	1	0	1	0	0	0	4	12
1998 1Q	WW	1	1	0	0	0	0	0	1	0	0	0	1	4	
	LW	0	0	0	0	1	0	0	0	0	1	0	0	2	
	WL	0	0	0	0	0	1	0	0	0	0	1	0	2	
	LL	0	0	1	1	0	0	1	0	1	0	0	0	4	12
CONSISTENCY OF INDIVIDUAL FUND PERFORMANCE															
1994-1997	WW	13	1	1	2	1	7	4	8	2	4	7	2	52	
	LW	1	3	6	4	4	2	3	2	3	4	3	3	38	
	WL	0	3	6	4	4	2	3	2	4	4	4	2	38	
	LL	1	8	2	5	6	4	5	3	6	3	1	8	52	180

Appendix 4.7

Quarterly estimates of the regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon_{p,t}$ relative to the PSI-20 index

A shadowed area relatively to the estimated alpha represents a value statistically significant at the 5 percent level. If is outlined, it is statistically significant at the 1 percent level. The statistical significance is based on the NEWEY and WEST [1987] t-statistics.

		A	B	C	D	F	G	H	I	J	K	L	M
2Q94	α	-0.00107	-0.00052	-0.00086	-0.00065	-0.00016	-0.00058	-0.00099	-0.00028	-0.00018	0.00050	0.00002	-0.00065
	β	0.32	0.36	0.42	0.15	0.57	0.35	0.16	0.42	0.70	0.48	0.48	0.15
	R^2	48.9%	64.7%	53.8%	44.0%	95.7%	58.2%	34.4%	73.7%	77.3%	79.3%	81.8%	58.3%
3Q94	α	-0.00014	0.00018	0.00076	0.00045	0.00016	0.00010	0.00074	0.00028	0.00056	-0.00004	0.00045	0.00066
	β	0.09	0.18	0.32	0.16	0.53	0.13	0.24	0.24	0.30	0.14	0.36	0.20
	R^2	19.0%	28.3%	22.0%	25.7%	79.2%	16.9%	19.4%	26.1%	26.4%	43.7%	47.1%	17.2%
4Q94	α	0.00044	-0.000004	0.00022	0.00009	-0.00007	0.00050	0.00044	0.00010	0.00032	-0.00029	0.00029	0.00050
	β	0.47	0.26	0.36	0.13	0.62	0.28	0.19	0.31	0.40	0.19	0.31	0.16
	R^2	40.4%	30.4%	15.9%	8.7%	64.2%	4.2%	10.9%	25.7%	17.9%	10.7%	11.0%	9.8%
1Q95	α	-0.00022	-0.00030	-0.00027	-0.00035	-0.00024	-0.00003	-0.00048	0.00018	-0.00059	-0.00053	-0.00053	-0.00045
	β	0.41	0.45	0.38	0.27	0.76	0.45	0.19	0.46	0.57	0.38	0.49	0.15
	R^2	49.8%	63.2%	31.3%	39.6%	80.9%	44.6%	20.1%	50.0%	46.3%	4.2%	66.3%	12.0%
2Q95	α	0.00029	0.00005	0.00017	0.00022	-0.00021	0.00011	-0.00010	0.00014	-0.00017	-0.00037	0.00009	-0.00016
	β	0.48	0.42	0.43	0.46	0.49	0.48	0.14	0.33	0.61	0.10	0.59	0.15
	R^2	49.8%	27.1%	15.5%	54.3%	35.5%	24.9%	3.2%	28.4%	53.5%	1.8%	46.4%	2.6%
3Q95	α	-0.00011	-0.00039	-0.00012	-0.00033	-0.00025	-0.00049	-0.00038	-0.00008	-0.00039	-0.00097	0.00024	-0.00061
	β	0.54	0.52	0.72	0.46	0.69	0.37	0.42	0.48	0.53	0.52	0.79	0.36
	R^2	51.4%	49.3%	56.8%	51.3%	69.2%	15.4%	28.7%	63.4%	32.3%	18.3%	70.3%	21.6%
4Q95	α	0.00045	-0.00038	-0.00044	-0.00039	-0.00097	0.00026	-0.00038	0.00010	-0.00139	-0.00170	-0.00044	-0.00060
	β	0.50	0.48	0.61	0.52	0.71	0.63	0.37	0.84	0.65	0.31	0.63	0.32
	R^2	61.4%	57.6%	63.1%	72.6%	59.9%	20.2%	5.3%	61.9%	55.8%	10.1%	72.4%	2.5%
1Q96	α	0.00085	-0.00016	0.00057	0.00015	-0.00006	0.00021	0.00044	0.00005	-0.00005	0.00057	0.00025	0.00035
	β	0.57	0.70	0.71	0.64	0.77	0.65	0.50	0.76	0.70	0.44	0.63	0.47
	R^2	67.2%	70.3%	65.5%	75.7%	78.7%	72.3%	38.1%	80.4%	70.2%	53.8%	77.3%	46.3%
2Q96	α	0.00075	0.00007	0.00036	0.00037	0.00027	0.00001	0.00034	-0.00015	0.00007	0.00083	0.00003	0.00013
	β	0.31	0.60	0.58	0.44	0.59	0.61	0.58	0.76	0.65	0.34	0.56	0.51
	R^2	27.2%	69.3%	54.9%	52.6%	59.6%	67.2%	55.0%	85.8%	70.0%	16.0%	75.9%	46.9%
3Q96	α	0.00047	-0.00013	0.00003	0.00041	-0.00012	-0.00009	0.00026	-0.00023	0.00023	0.00081	0.00000	-0.00006
	β	0.39	0.56	0.63	0.37	0.66	0.53	0.57	0.70	0.69	0.46	0.50	0.49
	R^2	43.3%	68.6%	43.0%	36.9%	65.1%	65.4%	64.3%	79.1%	70.8%	28.1%	66.4%	55.1%
4Q96	α	0.00066	0.00020	0.00011	-0.00001	0.00005	-0.00001	0.00033	0.00042	0.00004	0.00083	0.00004	0.00015
	β	0.80	0.70	0.70	0.69	0.91	0.89	0.70	0.76	0.98	0.73	0.84	0.60
	R^2	56.1%	70.9%	75.1%	68.8%	76.9%	80.3%	58.9%	52.4%	81.2%	55.6%	78.6%	49.2%
1Q97	α	0.00106	0.00005	-0.00002	0.00046	0.00049	0.00026	0.00034	0.00047	0.00012	0.00047	0.00026	0.00027
	β	0.79	0.69	0.75	0.59	0.89	0.76	0.73	0.71	0.82	0.72	0.75	0.73
	R^2	81.5%	92.3%	75.7%	76.3%	90.8%	76.0%	86.1%	88.0%	81.7%	73.0%	81.0%	86.6%
2Q97	α	0.00031	0.00017	0.00010	-0.00001	0.00017	0.00063	0.00028	0.00023	0.00018	0.00019	0.00028	0.00045
	β	0.74	0.78	0.82	0.74	0.84	0.80	0.67	0.79	0.88	0.60	0.87	0.69
	R^2	90.9%	91.9%	81.5%	91.5%	93.0%	87.1%	64.9%	94.1%	91.6%	87.1%	92.2%	62.2%
3Q97	α	0.00018	-0.00034	-0.00022	-0.00031	0.00036	-0.00016	-0.00039	0.00009	-0.00018	0.00052	-0.00010	-0.00012
	β	0.78	0.71	0.88	0.71	0.62	0.80	0.71	0.88	0.83	0.91	0.77	0.92
	R^2	96.5%	97.0%	82.3%	95.8%	18.5%	94.2%	91.7%	97.3%	94.2%	85.8%	93.8%	27.2%
4Q97	α	-0.00013	-0.00023	-0.00106	-0.00076	-0.00068	-0.00033	-0.00101	-0.00020	-0.00043	-0.00136	-0.00034	0.00024
	β	0.69	0.84	1.47	0.81	0.93	0.87	0.83	0.92	0.90	0.81	0.89	0.82
	R^2	93.4%	97.3%	84.9%	97.8%	97.5%	96.5%	95.4%	98.8%	96.1%	94.1%	96.7%	93.6%
1Q98	α	0.00043	0.00041	0.00005	0.00046	0.00061	0.00038	0.00006	0.00056	0.00033	0.00049	0.00024	0.00058
	β	0.83	0.80	0.93	0.87	0.86	0.82	0.84	0.81	0.81	0.75	0.86	0.87
	R^2	92.2%	90.7%	77.6%	92.8%	91.2%	90.2%	92.6%	91.0%	94.0%	89.9%	95.7%	92.5%

Appendix 4.8

Frequency of repeat winners and losers: alpha returns

A shadowed area for the quarterly estimated alpha represents a value statistically significant at the 5 percent level. If is outlined, it is statistically significant at the 1 percent level. The statistical significance is based on the NEWEY and WEST [1987] t-statistics.

	A	B	C	D	F	G	H	I	J	K	L	M	MEDIAN
2Q94	-0.00107	-0.00052	-0.00086	-0.00065	-0.00016	-0.00058	-0.00099	-0.00028	-0.00018	0.00050	0.00002	-0.00065	-0.00055
3Q94	-0.00014	0.00018	0.00076	0.00045	0.00016	0.00010	0.00074	0.00028	0.00056	-0.00004	0.00045	0.00066	0.00036
4Q94	0.00044	0.00000	0.00022	0.00009	-0.00007	0.00050	0.00044	0.00010	0.00032	-0.00029	0.00029	0.00050	0.00025
1Q95	-0.00022	-0.00030	-0.00027	-0.00035	-0.00024	-0.00003	-0.00048	0.00018	-0.00059	-0.00053	-0.00053	-0.00045	-0.00033
2Q95	0.00029	0.00005	0.00017	0.00022	-0.00021	0.00011	-0.00010	0.00014	-0.00017	-0.00037	0.00009	-0.00016	0.00007
3Q95	-0.00011	-0.00039	-0.00012	-0.00033	-0.00025	-0.00049	-0.00038	-0.00008	-0.00039	-0.00097	0.00024	-0.00081	-0.00035
4Q95	0.00045	-0.00038	-0.00044	-0.00039	-0.00097	0.00026	-0.00038	0.00010	-0.00139	-0.00170	-0.00044	-0.00060	-0.00041
1Q96	0.00085	-0.00016	0.00057	0.00015	-0.00006	0.00021	0.00044	0.00005	-0.00005	0.00057	0.00025	0.00035	0.00023
2Q96	0.00075	0.00007	0.00036	0.00037	0.00027	0.00001	0.00034	-0.00015	0.00007	0.00083	0.00003	0.00013	0.00020
3Q96	0.00047	-0.00013	0.00003	0.00041	-0.00012	-0.00009	0.00026	-0.00023	0.00023	0.00081	0.00000	-0.00006	0.00002
4Q96	0.00066	0.00020	0.00011	-0.00001	0.00005	-0.00001	0.00033	0.00042	0.00004	0.00083	0.00004	0.00015	0.00013
1Q97	0.00106	0.00005	-0.00002	0.00046	0.00049	0.00026	0.00034	0.00047	0.00012	0.00047	0.00026	0.00027	0.00030
2Q97	0.00031	0.00017	0.00010	-0.00001	0.00017	0.00063	0.00028	0.00023	0.00018	0.00019	0.00028	0.00045	0.00021
3Q97	0.00018	-0.00034	-0.00022	-0.00031	0.00036	-0.00016	-0.00039	0.00009	-0.00018	0.00052	-0.00010	-0.00012	-0.00014
4Q97	-0.00013	-0.00023	-0.00106	-0.00076	-0.00068	-0.00033	-0.00101	-0.00020	-0.00043	-0.00136	-0.00034	0.00024	-0.00038
1Q98	0.00043	0.00041	0.00005	0.00046	0.00061	0.00038	0.00006	0.00056	0.00033	0.00049	0.00024	0.00058	0.00042
TOTAL	0.00045	0.00009	0.00005	0.00024	0.00005	0.00024	0.00023	0.00026	-0.00002	0.00007	0.00018	0.00031	0.00021

REPEATS		A	B	C	D	F	G	H	I	J	K	L	M	SUM	TOTAL
1994 3Q	WW	0	0	0	0	0	0	0	0	1	0	1	0	2	
	LW	0	0	1	1	0	0	1	0	0	0	0	1	4	
	WL	0	1	0	0	1	0	0	1	0	1	0	0	4	
	LL	1	0	0	0	0	1	0	0	0	0	0	0	2	12
1994 4Q	WW	0	0	0	0	0	0	1	0	1	0	1	1	4	
	LW	1	0	0	0	0	1	0	0	0	0	0	0	2	
	WL	0	0	1	1	0	0	0	0	0	0	0	0	2	
	LL	0	1	0	0	1	0	0	1	0	1	0	0	4	12
1995 1Q	WW	1	0	0	0	0	1	0	0	0	0	0	0	2	
	LW	0	1	1	0	1	0	0	1	0	0	0	0	4	
	WL	0	0	0	0	0	0	1	0	1	0	1	1	4	
	LL	0	0	0	1	0	0	0	0	0	1	0	0	2	12
1995 2Q	WW	1	0	1	0	0	1	0	1	0	0	0	0	4	
	LW	0	0	0	1	0	0	0	0	0	0	1	0	2	
	WL	0	1	0	0	1	0	0	0	0	0	0	0	2	
	LL	0	0	0	0	0	0	1	0	1	1	0	1	4	12
1995 3Q	WW	1	0	1	1	0	0	0	1	0	0	1	0	5	
	LW	0	0	0	0	1	0	0	0	0	0	0	0	1	
	WL	0	0	0	0	0	1	0	0	0	0	0	0	1	
	LL	0	1	0	0	0	0	1	0	1	1	0	1	5	12
1995 4Q	WW	1	0	0	1	0	0	0	1	0	0	0	0	3	
	LW	0	1	0	0	0	1	1	0	0	0	0	0	3	
	WL	0	0	1	0	1	0	0	0	0	0	1	0	3	
	LL	0	0	0	0	0	0	0	0	1	1	0	1	3	12
1996 1Q	WW	1	0	0	0	0	0	1	0	0	0	0	0	2	
	LW	0	0	1	0	0	0	0	0	0	1	1	1	4	
	WL	0	1	0	1	0	1	0	1	0	0	0	0	4	
	LL	0	0	0	0	1	0	0	0	1	0	0	0	2	12
1996 2Q	WW	1	0	1	0	0	0	1	0	0	1	0	0	4	
	LW	0	0	0	1	1	0	0	0	0	0	0	0	2	
	WL	0	0	0	0	0	0	0	0	0	0	1	1	2	
	LL	0	1	0	0	0	1	0	1	1	0	0	0	4	12
1996 3Q	WW	1	0	1	1	0	0	1	0	0	1	0	0	5	
	LW	0	0	0	0	0	0	0	0	1	0	0	0	1	
	WL	0	0	0	0	1	0	0	0	0	0	0	0	1	
	LL	0	1	0	0	0	1	0	1	0	0	1	1	5	12
1996 4Q	WW	1	0	0	0	0	0	1	0	0	1	0	0	3	
	LW	0	1	0	0	0	0	0	1	0	0	0	1	3	
	WL	0	0	1	1	0	0	0	0	1	0	0	0	3	
	LL	0	0	0	0	1	1	0	0	0	0	1	0	3	12
1997 1Q	WW	1	0	0	0	0	0	1	1	0	1	0	0	4	
	LW	0	0	0	1	1	0	0	0	0	0	0	0	2	
	WL	0	1	0	0	0	0	0	0	0	0	0	1	2	
	LL	0	0	1	0	0	1	0	0	1	0	1	0	4	12
1997 2Q	WW	1	0	0	0	0	0	1	1	0	0	0	0	3	
	LW	0	0	0	0	0	1	0	0	0	0	1	1	3	
	WL	0	0	0	1	1	0	0	0	0	1	0	0	3	
	LL	0	1	1	0	0	0	0	0	1	0	0	0	3	12
1997 3Q	WW	1	0	0	0	0	0	0	1	0	0	1	1	4	
	LW	0	0	0	0	1	0	0	0	0	1	0	0	2	
	WL	0	0	0	0	0	1	1	0	0	0	0	0	2	
	LL	0	1	1	1	0	0	0	0	1	0	0	0	4	12
1997 4Q	WW	1	0	0	0	0	0	0	1	0	0	1	1	4	
	LW	0	1	0	0	0	1	0	0	0	0	0	0	2	
	WL	0	0	0	0	1	0	0	0	0	1	0	0	2	
	LL	0	0	1	1	0	0	1	0	1	0	0	0	4	12
1998 1Q	WW	1	0	0	0	0	0	0	1	0	0	0	1	3	
	LW	0	0	0	1	1	0	0	0	0	1	0	0	3	
	WL	0	1	0	0	0	1	0	0	0	0	1	0	3	
	LL	0	0	1	0	0	0	1	0	1	0	0	0	3	12
CONSISTENCY OF INDIVIDUAL FUND PERFORMANCE															
1994-1997	WW	13	0	4	3	0	2	7	8	2	4	5	4	52	
	LW	1	4	3	5	6	4	2	2	1	3	3	4	38	
	WL	0	5	3	4	6	4	2	2	2	3	4	3	38	
	LL	1	6	5	3	3	5	4	3	10	5	3	4	52	180

Appendix 4.9

Frequency of repeat winners and losers: standard deviation

	A	B	C	D	F	G	H	I	J	K	L	M	MEDIAN
2Q94	0.00530	0.00581	0.00740	0.00282	0.00733	0.00570	0.00340	0.00611	0.01001	0.00676	0.00681	0.00256	0.00596
3Q94	0.00283	0.00262	0.00536	0.00251	0.00459	0.00260	0.00425	0.00354	0.00456	0.00167	0.00406	0.00373	0.00364
4Q94	0.00237	0.00139	0.00303	0.00150	0.00252	0.00459	0.00187	0.00212	0.00307	0.00192	0.00310	0.00169	0.00225
1Q95	0.00361	0.00356	0.00418	0.00257	0.00521	0.00414	0.00253	0.00400	0.00513	0.00481	0.00361	0.00246	0.00381
2Q95	0.00267	0.00314	0.12503	0.00244	0.00326	0.00376	0.00317	0.00245	0.00329	0.00306	0.00337	0.00387	0.00321
3Q95	0.00262	0.00259	0.00316	0.00224	0.00285	0.00328	0.00269	0.00208	0.00319	0.00402	0.00338	0.00267	0.00277
4Q95	0.00306	0.00309	0.00378	0.00299	0.00453	0.00694	0.00782	0.00525	0.00424	0.00486	0.00368	0.01017	0.00439
1Q96	0.00402	0.00481	0.00510	0.00426	0.00503	0.00443	0.00473	0.00490	0.00486	0.00371	0.00418	0.00402	0.00458
2Q96	0.00253	0.00308	0.00336	0.00247	0.00327	0.00319	0.00336	0.00346	0.00333	0.00356	0.00273	0.00321	0.00324
3Q96	0.00208	0.00239	0.00336	0.00224	0.00283	0.00229	0.00252	0.00289	0.00287	0.00309	0.00216	0.00234	0.00246
4Q96	0.00389	0.00307	0.00294	0.00306	0.00383	0.00360	0.00332	0.00386	0.00396	0.00362	0.00343	0.00312	0.00351
1Q97	0.00710	0.00582	0.00699	0.00551	0.00761	0.00713	0.00633	0.00603	0.00734	0.00690	0.00669	0.00630	0.00680
2Q97	0.00729	0.00770	0.00864	0.00717	0.00820	0.00816	0.00791	0.00788	0.00883	0.00608	0.00870	0.00835	0.00803
3Q97	0.00926	0.00846	0.01134	0.00856	0.01688	0.00963	0.00869	0.01034	0.00990	0.01150	0.00924	0.02065	0.00977
4Q97	0.01106	0.01321	0.02474	0.01278	0.01469	0.01385	0.01317	0.01435	0.01432	0.01295	0.01403	0.01309	0.01353
1Q98	0.00921	0.00897	0.01128	0.00973	0.00960	0.00915	0.00927	0.00902	0.00894	0.00841	0.00933	0.00959	0.00924
GLOBAL PERIOD	0.00588	0.00602	0.00877	0.00573	0.00781	0.00668	0.00629	0.00657	0.00709	0.00646	0.00654	0.00807	0.00656

REPEATS		A	B	C	D	F	G	H	I	J	K	L	M	SUM	TOTAL
1994 3Q	WW	0	0	1	0	1	0	0	0	1	0	1	0	4	
	LW	0	0	0	0	0	0	1	0	0	0	0	1	2	
	WL	0	0	0	0	0	0	0	1	0	1	0	0	2	
	LL	1	1	0	1	0	1	0	0	0	0	0	0	4	12
1994 4Q	WW	0	0	1	0	1	0	0	0	1	0	1	0	4	
	LW	1	0	0	0	0	1	0	0	0	0	0	0	2	
	WL	0	0	0	0	0	0	1	0	0	0	0	1	2	
	LL	0	1	0	1	0	0	0	1	0	1	0	0	4	12
1995 1Q	WW	0	0	1	0	1	1	0	0	1	0	0	0	4	
	LW	0	0	0	0	0	0	0	1	0	1	0	0	2	
	WL	1	0	0	0	0	0	0	0	0	0	1	0	2	
	LL	0	1	0	1	0	0	1	0	0	0	0	1	4	12
1995 2Q	WW	0	0	1	0	1	1	0	0	1	0	0	0	4	
	LW	0	0	0	0	0	0	0	0	0	0	1	1	2	
	WL	0	0	0	0	0	0	0	1	0	1	0	0	2	
	LL	1	1	0	1	0	0	1	0	0	0	0	0	4	12
1995 3Q	WW	0	0	1	0	1	1	0	0	1	0	1	0	5	
	LW	0	0	0	0	0	0	0	0	0	1	0	0	1	
	WL	0	0	0	0	0	0	0	0	0	0	0	1	1	
	LL	1	1	0	1	0	0	1	1	0	0	0	0	5	12
1995 4Q	WW	0	0	0	0	1	1	0	0	0	1	0	0	3	
	LW	0	0	0	0	0	0	1	1	0	0	0	1	3	
	WL	0	0	1	0	0	0	0	0	1	0	1	0	3	
	LL	1	1	0	1	0	0	0	0	0	0	0	0	3	12
1996 1Q	WW	0	0	0	0	1	0	1	1	0	0	0	0	3	
	LW	0	1	1	0	0	0	0	0	1	0	0	0	3	
	WL	0	0	0	0	0	1	0	0	0	1	0	1	3	
	LL	1	0	0	1	0	0	0	0	0	0	1	0	3	12
1996 2Q	WW	0	0	1	0	1	0	1	1	1	0	0	0	5	
	LW	0	0	0	0	0	0	0	0	0	1	0	0	1	
	WL	0	1	0	0	0	0	0	0	0	0	0	0	1	
	LL	1	0	0	1	0	1	0	0	0	0	1	1	5	12
1996 3Q	WW	0	0	1	0	1	0	1	1	1	1	0	0	6	
	LW	0	0	0	0	0	0	0	0	0	0	0	0	0	
	WL	0	0	0	0	0	0	0	0	0	0	0	0	0	
	LL	1	1	0	1	0	1	0	0	0	0	1	1	6	12
1996 4Q	WW	0	0	0	0	1	0	0	1	1	1	0	0	4	
	LW	1	0	0	0	0	1	0	0	0	0	0	0	2	
	WL	0	0	1	0	0	0	1	0	0	0	0	0	2	
	LL	0	1	0	1	0	0	0	0	0	0	1	1	4	12
1997 1Q	WW	1	0	0	0	1	1	0	0	1	1	0	0	5	
	LW	0	0	1	0	0	0	0	0	0	0	0	0	1	
	WL	0	0	0	0	0	0	0	1	0	0	0	0	1	
	LL	0	1	0	1	0	0	1	0	0	0	1	1	5	12
1997 2Q	WW	0	0	1	0	1	1	0	0	1	0	0	0	4	
	LW	0	0	0	0	0	0	0	0	0	0	1	1	2	
	WL	1	0	0	0	0	0	0	0	0	1	0	0	2	
	LL	0	1	0	1	0	0	1	1	0	0	0	0	4	12
1997 3Q	WW	0	0	1	0	1	0	0	0	1	0	0	1	4	
	LW	0	0	0	0	0	0	0	1	0	1	0	0	2	
	WL	0	0	0	0	0	1	0	0	0	0	1	0	2	
	LL	1	1	0	1	0	0	1	0	0	0	0	0	4	12
1997 4Q	WW	0	0	1	0	1	0	0	1	1	0	0	0	4	
	LW	0	0	0	0	0	1	0	0	0	0	1	0	2	
	WL	0	0	0	0	0	0	0	0	0	1	0	1	2	
	LL	1	1	0	1	0	0	1	0	0	0	0	0	4	12
1998 1Q	WW	0	0	1	0	1	0	0	0	0	0	1	0	3	
	LW	0	0	0	1	0	0	1	0	0	0	0	1	3	
	WL	0	0	0	0	0	1	0	1	1	0	0	0	3	
	LL	1	1	0	0	0	0	0	0	0	1	0	0	3	12
CONSISTENCY OF INDIVIDUAL FUND PERFORMANCE															
1994-1997	WW	1	0	11	0	15	6	3	5	12	4	4	1	62	
	LW	2	1	2	1	0	3	3	3	1	4	3	5	28	
	WL	2	1	2	0	0	3	2	4	2	5	3	4	28	
	LL	10	13	0	14	0	3	7	3	0	2	5	5	62	180

Appendix 4.10

Frequency of repeat winners and losers: beta

	A	B	C	D	F	G	H	I	J	K	L	M	MEDIAN
2Q94	0.31864	0.36323	0.42456	0.14815	0.56535	0.35010	0.34421	0.41572	0.69543	0.47677	0.47908	0.15471	0.38947
3Q94	0.08986	0.17918	0.32399	0.16361	0.52891	0.13499	0.24185	0.23517	0.30099	0.14101	0.35701	0.20017	0.21767
4Q94	0.47134	0.26091	0.36438	0.13496	0.61844	0.27740	0.18925	0.31093	0.39829	0.19063	0.30886	0.16249	0.29313
1Q95	0.40882	0.45399	0.38154	0.26567	0.75929	0.45082	0.19246	0.46370	0.56593	0.37739	0.49249	0.14852	0.42982
2Q95	0.47691	0.41502	0.42595	0.45692	0.49229	0.47592	0.14066	0.32920	0.60855	0.10403	0.58826	0.15406	0.44144
3Q95	0.54209	0.52468	0.72218	0.46213	0.68758	0.37433	0.42384	0.48213	0.52724	0.52345	0.79397	0.36469	0.52407
4Q95	0.49798	0.47824	0.60907	0.51689	0.71266	0.63156	0.36618	0.83767	0.64563	0.30664	0.62929	0.32371	0.56298
1Q96	0.56667	0.69588	0.71171	0.63891	0.77039	0.64823	0.50408	0.75928	0.70011	0.43569	0.63366	0.47010	0.64357
2Q96	0.30697	0.60492	0.58321	0.44353	0.58737	0.61175	0.58017	0.76171	0.65484	0.34131	0.55568	0.51099	0.58169
3Q96	0:38968	0.56421	0.62954	0.37145	0.65634	0.53006	0.57146	0.70425	0.69143	0.46420	0.50279	0.49085	0.54713
4Q96	0.79783	0.70405	0.69824	0.69308	0.91408	0.89435	0.70016	0.76223	0.97527	0.73011	0.83710	0.59910	0.74617
1Q97	0.78691	0.68877	0.74569	0.58873	0.89308	0.76359	0.73494	0.71341	0.82499	0.72320	0.74732	0.73299	0.74031
2Q97	0.73785	0.78231	0.81549	0.74080	0.83561	0.79698	0.66584	0.79455	0.88445	0.60236	0.87457	0.69086	0.78843
3Q97	0.77607	0.71383	0.88214	0.70730	0.62190	0.80314	0.71362	0.87519	0.82571	0.91033	0.76885	0.92245	0.78961
4Q97	0.68919	0.83756	1.46724	0.81376	0.93221	0.87483	0.82639	0.91668	0.90224	0.81269	0.88721	0.81571	0.85620
1Q98	0.83030	0.80199	0.93292	0.87203	0.86136	0.81669	0.83762	0.80729	0.81380	0.74680	0.85684	0.86612	0.83396
TOTAL	0.61485	0.64913	0.85180	0.59564	0.76325	0.67521	0.59399	0.71480	0.77055	0.64629	0.71936	0.61609	0.66217

REPEATS		A	B	C	D	F	G	H	I	J	K	L	M	SUM	TOTAL
1994 3Q	WW	0	0	1	0	1	0	0	1	1	0	1	0	5	
	LW	0	0	0	0	0	0	1	0	0	0	0	0	1	
	WL	0	0	0	0	0	0	0	0	0	1	0	0	1	
	LL	1	1	0	1	0	1	0	0	0	0	0	1	5	12
1994 4Q	WW	0	0	1	0	1	0	0	1	1	0	1	0	5	
	LW	1	0	0	0	0	0	0	0	0	0	0	0	1	
	WL	0	0	0	0	0	0	1	0	0	0	0	0	1	
	LL	0	1	0	1	0	1	0	0	0	1	0	1	5	12
1995 1Q	WW	0	0	0	0	1	0	0	1	1	0	1	0	4	
	LW	0	1	0	0	0	1	0	0	0	0	0	0	2	
	WL	1	0	1	0	0	0	0	0	0	0	0	0	2	
	LL	0	0	0	1	0	0	1	0	0	1	0	1	4	12
1995 2Q	WW	0	0	0	0	1	1	0	0	1	0	1	0	4	
	LW	1	0	0	1	0	0	0	0	0	0	0	0	2	
	WL	0	1	0	0	0	0	0	1	0	0	0	0	2	
	LL	0	0	1	0	0	0	1	0	0	1	0	1	4	12
1995 3Q	WW	1	0	0	0	1	0	0	0	1	0	1	0	4	
	LW	0	1	1	0	0	0	0	0	0	0	0	0	2	
	WL	0	0	0	1	0	1	0	0	0	0	0	0	2	
	LL	0	0	0	0	0	0	1	1	0	1	0	1	4	12
1995 4Q	WW	0	0	1	0	1	0	0	0	1	0	1	0	4	
	LW	0	0	0	0	0	1	0	1	0	0	0	0	2	
	WL	1	1	0	0	0	0	0	0	0	0	0	0	2	
	LL	0	0	0	1	0	0	1	0	0	1	0	1	4	12
1996 1Q	WW	0	0	1	0	1	1	0	1	1	0	0	0	5	
	LW	0	1	0	0	0	0	0	0	0	0	0	0	1	
	WL	0	0	0	0	0	0	0	0	0	0	1	0	1	
	LL	1	0	0	1	0	0	1	0	0	1	0	1	5	12
1996 2Q	WW	0	1	1	0	1	1	0	1	1	0	0	0	6	
	LW	0	0	0	0	0	0	0	0	0	0	0	0	0	
	WL	0	0	0	0	0	0	0	0	0	0	0	0	0	
	LL	1	0	0	1	0	0	1	0	0	1	1	1	6	12
1996 3Q	WW	0	1	1	0	1	0	0	1	1	0	0	0	5	
	LW	0	0	0	0	0	0	1	0	0	0	0	0	1	
	WL	0	0	0	0	0	1	0	0	0	0	0	0	1	
	LL	1	0	0	1	0	0	0	0	0	1	1	1	5	12
1996 4Q	WW	0	0	0	0	1	0	0	1	1	0	0	0	3	
	LW	1	0	0	0	0	1	0	0	0	0	1	0	3	
	WL	0	1	1	0	0	0	1	0	0	0	0	0	3	
	LL	0	0	0	1	0	0	0	0	0	1	0	1	3	12
1997 1Q	WW	1	0	0	0	1	1	0	0	1	0	1	0	5	
	LW	0	0	1	0	0	0	0	0	0	0	0	0	1	
	WL	0	0	0	0	0	0	0	1	0	0	0	0	1	
	LL	0	1	0	1	0	0	1	0	0	1	0	1	5	12
1997 2Q	WW	0	0	1	0	1	1	0	0	1	0	1	0	5	
	LW	0	0	0	0	0	0	0	1	0	0	0	0	1	
	WL	1	0	0	0	0	0	0	0	0	0	0	0	1	
	LL	0	1	0	1	0	0	1	0	0	1	0	1	5	12
1997 3Q	WW	0	0	1	0	0	1	0	1	1	0	0	0	4	
	LW	0	0	0	0	0	0	0	0	0	1	0	1	2	
	WL	0	0	0	0	1	0	0	0	0	0	1	0	2	
	LL	1	1	0	1	0	0	1	0	0	0	0	0	4	12
1997 4Q	WW	0	0	1	0	0	1	0	1	1	0	0	0	4	
	LW	0	0	0	0	1	0	0	0	0	0	1	0	2	
	WL	0	0	0	0	0	0	0	0	0	1	0	1	2	
	LL	1	1	0	1	0	0	1	0	0	0	0	0	4	12
1998 1Q	WW	0	0	1	0	1	0	0	0	0	0	1	0	3	
	LW	0	0	0	1	0	0	1	0	0	0	0	1	3	
	WL	0	0	0	0	0	1	0	1	1	0	0	0	3	
	LL	1	1	0	0	0	0	0	0	0	1	0	0	3	12
CONSISTENCY OF INDIVIDUAL FUND PERFORMANCE															
1994-1997	WW	2	2	10	0	13	7	0	9	14	0	9	0	66	
	LW	3	3	2	2	1	3	3	2	0	1	2	2	24	
	WL	3	3	2	1	1	3	2	3	1	2	2	1	24	
	LL	7	7	1	12	0	2	10	1	0	12	2	12	66	180

Appendix 4.11

Frequency of repeat winners and losers based on excess returns(two year and one year periods)

		A	B	C	D	F	G	H	I	J	K	L	M	MEDIAN	
Apr94-Mar96		-0.00007	-0.00031	-0.00016	-0.00010	-0.00043	-0.00011	-0.00011	-0.00007	-0.00049	-0.00055	-0.00016	-0.00014	-0.00015	
Apr96-Mar98		0.00205	0.00162	0.00174	0.00164	0.00188	0.00180	0.00162	0.00184	0.00183	0.00185	0.00177	0.00185	0.00181	
TOTAL		0.00099	0.00065	0.00079	0.00077	0.00072	0.00084	0.00076	0.00089	0.00067	0.00065	0.00081	0.00085	0.00083	
REPEATS														SUM	TOTAL
	WW	1	0	0	0	0	0	0	1	0	0	0	1	3	
	LW	0	0	0	0	1	0	0	0	1	1	0	0	3	
	WL	0	0	0	1	0	1	1	0	0	0	0	0	3	
	LL	0	1	1	0	0	0	0	0	0	0	1	0	3	12

		A	B	C	D	F	G	H	I	J	K	L	M	MEDIAN	
Apr94-Mar95		-0.00061	-0.00056	-0.00046	-0.00028	-0.00066	-0.00040	-0.00022	-0.00037	-0.00064	-0.00056	-0.00047	-0.00016	-0.00047	
Apr95-Mar96		0.00048	-0.00005	0.00015	0.00008	-0.00021	0.00017	0.00001	0.00024	-0.00035	-0.00054	0.00016	-0.00013	0.00004	
Apr96-Mar97		0.00157	0.00091	0.00104	0.00107	0.00121	0.00104	0.00120	0.00104	0.00116	0.00155	0.00099	0.00095	0.00105	
Apr97-Mar98		0.00252	0.00231	0.00244	0.00220	0.00254	0.00255	0.00203	0.00263	0.00248	0.00214	0.00255	0.00274	0.00250	
TOTAL		0.00099	0.00065	0.00079	0.00077	0.00072	0.00084	0.00075	0.00089	0.00067	0.00065	0.00081	0.00085	0.00078	
REPEATS														SUM	TOTAL
2nd year	WW	0	0	1	1	0	1	0	1	0	0	0	0	4	
	LW	1	0	0	0	0	0	0	0	0	0	1	0	2	
	WL	0	0	0	0	0	0	1	0	0	0	0	1	2	
	LL	0	1	0	0	1	0	0	0	1	1	0	0	4	12
3rd year	WW	1	0	0	1	0	0	0	0	0	0	0	0	2	
	LW	0	0	0	0	1	0	1	0	1	1	0	0	4	
	WL	0	0	1	0	0	1	0	1	0	0	1	0	4	
	LL	0	1	0	0	0	0	0	0	0	0	0	1	2	12
4th year	WW	1	0	0	0	1	0	0	0	0	0	0	0	2	
	LW	0	0	0	0	0	1	0	1	0	0	1	1	4	
	WL	0	0	0	1	0	0	1	0	1	1	0	0	4	
	LL	0	1	1	0	0	0	0	0	0	0	0	0	2	12
					CONSISTENCY OF INDIVIDUAL FUND PERFORMANCE										
1994-1997	WW	2	0	1	2	1	1	0	1	0	0	0	0	8	
	LW	1	0	0	0	1	1	1	1	1	1	2	1	10	
	WL	0	0	1	1	0	1	2	1	1	1	1	1	10	
	LL	0	3	1	0	1	0	0	0	1	1	0	1	8	36

Appendix 4.12

Frequency of Repeat winners and losers based on excess returns

(Half year periods)

		A	B	C	D	F	G	H	I	J	K	L	M	MEDIAN	
A94-S94		-0.00109	-0.00078	-0.00069	-0.00029	-0.00077	-0.00089	-0.00028	-0.00066	-0.00086	-0.00056	-0.00052	-0.00019	-0.00068	
O94-M95		-0.00013	-0.00035	-0.00023	-0.00028	-0.00054	0.00009	-0.00017	-0.00009	-0.00041	-0.00057	-0.00042	-0.00013	-0.00025	
A95-S95		-0.00010	-0.00035	-0.00027	-0.00017	-0.00046	-0.00033	-0.00037	-0.00011	-0.00046	-0.00094	-0.00004	-0.00047	-0.00034	
O95-M96		0.00106	0.00025	0.00057	0.00034	0.00004	0.00068	0.00039	0.00059	-0.00023	-0.00015	0.00036	0.00022	0.00035	
A96-S96		0.00081	0.00041	0.00061	0.00073	0.00050	0.00042	0.00076	0.00036	0.00067	0.00110	0.00043	0.00045	0.00055	
O96-M97		0.00234	0.00142	0.00147	0.00141	0.00193	0.00167	0.00164	0.00173	0.00166	0.00201	0.00155	0.00145	0.00165	
A97-S97		0.00215	0.00180	0.00206	0.00164	0.00218	0.00230	0.00172	0.00232	0.00223	0.00218	0.00225	0.00217	0.00217	
O97-M98		0.00290	0.00282	0.00283	0.00276	0.00291	0.00281	0.00236	0.00295	0.00274	0.00210	0.00286	0.00332	0.00283	
TOTAL		0.00099	0.00065	0.00079	0.00077	0.00072	0.00084	0.00076	0.00089	0.00067	0.00065	0.00081	0.00085	0.00078	
														SUM	TOTAL
REPEATS															
O94-M95	WW	0	0	0	0	0	0	1	1	0	0	0	1	3	
	LW	1	0	1	0	0	1	0	0	0	0	0	0	3	
	WL	0	0	0	1	0	0	0	0	0	1	1	0	3	
	LL	0	1	0	0	1	0	0	0	1	0	0	0	3	12
A95-S95	WW	1	0	1	0	0	1	0	1	0	0	0	0	4	
	LW	0	0	0	1	0	0	0	0	0	0	1	0	2	
	WL	0	0	0	0	0	0	1	0	0	0	0	1	2	
	LL	0	1	0	0	1	0	0	0	1	1	0	0	4	12
O95-M96	WW	1	0	1	0	0	1	0	1	0	0	1	0	5	
	LW	0	0	0	0	0	0	1	0	0	0	0	0	1	
	WL	0	0	0	1	0	0	0	0	0	0	0	0	1	
	LL	0	1	0	0	1	0	0	0	1	1	0	1	5	12
A96-S96	WW	1	0	1	0	0	0	1	0	0	0	0	0	3	
	LW	0	0	0	1	0	0	0	0	1	1	0	0	3	
	WL	0	0	0	0	0	1	0	1	0	0	1	0	3	
	LL	0	1	0	0	1	0	0	0	0	0	0	1	3	12
O96-M97	WW	1	0	0	0	0	0	0	0	1	1	0	0	3	
	LW	0	0	0	0	1	1	0	1	0	0	0	0	3	
	WL	0	0	1	1	0	0	1	0	0	0	0	0	3	
	LL	0	1	0	0	0	0	0	0	0	0	1	1	3	12
A97-S97	WW	0	0	0	0	1	1	0	1	1	1	0	0	5	
	LW	0	0	0	0	0	0	0	0	0	0	1	0	1	
	WL	1	0	0	0	0	0	0	0	0	0	0	0	1	
	LL	0	1	1	1	0	0	1	0	0	0	0	1	5	12
O97-M98	WW	0	0	0	0	1	0	0	1	0	0	1	0	3	
	LW	1	0	1	0	0	0	0	0	0	0	0	1	3	
	WL	0	0	0	0	0	1	0	0	1	1	0	0	3	
	LL	0	1	0	1	0	0	1	0	0	0	0	0	3	12
						CONSISTENCY OF INDIVIDUAL FUND PERFORMANCE									
1994-1997	WW	4	0	3	0	2	3	2	5	2	2	2	1	26	
	LW	2	0	2	2	1	2	1	1	1	1	2	1	16	
	WL	1	0	1	3	0	2	2	1	1	2	2	1	16	
	LL	0	7	1	2	4	0	2	0	3	2	1	4	26	84

Appendix 4.13

Contingency table of excess returns: monthly periods

	WW	LW	WL	LL	PERCENT REPEAT W	MALKIEL Z-TEST W	p-value	ODDS RATIO	B&G Z-STAT		σ (LOG)	p-value
may-94	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
jun-94	5	1	1	5	0.83	1.63	0.102	25.00	2.08	*	1.55	0.038
jul-94	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13		1.22	0.258
aug-94	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
sept-94	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
oct-94	4	2	2	4	0.67	0.82	0.414	4.00	1.13		1.22	0.258
nov-94	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
dez-94	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
jan-95	5	1	1	5	0.83	1.63	0.102	25.00	2.08	*	1.55	0.038
fev-95	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13		1.22	0.258
mar-95	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13		1.22	0.258
apr-95	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
may-95	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
jun-95	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13		1.22	0.258
jul-95	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13		1.22	0.258
aug-95	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
sept-95	4	2	2	4	0.67	0.82	0.414	4.00	1.13		1.22	0.258
oct-95	4	2	2	4	0.67	0.82	0.414	4.00	1.13		1.22	0.258
nov-95	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
dez-95	5	1	1	5	0.83	1.63	0.102	25.00	2.08	*	1.55	0.038
jan-96	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
fev-96	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13		1.22	0.258
mar-96	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
apr-96	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
may-96	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13		1.22	0.258
jun-96	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
jul-96	4	2	2	4	0.67	0.82	0.414	4.00	1.13		1.22	0.258
aug-96	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
sept-96	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
oct-96	4	2	2	4	0.67	0.82	0.414	4.00	1.13		1.22	0.258
nov-96	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
dez-96	4	2	2	4	0.67	0.82	0.414	4.00	1.13		1.22	0.258
jan-97	5	1	1	5	0.83	1.63	0.102	25.00	2.08	*	1.55	0.038
feb-97	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13		1.22	0.258
mar-97	4	2	2	4	0.67	0.82	0.414	4.00	1.13		1.22	0.258
apr-97	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
may-97	4	2	2	4	0.67	0.82	0.414	4.00	1.13		1.22	0.258
jun-97	4	2	2	4	0.67	0.82	0.414	4.00	1.13		1.22	0.258
jul-97	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13		1.22	0.258
aug-97	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
sept-97	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13		1.22	0.258
oct-97	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
nov-97	4	2	2	4	0.67	0.82	0.414	4.00	1.13		1.22	0.258
dec-97	4	2	2	4	0.67	0.82	0.414	4.00	1.13		1.22	0.258
jan-98	3	3	3	3	0.50	0.00	1.000	1.00	0.00		1.15	1.000
feb-98	1	5	5	1	0.17	-1.63	0.102	0.04	-2.08	*	1.55	0.038
mar-98	2	4	4	2	0.33	-0.82	0.414	0.25	-1.13		1.22	0.258
TOTAL	147	135	135	147	0.52	0.71	0.475	1.19	1.01		0.17	0.312

(Cont.)

Appendix 4.13

(Continued)

	WW	LW	WL	LL	CHI-SQ		p-value	FISHER EXACT P-VALUE
may-94	3	3	3	3	0.000		1.000	1.000
jun-94	5	1	1	5	5.333	*	0.021	0.080
jul-94	2	4	4	2	1.333		0.248	0.567
aug-94	3	3	3	3	0.000		1.000	1.000
sept-94	3	3	3	3	0.000		1.000	1.000
oct-94	4	2	2	4	1.333		0.248	0.567
nov-94	3	3	3	3	0.000		1.000	1.000
dez-94	3	3	3	3	0.000		1.000	1.000
jan-95	5	1	1	5	5.333	*	0.021	0.080
fev-95	2	4	4	2	1.333		0.248	0.567
mar-95	2	4	4	2	1.333		0.248	0.567
apr-95	3	3	3	3	0.000		1.000	1.000
may-95	3	3	3	3	0.000		1.000	1.000
jun-95	2	4	4	2	1.333		0.248	0.567
jul-95	2	4	4	2	1.333		0.248	0.567
aug-95	3	3	3	3	0.000		1.000	1.000
sept-95	4	2	2	4	1.333		0.248	0.567
oct-95	4	2	2	4	1.333		0.248	0.567
nov-95	3	3	3	3	0.000		1.000	1.000
dez-95	5	1	1	5	5.333	*	0.021	0.080
jan-96	3	3	3	3	0.000		1.000	1.000
fev-96	2	4	4	2	1.333		0.248	0.567
mar-96	3	3	3	3	0.000		1.000	1.000
apr-96	3	3	3	3	0.000		1.000	1.000
may-96	2	4	4	2	1.333		0.248	0.567
jun-96	3	3	3	3	0.000		1.000	1.000
jul-96	4	2	2	4	1.333		0.248	0.567
aug-96	3	3	3	3	0.000		1.000	1.000
sept-96	3	3	3	3	0.000		1.000	1.000
oct-96	4	2	2	4	1.333		0.248	0.567
nov-96	3	3	3	3	0.000		1.000	1.000
dez-96	4	2	2	4	1.333		0.248	0.567
jan-97	5	1	1	5	5.333	*	0.021	0.080
feb-97	2	4	4	2	1.333		0.248	0.567
mar-97	4	2	2	4	1.333		0.248	0.567
apr-97	3	3	3	3	0.000		1.000	1.000
may-97	4	2	2	4	1.333		0.248	0.567
jun-97	4	2	2	4	1.333		0.248	0.567
jul-97	2	4	4	2	1.333		0.248	0.567
aug-97	3	3	3	3	0.000		1.000	1.000
sept-97	2	4	4	2	1.333		0.248	0.567
oct-97	3	3	3	3	0.000		1.000	1.000
nov-97	4	2	2	4	1.333		0.248	0.567
dec-97	4	2	2	4	1.333		0.248	0.567
jan-98	3	3	3	3	0.000		1.000	1.000
feb-98	1	5	5	1	5.333	*	0.021	0.080
mar-98	2	4	4	2	1.333		0.248	0.567
TOTAL	147	135	135	147	1.021		0.312	

* Statistically significant at the 5 percent level

Appendix 4.14

Two-Year estimates of the regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon_{p,t}$ relative to the PSI-20 index and frequency of repeat winners and losers

Fund		A	B	C	D	F	G	H	I	J	K	L	M
Apr94-Mar96	α	0.00006	-0.00015	0.00005	0.00000	-0.00020	0.00004	0.00000	0.00010	-0.00027	-0.00041	0.00005	-0.00004
		(0.40)	(-1.09)	(0.34)	(0.03)	(-1.87)	(0.25)	(-0.02)	(0.93)	(-1.64)	(-2.48)	(0.44)	(-0.28)
	β	0.358	0.392	0.466	0.270	0.615	0.383	0.250	0.451	0.595	0.387	0.498	0.224
	R^2	0.461	0.522	0.423	0.406	0.771	0.310	0.157	0.546	0.571	0.386	0.637	0.106
Apr96-Mar98	α	0.00048	-0.00004	-0.00051	0.00002	0.00011	0.00002	-0.00003	0.00004	-0.00001	0.00021	0.00000	0.00011
		(3.84)	(-0.39)	(-1.46)	(0.17)	(0.53)	(0.19)	(-0.20)	(0.35)	(-0.07)	(1.12)	(-0.00)	(0.57)
	β	0.737	0.776	1.052	0.758	0.833	0.822	0.766	0.847	0.855	0.767	0.830	0.810
	R^2	0.880	0.929	0.761	0.903	0.632	0.903	0.867	0.937	0.919	0.822	0.923	0.543

		A	*B*	*C*	*D*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*MEDIAN*	
Apr94-Mar96		0.00006	-0.00015	0.00005	0.00000	-0.00020	0.00004	0.00000	0.00010	-0.00027	-0.00041	0.00005	-0.00004	0.00000	
Apr96-Mar98		0.00048	-0.00004	-0.00051	0.00002	0.00011	0.00002	-0.00003	0.00004	-0.00001	0.00021	0.00000	0.00011	0.00002	
TOTAL		0.00027	-0.00009	-0.00023	0.00001	-0.00005	0.00003	-0.00002	0.00007	-0.00014	-0.00010	0.00002	0.00003	0.00001	
REPEATS														SUM	TOTAL
	WW	1	0	0	0	0	1	0	1	0	0	0	0	3	
	LW	0	0	0	0	1	0	0	0	0	1	0	1	3	
	WL	0	0	1	1	0	0	0	0	0	0	1	0	3	
	LL	0	1	0	0	0	0	1	0	1	0	0	0	3	12

Appendix 4.15

Frequency of repeat winners and losers based on alpha (one-year period)

A shadowed area for the estimated annual alpha represents a value statistically significant at the 5 percent level. If is outlined, it is statistically significant at the 1 percent level. The statistical significance is based on the NEWEY and WEST [1987] t-statistics.

		A	***B***	***C***	***D***	***F***	***G***	***H***	***I***	***J***	***K***	***L***	***M***	***MEDIAN***	
Apr94-Mar95		-0.00034	-0.00023	-0.00005	-0.00009	-0.00010	-0.00008	-0.00001	0.00001	-0.00012	-0.00020	0.00001	0.00006	-0.00009	
Apr95-Mar96		0.00038	-0.00016	0.00006	-0.00003	-0.00035	0.00007	-0.00007	0.00009	-0.00045	-0.00062	0.00000	-0.00022	-0.00005	
Apr96-Mar97		0.00070	0.00005	0.00010	0.00033	0.00015	0.00007	0.00031	0.00013	0.00013	0.00072	0.00010	0.00010	0.00013	
Apr97-Mar98		0.00031	-0.00004	-0.00086	-0.00013	0.00007	0.00005	-0.00031	0.00003	-0.00011	-0.00020	0.00000	0.00025	-0.00002	
TOTAL		0.00045	0.00009	0.00005	0.00024	0.00005	0.00024	0.00023	0.00026	-0.00002	0.00007	0.00018	0.00031	0.00021	
REPEATS														SUM	TOTAL
2nd year	WW	0	0	1	0	0	1	0	1	0	0	1	0	4	
	LW	1	0	0	1	0	0	0	0	0	0	0	0	2	
	WL	0	0	0	0	0	0	1	0	0	0	0	1	2	
	LL	0	1	0	0	1	0	0	0	1	1	0	0	4	12
3rd year	WW	1	0	0	1	0	0	0	0	0	0	0	0	2	
	LW	0	0	0	0	1	0	1	0	1	1	0	0	4	
	WL	0	0	1	0	0	1	0	1	0	0	1	0	4	
	LL	0	1	0	0	0	0	0	0	0	0	0	1	2	12
4th year	WW	1	0	0	0	1	0	0	0	0	0	0	0	2	
	LW	0	0	0	0	0	1	0	1	0	0	1	1	4	
	WL	0	0	0	1	0	0	1	0	1	1	0	0	4	
	LL	0	1	1	0	0	0	0	0	0	0	0	0	2	12
CONSISTENCY OF INDIVIDUAL FUND PERFORMANCE															
1994-1997	WW	2	0	1	1	1	1	0	1	0	0	1	0	8	
	LW	1	0	0	1	1	1	1	1	1	1	1	1	10	
	WL	0	0	1	1	0	1	2	1	1	1	1	1	10	
	LL	0	3	1	0	1	0	0	0	1	1	0	1	8	36

Appendix 4.16

Half-year estimates of the regression $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon_{p,t}$ relative to the PSI-20 index

A shadowed area for the estimated alpha represents a value statistically significant at the 5 percent level. If is outlined, it is statistically significant at the 1 percent level. The statistical significance is based on the NEWEY and WEST [1987] t-statistics.

		A	B	C	D	F	G	H	I	J	K	L	M
A94-S94	α	-0.00075	-0.00029	-0.00007	-0.00004	-0.00002	-0.00039	0.00001	-0.00013	-0.00010	-0.00007	0.00015	0.00010
	β	0.26	0.32	0.41	0.16	0.56	0.29	0.20	0.37	0.59	0.37	0.45	0.18
	R^2	42.2%	55.8%	43.9%	38.3%	91.0%	47.7%	27.0%	59.4%	63.4%	66.7%	71.9%	32.0%
O94-M95	α	0.00009	-0.00014	-0.00003	-0.00013	-0.00015	0.00024	-0.00003	0.00015	-0.00014	-0.00032	-0.00012	0.00001
	β	0.42	0.41	0.38	0.24	0.73	0.41	0.19	0.43	0.53	0.47	0.46	0.15
	R^2	46.4%	55.7%	26.3%	30.3%	77.3%	22.7%	19.4%	44.3%	38.0%	39.1%	42.8%	11.2%
A95-S95	α	0.00007	-0.00021	-0.00005	-0.00006	-0.00028	-0.00017	-0.00031	-0.00001	-0.00027	-0.00079	0.00011	-0.00044
	β	0.51	0.47	0.57	0.47	0.58	0.44	0.28	0.40	0.57	0.31	0.68	0.26
	R^2	51.2%	36.9%	30.6%	53.5%	50.0%	21.2%	12.4%	43.2%	43.1%	9.4%	57.4%	8.6%
O95-M96	α	0.00068	-0.00017	0.00012	-0.00006	-0.00048	0.00024	0.00010	0.00004	-0.00069	-0.00050	-0.00009	-0.00006
	β	0.55	0.61	0.69	0.60	0.76	0.64	0.46	0.79	0.70	0.42	0.64	0.43
	R^2												
A96-S96	α	0.00057	-0.00003	0.00017	0.00042	0.00003	-0.00002	0.00030	-0.00018	0.00015	0.00078	0.00003	0.00004
	β	0.35	0.59	0.61	0.41	0.63	0.58	0.58	0.74	0.67	0.39	0.53	0.51
	R^2	35.5%	70.3%	50.3%	46.7%	63.6%	67.5%	59.9%	83.7%	70.9%	21.5%	72.8%	51.6%
O96-M97	α	0.00085	0.00013	0.00003	0.00024	0.00026	0.00016	0.00033	0.00046	0.00012	0.00066	0.00017	0.00017
	β	0.79	0.69	0.74	0.61	0.90	0.79	0.73	0.72	0.85	0.72	0.76	0.71
	R^2	76.0%	87.8%	75.8%	74.6%	88.2%	77.0%	80.6%	78.2%	81.4%	69.3%	80.7%	79.3%
A97-S97	α	0.00022	-0.00006	-0.00011	-0.00015	0.00039	0.00021	-0.00010	0.00011	0.00002	0.00019	0.00014	0.00002
	β	0.76	0.74	0.86	0.72	0.70	0.80	0.70	0.84	0.85	0.79	0.81	0.83
	R^2	94.4%	94.4%	81.9%	93.9%	31.5%	91.3%	80.1%	96.0%	92.9%	82.7%	92.6%	31.3%
O97-M98		0.00039	0.00001	-0.00150	-0.00007	-0.00019	-0.00010	-0.00048	-0.00004	-0.00023	-0.00060	-0.00012	0.00049
	β	0.74	0.83	1.29	0.84	0.92	0.86	0.84	0.88	0.88	0.80	0.88	0.84
	R^2	92.3%	95.3%	80.9%	95.8%	95.5%	94.7%	94.6%	96.4%	95.5%	92.8%	96.5%	93.4%

Appendix 4.17

Frequency of repeat winners and losers based on alpha

(Half year periods)

A shadowed area for the estimated half-year alpha represents a value statistically significant at the 5 percent level. If is outlined, it is statistically significant at the 1 percent level. The statistical significance is based on the NEWEY and WEST [1987] t-statistics.

		A	B	C	D	F	G	H	I	J	K	L	M	MEDIAN	
A94-S94		-0.00075	-0.00029	-0.00007	-0.00004	-0.00002	-0.00039	0.00001	-0.00013	-0.00010	-0.00007	0.00015	0.00010	-0.00007	
O94-M95		0.00009	-0.00014	-0.00003	-0.00013	-0.00015	0.00024	-0.00003	0.00015	-0.00014	-0.00032	-0.00012	0.00001	-0.00008	
A95-S95		0.00007	-0.00021	-0.00005	-0.00006	-0.00028	-0.00017	-0.00031	-0.00001	-0.00027	-0.00079	0.00011	-0.00044	-0.00019	
O95-M96		0.00068	-0.00017	0.00012	-0.00006	-0.00048	0.00024	0.00010	0.00004	-0.00069	-0.00050	-0.00009	-0.00006	-0.00006	
A96-S96		0.00057	-0.00003	0.00017	0.00042	0.00003	-0.00002	0.00030	-0.00018	0.00015	0.00078	0.00003	0.00004	0.00009	
O96-M97		0.00085	0.00013	0.00003	0.00024	0.00026	0.00016	0.00033	0.00046	0.00012	0.00066	0.00017	0.00017	0.00021	
A97-S97		0.00022	-0.00006	-0.00011	-0.00015	0.00039	0.00021	-0.00010	0.00011	0.00002	0.00019	0.00014	0.00002	0.00007	
O97-M98		0.00039	0.00001	-0.00150	-0.00007	-0.00019	-0.00010	-0.00048	-0.00004	-0.00023	-0.00060	-0.00012	0.00049	-0.00011	
TOTAL		0.00045	0.00009	0.00005	0.00024	0.00005	0.00024	0.00023	0.00026	-0.00002	0.00007	0.00018	0.00031	0.00021	
REPEATS														SUM	TOTAL
O94-M95	WW	0	0	0	0	0	0	1	0	0	0	0	1	2	
	LW	1	0	1	0	0	1	0	1	0	0	0	0	4	
	WL	0	0	0	1	1	0	0	0	0	1	1	0	4	
	LL	0	1	0	0	0	0	0	0	1	0	0	0	2	12
A95-S95	WW	1	0	1	0	0	1	0	1	0	0	0	0	4	
	LW	0	0	0	1	0	0	0	0	0	0	1	0	2	
	WL	0	0	0	0	0	0	1	0	0	0	0	1	2	
	LL	0	1	0	0	1	0	0	0	1	1	0	0	4	12
O95-M96	WW	1	0	1	0	0	1	0	1	0	0	0	0	4	
	LW	0	0	0	0	0	0	1	0	0	0	0	1	2	
	WL	0	0	0	1	0	0	0	0	0	0	1	0	2	
	LL	0	1	0	0	1	0	0	0	1	1	0	0	4	12
A96-S96	WW	1	0	1	0	0	0	1	0	0	0	0	0	3	
	LW	0	0	0	1	0	0	0	0	1	1	0	0	3	
	WL	0	0	0	0	0	1	0	1	0	0	0	1	3	
	LL	0	1	0	0	1	0	0	0	0	0	1	0	3	12
O96-M97	WW	1	0	0	1	0	0	1	0	0	1	0	0	4	
	LW	0	0	0	0	1	0	0	1	0	0	0	0	2	
	WL	0	0	1	0	0	0	0	0	1	0	0	0	2	
	LL	0	1	0	0	0	1	0	0	0	0	1	1	4	12
A97-S97	WW	1	0	0	0	1	0	0	1	0	1	0	0	4	
	LW	0	0	0	0	0	1	0	0	0	0	1	0	2	
	WL	0	0	0	1	0	0	1	0	0	0	0	0	2	
	LL	0	1	1	0	0	0	0	0	1	0	0	1	4	12
O97-M98	WW	1	0	0	0	0	1	0	1	0	0	0	0	3	
	LW	0	1	0	1	0	0	0	0	0	0	0	1	3	
	WL	0	0	0	0	1	0	0	0	0	1	1	0	3	
	LL	0	0	1	0	0	0	1	0	1	0	0	0	3	12
CONSISTENCY OF INDIVIDUAL FUND PERFORMANCE															
1994-1997	WW	6	0	3	1	1	3	3	4	0	2	0	1	24	
	LW	1	1	1	3	1	2	1	2	1	1	2	2	18	
	WL	0	0	1	3	2	1	2	1	1	2	3	2	18	
	LL	0	6	2	0	3	1	1	0	5	2	2	2	24	84

CHAPTER 5

INVESTMENT IMPLICATIONS

5.1. INTRODUCTION

In the previous chapter we have provided empirical evidence on the performance and persistence of our sample of Portuguese equity funds. We have seen that, in the context of quarterly excess returns, there is some evidence of persistence for the overall sample period. The question we now address is whether an investor following a persistence strategy can earn abnormal returns, as suggested by previous authors (e.g.: HENDRICKS, PATEL and ZECKHAUSER [1993], BROWN and GOETZMANN [1995], MALKIEL [1995], ELTON, GRUBER and BLAKE [1996], GRUBER [1996], CARHART [1996]). We investigate this question by simulating strategies that exploit the persistence effect. The idea is simple: forming portfolios of mutual funds on lagged one-period returns and evaluating the resulting performance. In this way, we intend to assess the economic significance of the statistical persistence findings. In addition, this approach complements the analysis of the previous chapter, in the sense that it tests indirectly the degree of winning/losing in the subsequent period, which is not tested in the contingency tables (nor is the degree of winning/losing in the initial period).

5.2. METHODOLOGY FOR EVALUATING THE PERFORMANCE POTENTIAL OF STRATEGIES BASED ON THE PERSISTENCE OF FUND RETURNS

The methodology we use is broadly similar to that of HENDRICKS, PATEL and ZECKHAUSER [1993][120] and GRUBER [1996]. For every quarter (and year) in our sample, we ranked all mutual funds on the basis of their prior quarterly (or yearly) returns (unadjusted or adjusted for risk). Then, we form two equally weighted portfolios of funds: the winning funds and the losing funds.[121] These portfolios are referred to as rank portfolios. The performance of each rank portfolio over the following quarter (and year) is then evaluated according to some measure of performance. In the terminology of GRUBER [1996], this process involves two distinct periods: the selection period and the performance period. We estimate the statistical significance of the results based on the following criteria:

1) Calculating the correlation in the rank orderings of the selection and performance period. This is performed through the

---

120 This methodology was replicated by BROWN and GOETZMANN [1995] and CARHART [1997].

121 Other studies include rank portfolios of octiles or deciles. Since our limited sample size does not allow for such finer subdivisions, in addition to the portfolios of winners and losers, we rank the funds into 12 divisions (duodecimos), exhausting the sample of these twelve funds.

computation of the Spearman Rank Correlation Coefficient (ρ), defined as follows:

$$\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \qquad [5.1]$$

Where d_i are the differences of the ranked pairs, i.e., the differences for each portfolios rank from its performance rank in the next quarter. The null hypothesis here is that the performance rankings are randomly ordered, against the two-sided alternative of existence of some correlation. If ρ exceeds the $1-\alpha/2$ quantile or if ρ is less than the $1-\alpha/2$ quantile, the null hypothesis will be rejected at a level α.

2) Computing the means of the time series of the differences in excess returns on the rank portfolios and testing whether the differences are statistically different from zero. To infer the statistical significance of the differences in mean excess returns, we will use both the parametric t-test and its non-parametric analog, the Mann-Whitney U test.[122]

[122] The t-test is a more powerful test when both populations follow a normal distribution. Yet, the Mann-Whitney U test is a most useful alternative to the parametric t-test, particularly in the cases when we wish to avoid the restrictive assumption required by the parametric test relatively to the normal distribution of the populations. Indeed, the t-test has very little power compared to the Mann-Whitney U test when applied in the context of nonnormal distributions [CONOVER, 1980]. Also, when the population distributions are normal, at least in samples of moderate size, the nonparametric test is only a little less powerful than the corresponding t-test [NEWBOLD, 1991].

The t-test for two population means μ_X and μ_Y is used to test the hypothesis:

$$H_0 : \mu_X - \mu_Y = 0$$

against the alternative $\quad H_1 : \mu_X - \mu_Y \neq 0.$

If the two random samples of N_X and N_Y observations come from normal distribution and a common variance, the following statistic:

$$t = \frac{\overline{X} - \overline{Y}}{S\sqrt{\frac{1}{N_X} + \frac{1}{N_Y}}} \qquad [5.2]$$

has a *t* distribution with ($N_X + N_Y - 2$) degrees of freedom, where S^2 is an estimate of the common population variance (pooled variance), obtained through the observed sample variances S_X^2 and S_Y^2, as follows:

$$S^2 = \frac{(N_X - 1)S_X^2 + (N_Y - 1)S_Y^2}{N_X + N_Y - 2} \qquad [5.3]$$

On the other hand, the Mann-Whitney U tests the null hypothesis that the central locations of two population distributions are identical [CONOVER, 1980],[123] that is:

$$H_0 : \mu_X = \mu_Y$$

$$\text{and} \quad H_1 : \mu_X \neq \mu_Y \ .$$

Let n_1 be the number of cases in the smaller of two independent groups and n_2 the number of cases in the larger. If the sample observations are ranked, with R_1 denoting the sum of the ranks for the first population, the Mann-Whitney U test statistic is [SIEGEL, 1975]:[124]

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \qquad [5.4]$$

Under the null hypothesis, and for large sample sizes,[125] the distribution of U is well approximated by the standard normal distribution:

---

123 In testing the null hypothesis that the central location of both populations are the same, it is assumed that the two population distributions are identical.

124 Or equivalently, $U = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$.

125 The approximation is adequate for $n > 20$ observations [SIEGEL, 1975].

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} \sim N(0,1) \qquad [5.5]$$

3) Calculating Jensen's alpha resulting from the portfolios' performance in the post-formation period:

$$R_{p,t} = \alpha_p + \beta_p \left(R_{m,t} - R_{f,t} \right) + \varepsilon_{p,t} \qquad [5.6]$$

where **p** is either a winner/loser portfolio or a zero-investment best-worst fund. This type of strategy involves hypothetically investing in the top-performing fund and shorting the worst-performing fund, and corresponds to the potential maximum gain from exploiting performance persistence.

The significance of α is evaluated by a heteroscedastic-autocorrelated-consistent statistic following the NEWEY and WEST [1987] procedure.

5.3. EMPIRICAL EVIDENCE

5.3.1. One quarter evaluation period

Table 5.1 displays the results of a simulated strategy of investing in the top/bottom performing funds. Each quarter, and for the overall sample period (April 1994 - March 1998) rank portfolios of winners and losers are formed and reconstituted based on the preceding year performance ranks. We report on the summary statistics for each of the winners/losers portfolio, and a best/worst portfolio. Also, we consider a simulated strategy of buying winners and selling losers - a hypothetical situation of being long in the best performers and short in the worst performers. The statistical significance of this type of strategy is evaluated through the parametric t-test and the non-parametric Mann Whitney U test.

For each rank portfolio we also compute the CAPM alphas relatively to the PSI index, and an EWMF alpha, which refers to risk-adjusted returns relatively to an equally weighted average of returns of all funds that compose our sample.[126] The statistical significance of the alphas is evaluated following the NEWEY and WEST [1987] procedure, which is used to produce efficient

[126] The EWMF alpha corresponds to alpha in a context of fund betas equal to 1.

estimators of the sampling variances in the presence of heteroscedasticity and serial correlation.

Table 5.1

Rank portfolios: comparison of performance for different ranking criteria (One quarter evaluation period)

This table reports summary statistics for rank portfolios (formed according to the methodology described in the text). NEWEY and WEST [1987] t-statistics, which correct for heteroscedasticity and autocorrelation, are shown in parenthesis below the coefficient estimates. The p-values of observing the estimated statistic under the hypothesis of no differences in mean excess returns are reported in brackets (in the following order: first the parametric, and then the nonparametric p-value).

Portfolios formed on the basis of:		Top performing funds	Bottom performing funds	Top-Bottom	Best	Worst	Best fund-Worst fund
Excess Returns							
	Mean Excess Returns	0.00103	0.00088	0.00015	0.00110	0.00083	0.00026
				[0.595]			[0.379]
				[0.342]			[0.230]
	Standard Deviation	0.00639	0.00610	0.00189	0.00628	0.00658	0.00367
	CAPM Beta	0.74	0.73	0.02	0.70	0.73	-0.02
	CAPM alpha	0.00016	0.00002	0.00013	0.00025	-0.00006	0.00031
		(1.55)	(0.27)	(2.44)	(2.24)	(-0.44)	(2.46)
	EWMF Alpha	0.00006	-0.00006	0.00011	0.00017	-0.00013	0.00030
		(1.94)	(-1.94)	(1.94)	(2.53)	(-1.37)	(2.36)
Alpha							
	Mean Excess Returns	0.00100	0.00090	0.00010	0.00104	0.00065	0.00039
				[0.740]			[0.174]
				[0.466]			[0.088]
	Standard Deviation	0.00608	0.00643	0.00202	0.00627	0.00621	0.00340
	CAPM Beta	0.72	0.75	-0.03	0.70	0.68	0.02
	CAPM Alpha	0.00016	0.00002	0.00013	0.00022	-0.00015	0.00037
		(1.94)	(0.21)	(2.31)	(1.87)	(-1.13)	(3.18)
	EWMF Alpha	0.00008	-0.00008	0.00015	0.00013	-0.00022	0.00036
		(2.58)	(-2.58)	(2.58)	(1.93)	(-2.58)	(3.14)

At the aggregate level, the results are consistent with the possibility of performance persistence. We start by comparing the performance of the top-ranked funds to the bottom ranked funds. As can be seen, the winning funds exhibit, in all cases, higher risk-adjusted returns than the losing funds. Also, these differences in alpha performance are statistically significant at the 5 percent level, as corroborated by the respective NEWEY and WEST [1987] t-statistics. For the rank portfolios formed on the basis of excess returns:

$$t_{\text{top-bottom}} \approx 2.44 > t_{\text{critical}\ (923,0.025)} \approx 1.962$$

and for those formed on the basis of alpha returns:

$$t_{\text{top-bottom}} \approx 2.31 > t_{\text{critical}\ (923,0.025)} \approx 1.962$$

The statistical significance is further reinforced when we compare the differences in risk-adjusted returns from a strategy of investing in the best performing and in the worst performing fund. In this situation, the t-statistics are, for the rankings based on excess returns:

$$t_{\text{best-worst}} \approx 2.46 > t_{\text{critical}\ (923,0.025)} \approx 1.962$$

and

$$t_{\text{best-worst}} \approx 3.18 > t_{\text{critical}\ (923,0.025)} \approx 2.581$$

for the rankings based on excess returns and alphas, respectively.

The test statistics are, of course, even more evident when risk-adjusted returns are measured in relation to peer funds,[127] as shown by the EWMF alpha statistics.

We can also verify that the excess returns of the winning funds are superior to those of the losing funds. However, the difference has not been statistically significant whatever test (parametric or non-parametric) is used. In fact, the significance levels resulting from the t-test and the Mann-Whitney U-test (0.595 and 0.342, respectively) do not allow for rejection of the null hypothesis of no differences in means. Only for the case in which an investor purchases the best performing fund each quarter, the strategy is feasible, as the excess returns are statistically different from those resulting from the worst performing fund. Even so, this conclusion is valid only under the Mann-Whitney p-value (=0.088).

It is also interesting to examine more in detail the performance of a strategy in which funds are ranked and placed each quarter into duodecimos according to some performance criteria. Each column of Table 5.2 shows the average excess returns and alpha obtained from investing in each rank duodecimo.

[127] Which is, for many investors, the relevant benchmark for comparison purposes.

Table 5.2

Average performance by duodecimos for different ranking criteria

(One quarter evaluation period)

Mutual funds are sorted from 1 to 12 each quarter based on their previous quarter performance. The fund with the highest past one quarter performance comprises duodecimo 1 and the fund with the lowest performance duodecimo 12. The resulting performance for all subdivisions is presented. The statistical significance of the difference in mean excess returns is evaluated through parametric and non-parametric tests. The statistical significance of the differences in risk-adjusted returns is assessed through the NEWEY and WEST [1987] procedure for correction of heteroscedastic and autocorrelated disturbances.

	Portfolios formed on the basis of:			
	Excess Returns		Alphas	
	Excess Returns	Alphas	Excess Returns	Alphas
(BEST) 1	0.00110	0.00025	0.00104	0.00022
2	0.00111	0.00025	0.00116	0.00025
3	0.00092	-0.00004	0.00111	0.00031
4	0.00108	0.00028	0.00094	0.00009
5	0.00097	0.00016	0.00081	-0.00003
6	0.00098	0.00005	0.00092	0.00008
7	0.00097	0.00009	0.00104	0.00017
8	0.00093	0.00010	0.00095	0.00008
9	0.00089	0.00006	0.00088	-0.00014
10	0.00089	-0.00005	0.00099	0.00015
11	0.00081	-0.00001	0.00094	0.00004
(WORST) 12	0.00083	-0.00006	0.00065	-0.00015
Spearman Rank Coefficient	0.853***	0.706**	0.573*	0.706**
Best fund-Worst Fund	0.00026	0.00031**	0.00039*	0.00037***

* Statistically significant at the 10 percent level
** Statistically significant at the 5 percent level
*** Statistically significant at the 1 percent level

Analysis of the previous table shows that the past rankings of funds are highly correlated with future rankings, as shown by the

Spearman Rank Correlation Coefficient. Except for the case when portfolios are formed on the basis of alpha and evaluated according to excess returns, any other ranking criteria might suggest the possibility of performance persistence, as indicated by the significance of the rank correlation coefficients (at the 5 percent and 1 percent level).

However, these findings are not corroborated by the tests for differences in mean excess returns of the best and worst fund.[128] This contradicting evidence can be perhaps explained by the fact that the Spearman Rank Correlation test treats the order of each rank portfolio equally, therefore not considering the possibility that performance persistence is concentrated in the extremes (top or bottom performing funds).

The results for the individual quarters are presented in Tables 5.3 and 5.4.

[128] Only for the alpha criteria the differences in excess returns of the best and worst fund has statistical significance, although only at a 10 percent level.

Table 5.3.

Individual quarterly excess returns

Mutual funds are sorted into winners and losers each quarter based on their previous quarter performance. The fund with the highest past one quarter performance and the fund with the lowest performance are also considered. The resulting excess returns for all quarters are presented. The statistical significance of the difference in mean excess returns is evaluated through parametric and non-parametric tests.

	Portfolios formed on the basis of					
	Excess Returns			Alphas		
	Winners	Losers	Winners-Losers	Winners	Losers	Winners-Losers
3Q94	0.00043	0.00065	-0.00021	0.00054	0.00054	0.00000
4Q94	0.00021	0.00010	0.00011	0.00026	0.00004	0.00022
1Q95	-0.00073	-0.00061	-0.00012	-0.00073	-0.00061	-0.00012
2Q95	0.00030	-0.00002	0.00031	0.00022	0.00006	0.00015
3Q95	-0.00061	-0.00095	0.00034	-0.00061	-0.00095	0.00034
4Q95	-0.00030	-0.00107	0.00078	-0.00054	-0.00083	0.00028
1Q96	0.00134	0.00127	0.00007	0.00134	0.00127	0.00007
2Q96	0.00108	0.00121	-0.00013	0.00120	0.00109	0.00011
3Q96	0.00024	0.00001	0.00023	0.00028	-0.00004	0.00032
4Q96	0.00123	0.00103	0.00020	0.00121	0.00104	0.00017
1Q97	0.00237	0.00215	0.00022	0.00229	0.00223	0.00006
2Q97	0.00306	0.00290	0.00016	0.00281	0.00315	-0.00034
3Q97	0.00130	0.00121	0.00010	0.00125	0.00126	-0.00001
4Q97	0.00010	0.00013	-0.00003	0.00017	0.00007	0.00010
1Q98	0.00542	0.00538	0.00004	0.00542	0.00538	0.00004

	Portfolios formed on the basis of					
	Excess Returns			Alphas		
	Best	Worst	Best-Worst	Best	Worst	Best-Worst
3Q94	0.00059	0.00081	-0.00022	0.00010	-0.00033	0.00042
4Q94	0.00013	0.00033	-0.00019	0.00013	0.00033	-0.00019
1Q95	-0.00031	-0.00079	0.00048	-0.00031	-0.00079	0.00048
2Q95	0.00032	0.00026	0.00006	0.00032	-0.00005	0.00037
3Q95	-0.00055	-0.00150	0.00095	-0.00055	-0.00150	0.00095
4Q95	-0.00063	-0.00151	0.00088	-0.00063	-0.00151	0.00088
1Q96	0.00190	0.00112	0.00078	0.00190	0.00113	0.00077
2Q96	0.00117	0.00108	0.00010	0.00117	0.00108	0.00010
3Q96	0.00076	0.00000	0.00077 *	0.00076	-0.00014	0.00090 *
4Q96	0.00164	0.00097	0.00068	0.00164	0.00128	0.00037
1Q97	0.00238	0.00205	0.00033	0.00238	0.00205	0.00033
2Q97	0.00293	0.00292	0.00001	0.00293	0.00273	0.00020
3Q97	0.00118	0.00207	-0.00089	0.00118	0.00077	0.00042
4Q97	-0.00090	-0.00039	-0.00052	-0.00090	-0.00039	-0.00052
1Q98	0.00583	0.00505	0.00077	0.00583	0.00505	0.00077

* Statistically significant at the 10 percent level

For the individual quarters, it is not always the case that the winning group beats the losing group. In fact, we can examine the number of time periods the losers outperform the winners (or that the worst performing fund outperforms the best performing fund). While the mean excess returns to a strategy of investing in the best funds are, for most cases, greater than those of the worst funds, we can see that, with the exception of 3Q96,[129] investing in the winners (or best) fund does not yield excess returns statistically different from those generated by the loser (or worst) funds.

[129] Even so, only at a 10 percent significance level, as given by the respective p-values. For the excess returns criteria, the p-value$_{best-worst}$ is 0.093 according to the Mann-Whitney U-test. The parametric t-test is not significant (p-value$_{best-worst}$ ≈ 0.015). For the alpha criteria both tests allow for rejection of the null hypothesis: the p-values are 0.088 and 0.071 for the parametric and non-parametric tests, respectively.

Table 5.4.

Individual quarterly alphas

Mutual funds are sorted into winners and losers each quarter based on their previous quarter performance. The fund with the highest past one quarter performance and the fund with the lowest performance are also considered. The resulting risk-adjusted returns for all quarters are presented. The statistical significance of the differences in risk-adjusted returns is assessed through the NEWEY and WEST [1987] procedure for correction of heteroscedastic and autocorrelated disturbances.

	Portfolios formed on the basis of					
	Excess Returns			Alphas		
	Winners	Losers	Winners-Losers	Winners	Losers	Winners-Losers
3Q94	0.00032	0.00037	-0.00004	0.00027	0.00043	-0.00016
4Q94	0.00028	0.00014	0.00015	0.00031	0.00011	0.00020
1Q95	-0.00038	-0.00023	-0.00016	-0.00038	-0.00023	-0.00016
2Q95	0.00016	-0.00015	0.00032 *	0.00009	-0.00008	0.00017
3Q95	-0.00015	-0.00050	0.00035 **	-0.00015	-0.00050	0.00035 **
4Q95	-0.00007	-0.00091	0.00085 ***	-0.00028	-0.00070	0.00042
1Q96	0.00026	0.00027	-0.00001	0.00026	0.00027	-0.00001
2Q96	0.00021	0.00031	-0.00009	0.00041	0.00011	0.00030
3Q96	0.00027	-0.00001	0.00028 **	0.00031	-0.00005	0.00036 ***
4Q96	0.00031	0.00015	0.00016	0.00033	0.00014	0.00019
1Q97	0.00045	0.00025	0.00020	0.00044	0.00026	0.00018
2Q97	0.00028	0.00021	0.00007	0.00019	0.00030	-0.00011
3Q97	-0.00002	-0.00009	0.00008	-0.00008	-0.00003	-0.00005
4Q97	-0.00053	-0.00052	-0.00001	-0.00041	-0.00064	0.00023
1Q98	0.00043	0.00033	0.00010	0.00043	0.00033	0.00010

	Portfolios formed on the basis of					
	Excess Returns			Alphas		
	Best	Worst	Best-Worst	Best	Worst	Best-Worst
3Q94	0.00045	0.00056	-0.00011	-0.00004	-0.00014	0.00010
4Q94	0.00022	0.00044	-0.00022	0.00022	0.00044	-0.00022
1Q95	-0.00003	-0.00039	0.00036	-0.00003	-0.00039	0.00036
2Q95	0.00014	0.00009	0.00005	0.00014	-0.00018	0.00031
3Q95	-0.00011	-0.00097	0.00086	-0.00011	-0.00097	0.00086
4Q95	-0.00044	-0.00170	0.00126 *	-0.00044	-0.00170	0.00126 *
1Q96	0.00085	-0.00005	0.00091 *	0.00085	0.00057	0.00028
2Q96	0.00075	0.00007	0.00067 **	0.00075	0.00007	0.00067 **
3Q96	0.00081	0.00000	0.00081 **	0.00081	-0.00023	0.00105 **
4Q96	0.00083	0.00020	0.00063 *	0.00083	0.00042	0.00042
1Q97	0.00047	0.00046	0.00000	0.00047	0.00046	0.00000
2Q97	0.00031	0.00017	0.00014	0.00031	0.00011	0.00020
3Q97	-0.00016	0.00052	-0.00067	-0.00016	-0.00031	0.00016
4Q97	-0.00136	-0.00101	-0.00035	-0.00136	-0.00101	-0.00035
1Q98	0.00058	0.00049	0.00009	0.00058	0.00049	0.00009

* Statistically significant at the 10 percent level
** Statistically significant at the 5 percent level
*** Statistically significant at the 1 percent level

The differences in risk-adjusted return between the rank portfolios are more evident. Clearly there are some quarters for which an investor can obtain risk-adjusted excess returns by investing in the top-performing funds. While the differences in performance over these quarters are statistically significant (for 3Q96 this is very salient, at the 5 percent and even at the 1 percent level),[130] this is only verified for three of four quarters out of fifteen.

In summary, in this section we have seen that despite the weak evidence that a strategy of investing in the top performing funds (particularly at the individual quarter level) can statistically outperform the bottom performing funds (particularly at the individual quarter level), the overall results suggest there is some correlation between each fund past and future rankings.

Since the one quarter holding period is a theoretical situation, and not practical due to transaction costs, we will now examine the issue in terms of a one year holding period.

5.3.2. One year evaluation period

Table 5.5 summarizes the performance statistics for portfolios formed on the basis of excess returns and risk-adjusted returns.

[130] When portfolios are formed by alphas, the difference in risk-adjusted returns between the winning and losing funds is statistically significant at the 1 percent level. Curiously, we have already noted the possibility of performance persistence for this quarter when examining the quarterly contingency table of alpha returns (section 4.4.3).

Table 5.5

Rank portfolios: comparison of performance for different ranking criteria (One year evaluation period)

This table reports summary statistics for rank portfolios (formed according to the methodology described in the text). NEWEY and WEST [1987] t-statistics, which correct for heteroscedasticity and autocorrelation, are shown in parenthesis below the coefficient estimates. The p-values of observing the estimated statistic under the hypothesis of no differences in mean excess returns are reported in brackets (in the following order: first the parametric, and then the nonparametric p-value.)

Portfolios formed on the basis of:		Top performing funds	Bottom performing funds	Top-Bottom	Best	Worst	Best fund-Worst fund
Excess Returns							
	Mean Excess Returns	0.00118	0.00121	-0.00003	0.00133	0.00122	0.00011
				[0.927]			[0.769]
				[0.948]			[0.826]
	Standard Deviation	0.00625	0.00596	0.00275	0.00701	0.00683	0.00450
	CAPM Beta	0.77	0.82	-0.05	0.70	0.76	-0.06
	CAPM alpha	0.00005	-0.00003	0.00009	0.00028	0.00005	0.00023
		(0.63)	(-0.36)	(1.49)	(2.23)	(0.40)	(1.70)
	EWMF Alpha	0.00004	-0.00004	0.00009	0.00024	0.00005	0.00020
		(1.51)	(-1.51)	(1.51)	(2.17)	(0.53)	(1.41)
Alpha							
	Mean Excess Returns	0.00119	0.00121	-0.00002	0.00120	0.00145	-0.00025
				[0.949]			[0.490]
				[0.999]			[0.234]
	Standard Deviation	0.00627	0.00599	0.00294	0.00738	0.00664	0.00463
	CAPM Beta	0.76	0.83	-0.07	0.73	(0.74)	-0.01
	CAPM Alpha	0.00005	-0.00003	0.00008	0.00011	0.00034	-0.00024
		(0.59)	(-0.33)	(1.47)	(0.68)	(3.05)	(-1.60)
	EWMF Alpha	0.00004	-0.00004	0.00008	0.00006	0.00035	-0.00028
		(1.44)	(-1.44)	(1.44)	(0.56)	(4.12)	(-1.89)

These results show that when portfolios are constituted and held over one year periods, there is little evidence that might suggest the possibility of performance persistence. In fact, these results are perhaps more consistent with the hypothesis of performance reversals, since the bottom performing funds outperform, in many situations, the top performing funds. Anyhow, the possible differences (positive or negative) in post-selection performance measures of the rank portfolios are practically insignificant. This is verified by the

NEWEY and WEST t-statistics for alpha coefficients (less than the $t_{critical\ (735,0.025)} \approx 1.963$), and by the p-values (higher than 0.05) resulting from the parametric and non-parametric tests for the differences in mean excess returns.

Table 5.6 shows the post-selection excess returns and alphas for funds 1 to 12 and presents the statistical significance of the differences.

Table 5.6

Average performance by duodecimos for different ranking criteria

(One year evaluation period)

Mutual funds are sorted from 1 to 12 each year based on their previous year performance. The fund with the highest past one year performance comprises duodecimo 1 and the fund with the lowest performance duodecimo 12. The resulting performance for all subdivisions is presented. The statistical significance of the difference in mean excess returns is evaluated through parametric and non-parametric tests. The statistical significance of the differences in risk-adjusted returns is assessed through the NEWEY and WEST procedure for correction of heteroscedastic and autocorrelated disturbances.

	Portfolios formed on the basis of:			
	Excess Returns		Alphas	
	Excess Returns	Alphas	Excess Returns	Alphas
(BEST)1	0.00133	0.00028	0.00120	0.00011
2	0.00107	-0.00002	0.00125	0.00015
3	0.00122	0.00004	0.00116	0.00001
4	0.00109	-0.00004	0.00103	-0.00007
5	0.00124	0.00001	0.00123	0.00006
6	0.00114	0.00007	0.00125	0.00005
7	0.00134	0.00007	0.00131	0.00009
8	0.00094	-0.00004	0.00105	-0.00045
9	0.00116	-0.00050	0.00112	-0.00006
10	0.00142	0.00020	0.00108	-0.00012
11	0.00119	0.00001	0.00123	0.00002
(WORST)12	0.00122	0.00005	0.00145	0.00034
Spearman Rank Coefficient	-0.084	0.035	-0.112	0.203
Best fund-Worst Fund	0.00011	0.00023*	-0.00025	-0.00024

* Statistically significant at the 10 percent level

We find no evidence of performance persistence. Neither the Spearman Correlation Coefficient nor the tests performed for the differences in performance between the top performing and low performing funds are statistically significant.

Table 5.7 below reports summary statistics for the individual year results.

Table 5.7.

Individual annual excess returns

Mutual funds are sorted into winners and losers each year based on their previous year performance. The fund with the highest past one year performance and the fund with the lowest performance are also considered. The resulting performance for all years is presented. The statistical significance of the difference in mean excess returns is evaluated through parametric and non-parametric tests.

	Portfolios formed on the basis of					
	Excess Returns			Alphas		
	Winners	Losers	Winners-Losers	Winners	Losers	Winners-Losers
Apr.95-Mar.96	0.00009	-0.00009	0.00017	0.00010	-0.00010	0.00020
Apr.96-Mar.97	0.00112	0.00116	-0.00004	0.00112	0.00116	-0.00004
Apr.97-Mar.98	0.00232	0.00254	-0.00022	0.00232	0.00254	-0.00022

	Portfolios formed on the basis of					
	Excess Returns			Alphas		
	Best	Worst	Best-Worst	Best	Worst	Best-Worst
Apr.95-Mar.96	-0.00013	-0.00021	0.00008	-0.00013	0.00048	-0.00061 *
Apr.96-Mar.97	0.00157	0.00155	0.00002	0.00157	0.00155	0.00002
Apr.97-Mar.98	0.00252	0.00231	0.00021	0.00214	0.00231	-0.00017

* Statistically significant at the 10 percent level

Analysis of the above table shows that almost only by chance will a strategy of investing in the previous' year top performing funds outperform the bottom performing funds. Where the mean excess returns to the strategy is positive, it is insignificant. It is also possible to find performance reversals, in the sense that the worst performing fund in the selection period produced higher returns than the best performing fund. Moreover, for the first year, this difference is statistically significant at the 10 percent level, as supported by the respective nonparametric p-value (≈ 0.067).[131]

Table 5.8.

Individual annual alphas

Mutual funds are sorted into winners and losers each year based on their previous year performance. The fund with the highest past one year performance and the fund with the lowest performance are also considered. The resulting risk-adjusted returns for all years are presented. The statistical significance of the differences in risk-adjusted returns is assessed through the NEWEY and WEST [1987] procedure for correction of heteroscedastic and autocorrelated disturbances.

	Portfolios formed on the basis of					
	Excess Returns			Alphas		
	Winners	Losers	Winners-Losers	Winners	Losers	Winners-Losers
Apr.95-Mar.96	-0.00003	-0.00019	0.00016	-0.00001	-0.00021	0.00019 **
Apr.96-Mar.97	0.00024	0.00024	-0.00001	0.00024	0.00024	-0.00001
Apr.97-Mar.98	-0.00003	-0.00012	0.00009	-0.00006	-0.00009	0.00003

	Portfolios formed on the basis of					
	Excess Returns			Alphas		
	Best	Worst	Best-Worst	Best	Worst	Best-Worst
Apr.95-Mar.96	-0.00022	-0.00045	0.00023	-0.00022	0.00038	-0.00060 **
Apr.96-Mar.97	0.00071	0.00072	-0.00002	0.00071	0.00072	-0.00002
Apr.97-Mar.98	0.00031	-0.00004	0.00035 *	-0.00020	-0.00004	-0.00016

* Statistically significant at the 10 percent level
** Statistically significant at the 5 percent level

131 The corresponding t-statistic ≈ -1.418, less than the $t_{critical\ (243,0.05)} \approx$ 1.6551, and p-value ≈ 0.157, therefore not allowing for rejection of the null hypothesis.

The scenario in terms of risk adjusted-returns is broadly similar. The winning (or best) funds outperform the losing (or worst) funds just as many times as the losers (or worst funds) outperform the winners (or best funds). Although for two of the years we can find statistical evidence (at the 5 and 10 percent level, respectively) that the top-performing (best) funds generate higher returns than the bottom performing (worst) funds, there is also evidence of a reversal pattern in performance. Indeed, we find statistical significance (at the 5 percent level) of a reversal in performance for the first year, relatively to best/worst funds (ranked by alphas).

5.4 CONCLUSIONS

In this chapter we examined the potential performance relatively to simulated strategies of investing in the past winning funds. This is a question of practical significance for investors. Besides investigating whether past rankings of funds help predict future rankings, it is important to further question whether is worthwhile exploiting persistence strategies. Although at the quarterly level the results indicate that there is a significant correlation between fund rankings in the selection and post-selection period, we have seen that this is not reflected in practical advantages for investors. In fact, in general, the resulting performance has not turned out to be statistically different from the returns of a strategy of investing in the losing pool of funds.

These results are not robust to the time period chosen. In terms of the one year evaluation period, we find no evidence whatsoever of performance persistence. On the contrary, we even find evidence of performance reversals, indicating that the worst performing fund in the selection period outperformed the best performing funds.

These conclusions contradict those documented in the literature. MALKIEL [1995], BROWN and GOETZMANN [1995] and GRUBER [1996] generally find that past performance can help to obtain positive performance in the future.

Anyhow, even if these types of simulated strategies worked well, the results would hardly ever be achievable, in the sense that investors would consistently outperform the market by investing in the recently top-performing funds. This is because most funds charge load (sales) fees.[132] Strategies that involve switching periodically to the best performing fund would incur very high expenditure, and therefore are impractical.

[132] For our sample of funds, all except one charge load fees, ranging from 1% to 2% in the situations when funds are held for less than a one-year period.

CHAPTER 6

SUMMARY, CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

Over the last three decades, the performance evaluation of managed funds has been one of the most studied topics in the field of finance. Nevertheless, only recently attention has focused on the investigation of predictable patterns in mutual fund performance. Despite recent empirical research supporting the idea that future performance is (at least in part) predictable from past performance, it must be interpreted with caution in light of the somewhat mixed evidence documented in the literature.

In this research we addressed this major issue of mutual fund performance persistence. In chapter 2 we conducted a review of the literature on performance evaluation. It was our concern to describe and discuss the theoretical frameworks used for performance evaluation purposes, in general, and to focus on the most pertinent and controversial issues in debate, which ultimately can constitute the basis for many questions regarding performance persistence. In this sense, we emphasized topics such as the benchmark problem and investment style. We then presented and discussed the most relevant evidence and methodologies for the assessment of performance persistence. Among those discussed, and due to limitations related to the size of the sample, the methodology based on contingency table of winners and losers became the focus of our attention. The next logical step, carried out in chapter 3, was to describe more in detail the method of analysis based on contingency tables. In this chapter, besides having described and compared several criteria to assess performance persistence, we gave special attention to the issue of small expected frequencies and presented the adjustments required in the case of a limited sample size.

In chapter 4 we investigated whether the persistence phenomenon detected in the U.S. and the U.K. also exists in relation to the Portuguese fund market (a small market). We first described the dataset used in the study and, then, provided empirical evidence on the performance persistence of a sample of Portuguese equity funds. We have documented some evidence of performance persistence in terms of total returns, which disappears after returns are adjusted for risk. However, when considering funds' risk characteristics, we found strong evidence of persistent "risk winners" and "risk losers". Furthermore, identification of the term structure of performance persistence allowed us to observe that the persistence phenomenon is not robust to the consideration of different time periods of return measurement. Finally, despite the fact that average performance persistence is relatively small, it is the case that some individual funds are persistent winning and losing funds.

In the course of the research we have identified several problems resulting from the application of alternative criteria for evaluating performance persistence. The consideration of the most appropriate criteria to apply in various circumstances has also been discussed. In addition, we have also emphasized the importance of correcting for the small sample bias (through the Yates continuity correction, Fisher's exact p-value or, alternatively, the bootstrap), and shown that the unadjusted significance statistics overstate the persistence phenomenon.

The discussion of our contribution in methodology is important for several reasons. In fact, although it is applied to Portuguese mutual funds, the extensions of much of this methodology (along with adjustments for small samples) are important to other small markets and even to contexts other than portfolio funds (v.g: companies in one industry, dealers within a bank, divisions of a company, etc.).

In chapter 5 we investigated the implications of following performance persistence strategies. Although we have observed some correlation between fund rankings in the selection period and the postselection period, we have found no evidence of exploitable persistence.

In conclusion, while over the observed sample period some fund managers did outperform the market, many did not. Furthermore, on average, we find little evidence supportive of performance persistence. In this way our results relative to the Portuguese fund market contradict the apparent strength of the persistence evidence found in the U.S. and the U.K. markets.

Overall, the evidence is not inconsistent with market efficiency. We cannot advise investors that there is a reliable strategy (based on historical information) for selecting funds expected to perform well in the future. In this line of thought, index funds would constitute an attractive alternative for investors, which would have to be satisfied with the diversification, professional management and low transaction costs services provided by mutual fund ownership.

On the other hand, despite the fact that, on average, there is weak evidence of performance persistence, we have found that some individual fund managers appear to have "hot hands". In particular, it seems that a fund manager (fund A) does have sustained superior performance on a risk-adjusted basis. Whether this fact is evidence of market inefficiency is a debatable question. We suggest this result is more consistent with a GROSSMAN and STIGLITZ [1980] view of efficient markets, in which some managers may just earn enough to be compensated for their time and effort in obtaining costly information.

We also have detected managers with "cold hands". The existence of persistent losing funds is similar to the results of previous studies and constitutes a puzzling finding. Why do investors remain in funds that consistently perform poorly? Possible explanations such as the existence of disadvantaged investors (GRUBER [1996]), psychological reasons (GOETZMANN and PELES [1997]) or immunity from periodic performance review (BROWN, GOETZMANN, IBBOTSON and ROSS [1992]) might help to solve the puzzle.

Of course we can always question the fact that the CAPM may not constitute an appropriate framework for risk adjustment, and propose the use of multi-index models. While this is consensual, it would only be pertinent if we had found clear evidence of persistence of performance, which could be explained by the returns being attributable to undetected sources of systematic risk. Furthermore, and at least in relation to the last two years of the sample period, the return-generating model seems to perform quite well.

The issue of performance persistence is far from being solved, considering all these points of discussion. As Brown and Goetzmann comment: "(...) *the nature of mutual fund persistence is more complicated than previous researchers, including the current authors, have understood*" [BROWN and GOETZMANN, 1995, p. 697]. A possible path of future research might be the consideration of the cyclical economic conditions which affect the activities of fund managers at different time periods. The additional information used by conditional measures of performance evaluation (FERSON and SCHADT [1996], CHRISTOPHERSON, FERSON and GLASSMAN [1998]) might allow for a better assessment of persistence of performance.

Finally, because studies to date on performance persistence do not thoroughly consider the non-normality of returns, the degree of winning and losing, and explanations for performance, no doubt performance persistence evaluation will be a dynamic research in Portugal and elsewhere in the process of raising and solving such problems.

BIBLIOGRAPHY

ADMATI, Anat R.; **BHATTACHARYA**, Sudipto; **PFLEIDERER**, Paul; **ROSS**, Stephen **(1986)** "On timing and selectivity." *Journal of Finance* (July), vol. XLI, pp. 715-732.

ALEXANDER, Gordon J.; **STOVER**, Roger D. **(1980)** "Consistency of mutual fund performance during varying market conditions." *Journal of Economics and Business* (Spring), pp. 219-226.

ANDERSEN, Erling B. **(1990)** *The statistical analysis of categorical data,* Springer-Verlag, Berlin, 522 pages.

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

ANKRIM, Ernest M. **(1992)** "Risk-adjusted performance attribution." *Financial Analysts Journal* (March-April), vol. 48, no. 2, pp.75-82.

ARMADA, Manuel J. R. **(1992)** "On the investigation of timing and selectivity in portfolio management." PhD Dissertation (December), Manchester Business School.

BAILEY, Jeffery V. **(1992a)** "Are manager universes acceptable performance benchmarks?" *Journal of Portfolio Management* (Spring), vol.18, no. 3, pp.9-13.

BAILEY, Jeffery V. **(1992b)** "Evaluating benchmark quality." *Financial Analysts Journal* (May-June), vol.48, no.2, pp.33-39.

BAILEY, Jeffery V. **;** **TIERNEY**, David E. **(1993)** "Gaming manager benchmarks." *Journal of Portfolio Management* (Summer), vol. 19, no. 4, pp.37-40.

BAL, Yasemin**;** **LEGER**, Lawrence A. **(1996)** "The performance of UK investment trusts." *Services Industries Journal* (January), vol. 16, no. 1, pp.67-81.

BALZER, Leslie A. **(1995)** "Measuring investment risk: a review." *Journal of Investing* (Fall), vol. 4, no. 3, pp.5-16.

BANZ, Rolf **(1981)** "The relationship between return and market value of common stocks." *Journal of Financial Economics* (March), 9, no. 1, pp. 3-18.

BAUMAN, W. Scott; **MILLER**, Robert E. **(1994)** "Can managed portfolio performance be predicted?" *Journal of Portfolio Management* (Summer), vol. 20, no. 4, pp.31-40.

BAUMAN, W. Scott; **MILLER**, Robert E. **(1995)** "Portfolio performance rankings in stock market cycles." *Financial Analysts Journal* (March-April), vol. 51, no. 2, pp.79-87.

BASU, Sanjoy **(1977)** Investment performance of common stocks in relation to their price-earnings ratios: a test of the efficient market hypothesis." *Journal of Finance* (June), vol. XXXII, no.3, pp. 663-682.

BERRY, Michael A.; **BURMEISTER**, Edwin; **McELROY**, Marjorie B. **(1988)** "Sorting out risks using known APT factors." Financial Analysts Journal (March-April), vol.44, no. 2, pp.29-42.

BLACK, Fischer **(1993a)** "Beta and return." *Journal of Portfolio Management* (Fall), vol. 20, no. 1, pp.8-18.

BLACK, Fischer **(1993b)** "Estimating expected return." *Financial Analysts Journal* (September-October), vol. 49, no. 5, pp.36-38.

BLAKE, Christopher R.; **ELTON**, Edwin J.; **GRUBER**, Martin J. **(1993)** "The performance of bond mutual funds." *Journal of Business*, vol. 66, no. 3, pp. 371-403.

BOGLE, John C. **(1992)** "Selecting equity mutual funds." *Journal of Portfolio Management* (Winter), vol. 18, no. 2, pp.94-100.

BOWER, Richard S.; **WIPPERN**, Ronald F. **(1969)** "Risk-return measurement in portfolio selection and performance appraisal models: progress report." *Journal of Financial and Quantitative Analysis*, Vol. IV, pp. 417-447.

BREEN, W.; **JAGANATHAN**, R.; **OFER**, A. **(1986)** "Correcting for heteroscedasticity in tests for market timing ability." *Journal of Business*, 59, pp. 585-598.

BREUSCH, T. **(1978)** "Testing for autocorrelation in dynamic linear models." *Australian Economic Papers*, vol. 17, pp. 334-355.

BRINSON, Gary P.; **HOOD**, L. Randolph ; **BEEBOWER**, Gilbert L. **(1986)** "Determinants of portfolio performance." *Financial Analysts Journal* (July-August), pp.39-44.

BRINSON, Gary P.; **HOOD**, L. Randolph ; **BEEBOWER**, Gilbert L. **(1991)** "Determinants of portfolio performance II: an update." *Financial Analysts Journal* (May-June), pp.40-48.

BROWN, Keith C.; **BROWN**, Gregory D. **(1987)** "Does the composition of the market portfolio really matter?" *Journal of Portfolio Management* (Winter), pp. 26-32.

BROWN, Stephen J.; **GOETZMANN**, William N. **(1995)** "Performance persistence." *Journal of Finance* (June), vol. L, no. 2, pp.679-698.

BROWN, Stephen J.; **GOETZMANN**, William N. **(1997)** "Mutual fund styles." *Journal of Financial Economics*, vol.43, no. 3, pp.373-399.

BROWN, Stephen J.; **GOETZMANN**, William; **IBBOTSON**, Roger G. **(1998)** "Offshore hedge funds: survival and performance 1989-1995" NBER Working Paper, 37 pages.

BROWN, Stephen J.; **GOETZMANN**, William; **IBBOTSON**, Roger G.; **ROSS**, Stephen A. **(1992)** "Survivorship bias in performance studies." *Review of Financial Studies*, vol. 5, no. 4, pp.553-580.

CAPAUL, Carlo.; **ROWLEY**, Ian; **SHARPE**, William F. **(1993)** "International value and growth stock returns." *Financial Analysts Journal* (January-February), vol. 49, no.1, pp. 27-36.

CARHART, Mark M.**(1995)** "Survivor bias and mutual fund performance." Working Paper 95-23 (September), University of Southern California, Los Angeles, 36 pages.

CARHART, Mark M.**(1997)** "On persistence in mutual fund performance." *Journal of Finance* (March), vol. 52, no. 1, pp.57-83.

CHAN, Louis K.C.; **LAKONISHOK**, Josef **(1993)** "Are the reports of beta's death premature?" *Journal of Portfolio Management* (Summer), vol.19, no.4, pp.51-62.

CHANG, Eric C.; **LEWELLEN**, Wilbur G. **(1984)** "Market timing and mutual fund performance." *Journal of Business*, vol. 57, no. 1, pp.57-72.

CHEN, Carl R.; **STOCKUM**, Steve (1986) "Selectivity, market timing, and random beta behaviour of mutual funds: a generalized model." *Journal of Financial Research* (Spring), vol. IX, no. 1, pp. 87-96.

CHEN, Nai-Fu **(1983)** "Some empirical tests of the theory of asset pricing." *Journal of Finance* (December), vol. XXXVIII, no. 5, pp. 1393-1414.

CHEN, Nai-Fu; **COPELAND**, Thomas E.; **MAYERS**, David **(1987)** "A comparison of single and multifactor portfolio performance methodologies." *Journal of Financial and Quantitative Analysis* (December), vol. 22, no. 4, pp.401-417.

CHEN, Nai-Fu; **ROLL**, Richard; **ROSS**, Stephen **(1986)** "Economic Forces and the stock market." *Journal of Business* (July), vol. 59, pp. 386-403.

CHEN, Son-Nan; **LEE**, Cheng F. **(1981)** "The sampling relationship between Sharpe's performance measure and its risk proxy: sample size, investment horizon and market conditions." *Management Science* (June), vol. 27, no. 6, pp.607-618.

CHRISTENSEN, Ronald **(1990)** *Log-linear models*, Springer-Verlag, New York, 409 pages.

CHRISTOPHERSON, Jon A. **(1995)** "Equity Style Classifications." *Journal of Portfolio Management* (Spring), vol. 21, no. 3, pp.32-43.

CHRISTOPHERSON, Jon A.; **FERSON**, Wayne E.; **GLASSMAN**, Debra A. **(1998)** "Conditioning manager alphas on economic information: another look at the persistence of performance." Review of Financial Studies (Spring), vol. 11, no. 1, pp. 111-142.

CHRISTOPHERSON, Jon A.; **TRITTIN**, D. **(1995)** "Equity style classification system." In *The Handbook of Equity Style Management*, edited by T.D. Coggin and F.J. Fabozzi, pp. 69-98.

CHRISTOPHERSON, Jon A.; **TURNER**, Andrew L. **(1991)** "Volatility and predictability of manager alpha." *Journal of Portfolio Management* (Fall), vol.18, no. 1, pp.5-12

CHRISTOPHERSON, Jon A.; **WILLIAMS**, C. Nola **(1995)** "Equity style: what is it and why it matters." In *The Handbook of Equity Style Management*, edited by T.D. Coggin and F.J. Fabozzi, pp. 1-22.

COGGIN, T. Daniel; **FABOZZI**, Frank J; **RAHMAN**, Shafiqur **(1993)** "The investment performance of US equity pension fund managers: an empirical investigation. *Journal of Finance* (July), vol. XLVIII, no. 3, pp.1039-1055.

CONNOR, Gregory; **KORAJCZYK**, Robert A. **(1986)** "Performance measurement with the Arbitrage Pricing theory: a new framework for analysis." *Journal of Financial Economics* 15, pp.374-394.

CONNOR, Gregory; **KORAJCZYK**, Robert A. **(1991)** "The attributes, behaviour, and performance of U.S. mutual funds." *Review of Finance and Accounting*, 1, pp.5-26.

CONNOR, Gregory; **KORAJCZYK**, Robert A. **(1995)** "The arbitrage pricing theory and multifactor models of asset returns." In *Handbooks in Operations Research and Management Science*, vol. 9, Finance, edited by R.A. Jarrow, V. Maksimovic and W.T. Ziemba, Elsevier Science, pp. 87-144.

CONOVER, W.J. **(1980)** *Practical nonparametric statistics*, Second Edition, John Wiley & sons, New York, 493 pp.

CORTEZ, Maria do Céu R.; **ARMADA**, Manuel J.R. **(1997)** "On mutual fund performance evaluation" *Estudos de Gestão*, vol. 3, no.3, pp.145-163.

CUMBY, Robert E.; **GLEN**, Jack D. **(1990)** "Evaluating the performance of International mutual funds." *Journal of Finance* (June), vol. XLV, no. 2, pp.497-521.

DHRYMES, Pheobus J.; **FRIEND**, Irwin; **GULTENKIN**, N. Bulent (1984) "A critical reexamination of the empirical evidence on the Arbitrage Pricing Theory." *Journal of Finance* (June), vol. XXXIX, no. 2, pp.323-346.

DiBARTOLOMEO, Dan; **WITKOWSKI**, Erik **(1997)** "Mutual fund misclassification: evidence based on style analysis." *Financial Analysts Journal* (September-October), vol. 53, no.5, pp.32-43.

DIVECHA, Arjun; **GRINOLD**, Richard C. **(1989)** "Normal portfolios: issues for sponsors, managers and consultants." *Financial Analysts Journal* (March-April), vol. 21, no.1, pp.7-13.

DUNN, Patricia C.; **THEISEN**, Rolf D. **(1983)** "How consistently do active managers win?" *Journal of Portfolio Management* (Summer), vol. 9, no. 4, pp.47-50.

DYBVIG, Philip H.; **ROSS**, Stephen A. **(1985a)** "Differential information and performance measurement using a security market line." *Journal of Finance* (June), vol. XL, no. 2, pp. 383-398.

DYBVIG, Philip H.; **ROSS**, Stephen A. **(1985b)** "The analytics of performance measurement using a security market line." *Journal of Finance* (June), vol. XL, no. 2, pp. 401-416.

EFRON, Bradley **(1979)** "Bootstrap methods: another look at the jacknife." *Annales of Statistics*, vol. 7, pp.1-26.

ELTON, Edwin J.; **GRUBER**, Martin J. **(1995)** *Modern Portfolio Theory and Investment Analysis*, Fifth Edition, John Wiley & Sons, Inc, 703 pages.

ELTON, Edwin J.; **GRUBER**, Martin J. **(1997)** "Modern portfolio theory, 1950 to date." Working Paper, 97-3 (March), New York University, L. Stern School of Business, 33 pages.

ELTON, Edwin J.; **GRUBER**, Martin J.; **BLAKE**, Christopher R. **(1996a)** "The persistence of risk-adjusted mutual fund performance." *Journal of Business*, vol. 69, no. 2, pp.133-157.

ELTON, Edwin J.; **GRUBER,** Martin J.; **BLAKE,** Christopher R. **(1996b)** "Common factors in mutual fund returns." Paper presented at the 23rd Annual Meeting organized by the European Finance Association, Oslo, 25 pages.

ELTON, Edwin J.; **GRUBER**, Martin J.; **BLAKE**, Christopher R. **(1996c)** "Survivorship bias and mutual fund performance." *Review of Financial Studies* (Winter), vol. 9, no. 4, pp.1097-1120.

ELTON, Edwin J.; **GRUBER**, Martin J.; **DAS**, Sanjiv; **HLAVKA**, Matthew **(1993)** "Efficiency with costly information: a reinterpretation of evidence from managed portfolios." *Review of Financial Studies*, vol.6, no. 1, pp.1-22.

ELTON, Edwin J.; **GRUBER**, Martin J.; **RENTZLER**, Joel **(1990)** "The performance of publicly offered commodity funds." *Financial Analysts Journal* (July-August), vol. 46, no. 4, pp.23-30.

EVERITT, B.S. **(1992)** *The analysis of contingency tables*, Second Edition, Chapman & Hall, london, 164 pages.

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

FABOZZI, Frank J.; **FRANCIS**, Jack C. **(1979)** "Mutual fund systematic risk for bull and bear markets: an empirical investigation." *Journal of Finance* (December), vol. XXXIV, no. 5, pp. 1243-1250.

FABOZZI, Frank J.; **FRANCIS**, Jack C. **(1980)** "Stability of mutual fund systematic risk statistic." *Journal of Business Research* (June), vol. 8, no. 2, pp. 263-275.

FAMA. Eugene F. **(1972)** "Components of investment performance." *Journal of Finance* (June), vol. XXVII, no. 3, pp. 551-567.

FAMA. Eugene F. **(1991)** "Efficient capital markets II." *Journal of Finance* (December), vol. XLVI , no. 5, pp. 1575-1618.

FAMA, Eugene F.; **FRENCH**, Kenneth R. **(1992)** "The cross-section of expected stock returns." *Journal of Finance* (June), vol. XLVII, no. 2, pp.427-465.

FAMA, Eugene F.; **FRENCH**, Kenneth R. **(1993)** "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics*, vol. 33, pp.3-56.

FAMA, Eugene F.; **FRENCH**, Kenneth R. **(1995)** "Size and book-to-market factors in earnings and returns." *Journal of Finance* (March), vol. L, no. 1, pp.131-155.

FERGUSON, Robert **(1980)** "Performance measurement doesn't make sense." *Financial Analysts Journal* (May-June), pp.59-69.

FERGUSON, Robert **(1986)** "The trouble with performance measurement." *Journal of Portfolio Management* (Spring), pp.4-9.

FERSON, Wayne E.; **SCHADT**, Rudi W. **(1996)** "Measuring fund strategy and performance in changing economic conditions." Journal of Finance (June), vol. 51, no. 2, pp..425-561.

FIELITZ, Bruce D.; **GREENE**, Myron T. **(1980)** "Shortcomings in portfolio evaluation via MPT." *Journal of Portfolio Management* (Summer), pp.13-19.

FLETCHER, Jonathan **(1995)** "An examination of the selectivity and market timing performance of UK unit trusts." *Journal of Finance and Accounting* (January), vol. 22, no. 1, pp-143-156.

FRIEND, Irwin; **BLUME**, Marshall **(1970)** "Measurement of portfolio performance under uncertainty." *American Economic Review* (September), vol. LX, pp. 561-575.

GODFREY, L. **(1978)** "Testing against general autoregressive and moving average error models when regressors include lagged dependent variables." *Econometrica*, vol. 48, pp. 1293-1302.

GOETZMANN, William N.; **IBBOTSON**, Roger G. **(1994)** "Do winners repeat? Patterns in mutual fund performance." *Journal of Portfolio Management* (Winter), vol. 20, no. 2, pp.9-18.

GOETZMANN, William N.; **PELES**, Nadav **(1997)** "Cognitive dissonance and mutual fund investors." *Journal of Financial Research* (Summer), vol. 20, no. 2, pp.145-158.

GRANT, Dwight **(1977)** "Portfolio performance and the "cost" of timing decisions." *Journal of Finance* (June), vol. XXXII), no. 3, pp. 837-845.

GREENE, William H. **(1993)** *Econometric analysis*, Second Edition, Prentice-Hall, Englewood Cliffs, N.J., 791 pages.

GRINBLATT, Mark; **TITMAN**, Sheridan **(1988)** "The evaluation of mutual fund performance: an analysis of monthly returns." Working Paper, University of California - UCLA.

GRINBLATT, Mark; **TITMAN**, Sheridan **(1989a)** "Mutual fund performance: an analysis of quarterly portfolio holdings." *Journal of Business* (July), vol. 62, no. 3, pp. 393-416.

GRINBLATT, Mark; **TITMAN**, Sheridan **(1989b)** "Portfolio performance evaluation: old issues and new insights." *Review of Financial Studies*, vol. 2, no. 3, pp.393-421.

GRINBLATT, Mark; **TITMAN**, Sheridan **(1992)** "The persistence of mutual fund performance." *Journal of Finance* (December), vol. 47, no. 5, pp. 1977-1984.

GRINBLATT, Mark; **TITMAN**, Sheridan **(1993)** "Performance measurement without benchmarks: an examination of mutual fund returns." *Journal of Business*, vol. 66, no. 1, pp.47-68.

GRINBLATT, Mark; **TITMAN**, Sheridan **(1994)** "A study of monthly fund returns and performance evaluation techniques." *Journal of Financial and Quantitative Analysis* (September), vol. 29, no. 3, pp.419-444.

GROSSMAN S.; **STIGLITZ**, J. **(1980)** "On the informationally efficient markets." *American Economic Review*, vol. 70, pp. 393-408.

GRUBER, Martin J. **(1996)** "Another puzzle: the growth in actively managed mutual funds." *Journal of Finance* (July), vol. LI, no. 3, pp.783-810.

GRUNDY, Kevin; **MALKIEL**, Burton G. **(1996)** "Reports on beta's death have been greatly exaggerated." *Journal of Portfolio Management* (Spring), vol. 22, no. 3, pp. 36-44.

GUNTHORPE, Deborah; **LEVY**, Haim **(1994)** "Portfolio composition and the investment horizon." *Financial Analysts Journal* (January-February), vol. 50, no. 1, pp.51-56.

HAMMER, David A. (1991) *Dynamic Asset Allocation: strategies for the stock, bond, and money markets*, John Wiley & Sons, 336 pages.

HANDA, Puneet; **KOTHARI**, S.P.; **WASLEY**, Charles **(1993)** "Sensitivity of multivariate tests of the Capital Asset Pricing Model to the return measurement interval." *Journal of Finance* (September), vol. XLVIII, no. 4, pp. 1543-1551.

HARLOW, W.V. **(1991)** "Asset allocation in a downside-risk framework." *Financial Analysts Journal* (September-October), vol. 47, no. 5, pp.28-40.

HENDRICKS, Darryll; **PATEL**, Jayendu; **ZECKHAUSER**, Richard **(1993)** "Hot hands in mutual funds: short-run persistence of relative performance." *Journal of Finance* (March), vol. XLVIII, no. 1, pp.93-130.

HENRIKSSON, Roy D.; **MERTON**, Robert 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.

HENRIKSSON, Roy D. **(1984)** "Market timing and mutual fund performance: an empirical investigation." *Journal of Business* (January), vol. 57, no. 1, pp. 73-96.

HENSEL, Chris R.; **EZRA**, D. Don; **ILKIW**, John H. **(1991)** "The importance of the asset allocation decision." *Financial Analysts Journal* (July-August), vol. 47, no. 4, pp 65-72.

IPPOLITO, Richard A. **(1989)** "Efficiency with costly information: a study of mutual fund performance, 1965-1984." *Quarterly Journal of Economics* (February), vol. 194, no. 1, pp.1-23.

JAGADEESH, Narasimhan; **TITMAN**, Sheridan **(1993)** "Returns to buying winners and selling losers: implications for stock market efficiency." *Journal of Finance* (March), vol. XLVIII, no. 1, pp. 65-91.

JENSEN, Michael C. **(1968)** "The performance of mutual funds in the period 1945-1964." *Journal of Finance* (May), vol. XXIII, pp.389-416.

JENSEN, Michael C. **(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.

JOHNSTON, John **(1987)** *Econometric methods*, McGraw-Hill, Singapore, 568 pages.

KAHN, Ronald N.; **RUDD**, Andrew **(1995)** "Does historical performance predict future performance?" *Financial Analysts Journal* (November-December), vol. 51, no. 6, pp.43-52.

KLEMKOSKY, Robert C. **(1973)** "The bias in composite performance measures." *Journal of Financial and Quantitative Analysis* (June), vol. VIII, pp.505-514.

KLEMKOSKY, Robert C. **(1977)** "How consistently do managers manage?" *Journal of Portfolio Management* (Winter), pp. 11-15.

KLEMKOSKY, Robert C.; **MANESS**, Terry S. **(1978)** "The predictability of real portfolio risk levels." *Journal of Finance* (May), vol. XXXIII, pp. 631-639.

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

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

KON, Stanley J.; **JEN**, Frank C. **(1979)** "The investment performance of mutual funds: an empirical investigation of timing, selectivity and market efficiency." *Journal of Business* (April), vol. 52, no. 2, pp. 263-289.

KOTHARI, S. P.; **SHANKEN**, Jay; **SLOAN**, Richard G. **(1995)** "Another look at the cross-section of expected returns." *Journal of Finance* (March), vol. L, no. 1, pp.185-224.

KRITZMAN, Mark **(1983)** "Can bond managers perform consistently?" *Journal of Portfolio Management*, vol. 9, pp. 54-56.

KRITZMAN, Mark **(1987)** "How to build a normal portfolio in three easy steps." *Journal of Portfolio Management* (Summer), pp.21-23.

LAKONISHOK, Josef; **SHLEIFER**, Andrei; **VISHNY**, Robert W. **(1992)** "The structure and performance of the money management industry." *Brookings Papers on Economic Activity: Microeconomics*, pp.339-379.

LEE, Cheng F.; **JEN**, Frank C. **(1978)** "Effects of measurement errors on systematic risk an performance measure of a portfolio." *Journal of Financial and Quantitative Analysis* (June), p. 299-312.

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

LEHMANN, Bruce L.; **MODEST**, David M. **(1987)** "Mutual fund performance evaluation: a comparison of benchmarks and benchmark comparison." *Journal of Finance* (June), vol. XLII, no. 2, pp. 233-265.

LEHMANN, Bruce L.; **MODEST**, David M. **(1988)** "The empirical foundations of the Arbitrage Pricing Theory." *Journal of Financial Economics*, vol. 21, pp.213-254.

LEVY, Haim **(1981)** "The CAPM and the investment horizon." *Journal of Portfolio Management* (Winter), pp.32-40.

LEVY, Haim **(1984)** "Measuring risk and performance over alternative investment horizons." *Financial Analysts Journal* (March-April), pp. 61-68.

LINTNER, John **(1965)** "The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets." *Review of Economics and Statistics* (February), vol. XLVII, no. 1, pp.13-37.

LUCK, Christopher G. **(1995)** "Style in indexes and benchmarks." In *Performance Evaluation, Benchmarks, and Attribution Analysis*, ICFA Continuing Education (June), Association for Investment Management and Research, pp. 42-47.

MACHADO-SANTOS, Carlos M.; **ARMADA**, Manuel J.R. **(1996)** "Assimetria de Informação e Avaliação da Performance de Gestores de Investimentos." *Revista de Administração Contemporânea (RAC-Brasil)*, vol. 1, no. 1 (Janeiro-Abril), pp.121-143.

MALKIEL, Burton G. **(1995)** "Returns from investing in equity mutual funds 1971 to 1991." *Journal of Finance* (June)., vol. L, no. 2, pp.549-572.

MARKOWITZ, Harry M. **(1952)** "Portfolio selection." *Journal of Finance* (March), vol. VIII, no. 1, pp.77-91.

MARKOWITZ, Harry M. **(1959)** *Portfolio selection: efficient diversification of investment*, John Wiley & Sons Inc., New York.

MARMER, Harry S.; **NG**, F.K. Louis **(1993)** "Mean-semivariance analysis of option-based strategies: a total asset mix perspective." *Financial Analysts Journal* (May-June), vol. 49, no. 3, pp.47-54.

MAYERS, David; **RICE**, Edward M. **(1979)** "Measuring portfolio performance and the empirical content of asset pricing models." *Journal of Financial Economics* (March), 7, pp. 3-28.

MERRIKEN, Harry E. **(1994)** Analytical approaches to limit downside risk: the semivariance and the need for liquidity." *Journal of Investing* (Fall), vol. 3, no. 3, pp. 65-72.

MERTON, Robert C. **(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.

MOOD, Alexander M.; **GRAYBILL**, Franklin A.; **BOES**, Duane C. **(1974)** *Introduction to the theory of statistics*, McGraw-Hill New York, 564 pages.

MOSES, Edward A.; **CHENEY**, John M. **(1989)** *Investments: Analysis, Selection and Management*, West Publishing Company, 788 pages.

MOSSIN, Jan **(1966)** "Equilibrium in a capital asset market." *Econometrica* (October), vol. 34, no. 4, pp.768-783.

MURTEIRA, Bento J.F. **(1990)** *Probabilidades e Estatística*, vol. II, second edition, McGraw-Hill de Portugal, 480 pages.

NEWBOLD, Paul **(1991)** *Statistics for business and economics*, Third Edition, Prentice-Hall international, 930 pages.

NEWEY, Whitney K.; **WEST**, Kenneth **(1987)** "A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix." *Econometrica*, vol. 55, no. 3, pp. 703-708.

NOREEN, Eric W. **(1989)** *Computer-intensive methods for testing hypothesis: an introduction*, John Wiley & Sons, New York, 229 pages.

PESARAN, M. Hashem; **TIMMERMANN**, Allan G. **(1992a)** "A generalization of the non-parametric Henriksson-Merton test of market timing." Working Paper, (July) Birkbeck College, University of London, 12 pages.

PESARAN, M. Hashem; **TIMMERMANN**, Allan G. **(1992b)** "A simple nonparametric test of predictive performance." *Journal of Business and Economic Statistics* (October), vol. 10, no. 4, pp. 461-465.

PFLEIDERER, Paul; **BHATTACHARYA**, Sudipto **(1983)** "A note on performance evaluation." Technical report no. 714 (October), Stanford University, Graduate School of Business, 23 pages.

PHELPS, Shawn; **DETZEL**, Larry **(1997)** "The nonpersistence of mutual fund performance." *Quarterly Journal of Business and Economics* (Spring), vol. 36, no. 2, pp.55-69.

RENNIE, Edward P.; **COWHEY**, Thomas J. **(1990)** "The successful use of benchmark portfolios: a case study." *Financial Analysts Journal* (September-October), vol. 46, no. 5, pp.18-26.

ROLL, Richard **(1977)** "A critique of the asset pricing theory's tests. Part I: on past and potential testability of the theory." *Journal of Financial Economics* (March), vol. 4, no. 2, pp.129-176.

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

ROLL, Richard **(1979)** "A reply to Mayers and Rice (1979)." *Journal of Financial Economics*, 7, pp.391-400.

ROLL, Richard **(1980)** "Performance evaluation and benchmark errors (I)." *Journal of Portfolio Management* (Summer), 6, pp.5-12.

ROLL, Richard **(1981)** "Performance evaluation and benchmark errors (II)." *Journal of Portfolio Management* (Winter), pp.17-22.

ROLL, Richard **(1995)** "Style return differentials: illusions, risk premiums, or investment opportunities?" In *The Handbook of Equity Style Management*, edited by T.D. Coggin and F.J. Fabozzi, pp.99-128.

ROLL, Richard; **ROSS**, Stephen A. **(1994)** "On the cross-sectional relation between expected returns and betas." *Journal of Finance* (March), vol. XLIX, no.1, pp.101-121.

ROM, Brian M.; **FERGUSON**, Kathleen W. **(1994a)** "Post-modern portfolio theory comes of age." *Journal of Investing* (Fall), vol. 3, no. 3, pp.11-17.

ROM, Brian M.; **FERGUSON**, Kathleen W. **(1994b)** "Portfolio theory is alive and well: a response." *Journal of Investing* (Fall), vol. 3, no. 3, pp.24-44.

ROSS, Stephen A. **(1976)** "The arbitrage theory of capital asset pricing." *Journal of Economic Theory* (December), vol. 13, pp. 341-360.

ROSS, Stephen A. **(1977)** "Risk, return and arbitrage." In *Risk and Return in Finance*, Friend & Bicksler Ed., Cambridge, MA, pp.189-218.

SHARPE, William F. **(1964)** "Capital asset prices: a theory of market equilibrium under conditions of risk." *Journal of Finance* (September), vol. XIX, no. 3, pp. 425-444.

SHARPE, William F. **(1966)** "Mutual fund performance." *Journal of Business* (January), vol. 39, no. 1, Part II, pp.119-138.

SHARPE, William F. **(1984)** "Factor models, CAPMs, and the ABT."(*sic*) *Journal of Portfolio Management* (Fall), pp. 21-25.

SHARPE, William F. **(1988)** "Determining a fund's effective asset mix." *Investment Management Review* (December), pp. 59-69.

SHARPE, William F. **(1992)** "Asset allocation: Management style and performance measurement. *Journal of Portfolio Management* (Winter), vol. 18, pp. 7-19.

SHARPE, William F. **(1994)** "The Sharpe ratio." *Journal of Portfolio Management* (Fall), vol. 21, no. 1, pp.49-58.

SHARPE, William F. **(1996)** "The styles and performance of large seasoned U.S. mutual funds." *World Wide Web* http://gsb-www.stanford.edu/~wfsharpe/ls100.htm.

SHARPE, William F.; **ALEXANDER**, Gordon J.; **BAILEY**, Jeffrey V. **(1995)** *Investments*, Fifth Edition, Prentice Hall International, 1058 pages.

SHUKLA, Ravi; **TRZINKA**, Charles **(1991)** "Research on risk and return: can measures of return explain anything?" *Journal of Portfolio Management* (Spring), vol.17, no.3, pp.15-21.

SHUKLA, Ravi; **TRZCINKA**, Charles **(1994)** "Persistence performance in the mutual fund market: tests with funds and investment advisors." *Review of Quantitative Finance and Accounting*, 4, pp.114-135.

SIEGEL, Sydney **(1975)** *Estatística não paramétrica*, McGraw-Hill do Brasil, São Paulo, 350 pages.

SINGER, Brian D.; **KARNOSKY**, Dennis S. **(1995)** "The global framework for global investment management and performance attribution." *Journal of Portfolio Management* (Winter), vol. 21, no. 2, pp.84-92.

SMITH, Keith V.; **TITO**, Dennis A. (1969) "Risk return measures of ex-post portfolio performance." *Journal of Financial and Quantitative Analysis* (December), vol. 4, pp. 449-471.

TIERNEY, David E.; **BAILEY**, Jeffery V. **(1995)** "Benchmark orthogonality properties." Journal of Portfolio Management (Spring), vol. 21, no. 3, pp.27-31.

TIERNEY, David E.; **WINSTON**, Kenneth **(1990)** "Defining and using dynamic completeness to enhance total fund efficiency." *Financial Analysts Journal* (July-August), vol. 46, no. 4, pp.49-54.

TIERNEY, David E.; **WINSTON**, Kenneth **(1991)** "Using generic benchmarks to present manager styles." *Journal of Portfolio Management* (Summer), vol. 17, no. 2, pp.33-36.

TREYNOR, Jack L. **(1965)** "How to rate management of investment funds." *Harvard Business Review* (January-February), vol. 43, no. 1, pp.63-75.

TREYNOR, Jack L.; **MAZUY**, Kay K. **(1966)** "Can mutual funds outguess the market?" *Harvard Business Review* (July-August), vol. 44, pp. 131-136.

TOBIN, James **(1958)** "Liquidity preferences or behaviour towards risk." *Review of Economic Studies* (February), pp.65-85.

TRZCINKA, Charles A. **(1995a)** "Equity style classifications: comment." *Journal of Portfolio Management* (Spring), vol.21, no.3, pp.44-46.

TRZCINKA, Charles A. **(1995b)** "Measuring and monitoring equity investment style: is it worth the effort?" In *The Handbook of Equity Style Management*, edited by T.D. Coggin and F.J. Fabozzi, pp. 241-253.

VOLKMAN, David A.; **WOHAR**, Mark E. **(1995)** "Determinants of persistence in relative performance of mutual funds." *Journal of Financial Research* (Winter), vol. XVIII, no. 4, pp.415-430.

VOLKMAN, David A.; **WOHAR**, Mark E. **(1996)** "Abnormal profits and relative strength in mutual fund returns." *Review of Financial Economics*, vol. 5, no.2, pp.101-116.

WERMERS, Russ **(1997)** "Momentum investment strategies of mutual funds, performance persistence and survivorship bias." Working Paper (March), Graduate School of Business and Administration, University of Colorado at Boulder, 46 pages.

WHITE, Halberth **(1980)** "A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity." *Econometrica*, vol. 48, no. 4, pp. 149-170.

WILSON, Jack W.; **JONES**, Charles P. **(1981)** "The relationship between performance and risk: whence the bias?" *Journal of Financial Research* (Summer), vol. IV, no. 2, pp.103-107.

ZIMMERMANN, Heinz; **ZOGG-WETTER**, Claudia **(1992)** "On detecting selection and timing ability: the case of stock market indexes." *Financial Analysts Journal* (January-February), vol. 48, no. 1, pp.80-83.