LAPPEENRANTA UNIVERSITY OF TECHNOLOGY

School of Business

Finance

PERFORMANCE PERSISTENCE OF EQUITY FUNDS IN THE RUSSIAN STOCK MARKET

Examiners: D. Sc. Eero Pätäri

Professor Mika Vaihekoski

Lappeenranta, 20.11.2008

Mika Rossi

ABSTRACT

Author: Mika Rossi

Title: Performance persistence of equity funds in the Russian stock market

Faculty: LUT, School of Business

Major: Finance

Year: 2008

Master's Thesis: Lappeenranta University of Technology. 86 pages, 2 figures, 12 tables, 6 appendices.

Examiners: D. Sc. Eero Pätäri
Professor Mika Vaihekoski

Keywords: Russia, equity funds, performance, persistence, Bayesian estimation

This thesis investigates performance persistence among the equity funds investing in Russia during 2003-2007. Fund performance is measured using several methods including the Jensen alpha, the Fama-French 3-factor alpha, the Sharpe ratio and two of its variations. Moreover, we apply the Bayesian shrinkage estimation in performance measurement and evaluate its usefulness compared with the OLS 3-factor alphas. The pattern of performance persistence is analyzed using the Spearman rank correlation test, cross-sectional regression analysis and stacked return time series.

Empirical results indicate that the Bayesian shrinkage estimates may provide better and more accurate estimates of fund performance compared with the OLS 3-factor alphas. Secondly, based on the results it seems that the degree of performance persistence is strongly related to length of the observation period. For the full sample period the results show strong signs of performance reversal whereas for the subperiod analysis the results indicate performance persistence during the most recent years.

TIIVISTELMÄ

Tekijä:	Mika Rossi
Tutkielman nimi:	Osakerahastojen menestyksen pysyvyys Venäjän osakemarkkinoilla
Osasto:	Kauppatieteellinen tiedekunta
Pääaine:	Rahoitus
Vuosi:	2008
Pro gradu -tutkielma:	Lappeenrannan teknillinen yliopisto. 86 sivua, 2 kuvaa, 12 taulukkoa, 6 liitettä
Tarkastajat:	KTT Eero Pätäri Professori Mika Vaihekoski
Hakusanat:	Venäjä, osakerahastot, menestyksen mittaus, menestyksen pysyvyys, Bayesilainen estimointi

Tämä pro gradu -tutkielma tutkii Venäjän osakemarkkinoille sijoittavien osakerahastojen menestyksen pysyvyyttä vuosina 2003-2007. Rahastojen menestystä mitataan useilla eri mittareilla; Jensenin alfalla, Fama-French 3-faktori alfalla, Sharpen mittarilla sekä kahdella Sharpen mittarin variaatiolla. Lisäksi menestyksen mittauksessa käytetään Bayesilaista estimointia sekä arvioidaan sen hyödyllisyyttä ja ennustetarkkuutta suhteessa pienimmän neliösumman- menetelmän estimaatteihin. Menestyksen pysyvyyttä tutkitaan Spearmanin järjestyskorrelaatiotestillä, poikkileikkausregressiolla sekä ns. yhdistetyn tuottoaikasarjan menetelmällä.

Tulokset osoittavat, että Bayesilaiset alfat ennustavat rahastojen menestystä hieman tarkemmin kuin pienimmän neliösumman menetelmään perustuvat 3-faktori alfat. Lisäksi tulokset osoittavat, että menestyksen pysyvyys on vahvasti sidoksissa kulloinkin käytettävään tarkasteluajanjakson pituuteen. Kun testeissä käytetään havaintoja koko tarkasteluajanjaksolta, tulokset osoittavat vahvaa ns. käänteistä pysyvyyttä. Toisaalta, kun aineisto jaetaan alaperioideihin, tulokset indikoivat menestyksen pysyvyyttä viimeisten vuosien aikana.

ACKNOWLEDGEMENTS

Now my studies are ending and it is time for new challenges. However, I would not have succeeded alone. First, regarding to my thesis I want to thank my excellent supervisor Mr. Eero Pätäri for the guidance and good advice throughout the writing process. Moreover, I want to thank my friends for the great time I have spent here in Lappeenranta.

Finally, I want to thank my wonderful parents, my sister and my other family members for supporting me through my studies and helping me through the hard times.

Lappeenranta, 20.11.2008

Mika Rossi

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Background

Measuring persistence in fund performance has been the goal of several academic studies over the last decades. Performance persistence refers to the ability of a fund to maintain its performance ranking against a specific benchmark or against some fund over time. Therefore, persistency in performance is sometimes related to superior stock picking skills or market timing skills of a fund manager and it is sometimes called as “hot hands” phenomena.[1] On the other hand, good track record of a portfolio manager is often used in marketing funds to investors. Obviously, it would be rather difficult to sell a mutual fund with a poor performance record to the public.

One issue closely related to performance persistence evaluation is the significance of the past information set in predicting the future performance. According to the efficient market hypothesis it should not be possible to predict future performance of any security using past price data after adjusting for the risk and other pricing factors. Therefore, studies on performance persistence and future performance prediction using the past information set are partly tests of stock market efficiency.

From investor’s point of view, if fund returns show a pattern of predictable behavior, active selection among mutual funds could be a profitable strategy. More specifically, if an investor is able to identify e.g. using past information the funds that will be superior performers in the future, the expected return on his or her portfolio can be increased. On the other hand, if past performance does not contain any information about the future, all the data processing and performance measurement would be a useless procedure for an investor. (ter Horst and Verbeek, 2000)

[1] Usually this expression is used to describe short-term performance persistence. See e.g. Droms (2006) and Hendricks et al. (1993).

A great number of papers have studied the performance persistence of mutual funds. For example, papers devoted to the US equity funds are authored by Hendricks, Zeckhauser and Patel (1993), Grinblatt and Titman (1994), Elton et al. (1996), Carhart (1997) and Deaves (2004). As far as European markets are considered, persistence phenomenon has been widely studied among UK mutual funds e.g. by Fletcher (1998) and Allen and Tan (1999). Moreover, quite recently studies have been conducted for example, by Busse and Irvine (2006) and Huij and Verbeek (2007) using more innovative performance measures. Despite the extensive analysis, the results of the previous studies reveal that it is difficult to make any unanimous conclusion whether performance persists or not. A number of studies have been published both for and against the prediction power of return history. On the other hand, the authors seem to be quite unanimous that if persistence exists, it is rather a short-term phenomenon.

However, despite the numerous papers devoted to examining the pattern of performance persistence, most of them seem to concentrate on the US equity markets or on other developed markets such as the UK. Due to limited data available or for some other reason, there seems to be significant lack of performance persistence studies of mutual funds concentrating in emerging equity markets. To our knowledge, the only paper that investigates emerging market funds is conducted by Huij and Post (2008). However, there has been outstanding growth of emerging market funds over the last years and although some of the emerging markets have undergone remarkable recent development, mutual funds may still offer the best and the easiest way for an individual investor to invest in these markets.

In general, investing in the emerging markets has proven to be attractive to investors' since many emerging economies have experienced rapid growth and hence offered considerable opportunities for high returns. For example, according to Kauppalehti (2008) the equity funds investing in

Russia have profited around 153 percent during 2005-2007 being the most profitable emerging market fund category.[2] Besides that, emerging markets have provided greater scope for investors' portfolio risk reduction compared with the one that can be achieved by developed markets alone. These diversification benefits stem from the low correlations with the developed markets. (Sharpe et al. 1999, 880)

Due to the great attractiveness of these markets among the investors, we consider it important to conduct studies related to emerging markets as well. Therefore, this paper adds new empirical evidence to the existing literature by providing a new and an interesting insight of mutual fund performance persistence evaluation.[3]

1.2 Objects, limitations and methodologies of the study

This thesis will analyze the equity funds investing in the Russian stock market. To our best knowledge, this persistence study is the first one concentrating on the funds investing in Russia and one of the very first studies concentrating on the emerging market funds at all. Hence, we consider that our study would have great novelty value to the performance persistence literature.

Firstly, the general objective of this study is to fill the existing gap in the financial literature between studies concentrating on the developed and on the emerging mutual fund markets. Secondly, the empirical objective of this thesis is to provide evidence if the equity funds investing in Russia exhibit relative performance persistence over a period from 2003 to 2007. The study is made from the European investor's point of view since the

[2] Kauppalehti is a Finnish newspaper concentrating on business.
[3] Interestingly, Sandvall (1999) suggests that the performance persistence mutual funds may be stronger and more evident on the emerging markets due to a potential "first mover" advantage.

sample consists of European equity funds and all the data needed in this study is quoted in euros.

In this study we explore the risk adjusted returns. To prevent spurious results arising due to a model misspecification, fund performance is evaluated using several performance metrics. These include the Sharpe ratio, downside deviation based Sharpe ratio, the modified Sharpe ratio, the Jensen alpha and the three-factor model developed by Fama and French. Besides the traditional performance metrics presented above, we also employ the Bayesian shrinkage estimation in analyzing the fund performance. The Bayesian estimation has gained ground among the academics since it tries to exploit prior information contained in the group of mutual fund returns. Using this method, we calculate the Bayesian alphas for each fund and then compare them with the standard frequentist Fama-French three-factor alphas estimated through the ordinary least squares regression (OLS). Our main objective in employing the Bayesian method is to study whether the Bayesian alphas could provide better and more accurate estimates of fund performance than the traditional OLS estimates and on the other hand, how the Bayesian alphas detect performance persistence.

To study performance persistence we also employ several methods. First, we start with the Spearman correlation test to investigate if the fund rankings form the selection period correlate with the ones from the holding period. Second, we employ a cross-sectional regression analysis in order to detect whether the performance metrics from the ranking period explain those from the holding period. Third, we apply so-called stacked return time series analysis. We form top and bottom portfolios based on the ranking period performance and compare their performance in the following period in order to study whether the performance difference between top and bottom performers remains.

1.3 Structure of the study

This study is organised into six sections and the remainder of this thesis is structured as follows: section 2 provides the theoretical background for this thesis. Section 3 presents the previous literature related to performance persistence of mutual funds. The fourth section describes the data and the methodology applied in this study. Section 5 introduces the empirical results. Finally, in the sixth section we conclude this thesis and suggest a couple of ways to further extend this study.

2 THEORETICAL BACKGROUND

2.1 The Efficient market hypothesis

As mentioned before, one issue closely related to mutual fund performance persistence evaluation is the employment of the past information set in predicting the performance in the future. Therefore, the persistence studies are partly tests of the stock market efficiency. In the classic article, Fama (1970) suggests that on an efficient market at any given time, all securities fully reflect all available information. More specifically, he states that it is not consistently possible to beat the market by using the information that is already known. Therefore, when the prices fully contain all the information, they only change in response to new information, which must be something unpredictable. This makes securities prices to move unpredictably. In finance this movement is often referred to as random walk process (Bodie et al. 2005, 370-371).

Fama (1970) subdivides the efficient market hypothesis into three categories, each of them dealing with a different type of information. In the first category security prices reflect all the information contained in the record of past security prices. This is called weak form of market efficiency. If a market meets the weak form criteria, it is not possible to make superior profits by studying the past returns. Therefore, according to the weak form criteria it should not be possible to use e.g. historical fund returns to predict the fund performance in the future and make superior profits.

The second form of efficiency states that security prices reflect both the past information and all the other published information. This form is better known as the semi-strong form of market efficiency. If markets meet the semi strong criteria, then the prices will immediately adjust for public announcements such as the announcement of the last quarter’s earnings.

Finally, the third form of market efficiency is called strong form efficiency. This means that the prices reflect both public and private information of the certain security. Therefore, not even insider information could be used to gain superior profits.

2.2 The Capital Asset Pricing Model

The Capital Asset Pricing Model (CAPM), which is the standard form of the general relationship for asset return and risk was developed independently by Sharpe (1964), Lintner (1965) and Mossin (1966). All three authors make a similar conclusion about the equilibrium model that determines the relationship between the expected return and risk for any asset. The basic idea behind the CAPM is that the expected returns on securities are a positive linear function of their market risk. The model can be given as follows:

$$E(r_i) = r_f + \beta_i [E(r_m) - r_f] \tag{1}$$

where $E(r_i)$ is the expected return for asset i, r_f is the return of the risk-free asset, β_i stands for the beta coefficient for security i and $E(r_m)$ is the expected return for the market portfolio.

The risk-free asset is considered as a certain return. Therefore, this type of asset must be some kind of fixed income security with no possibility of default. Generally accepted proxy for the risk-free asset is Treasury security with a maturity that matches the length of the investor's holding period. (Sharpe et al. 1999:204-205)

The beta coefficient measures the security's sensitiveness to the changes in return of the market portfolio. It assumes that any additional variables such as price ratio or the firm size do not have an effect on expected excess return. Therefore, it is the index of systematic risk. The higher the beta is for any security, the higher the equilibrium returns is expected to be. On the other hand, higher beta coefficient would mean higher losses when the market is going down. The beta coefficient can be calculated as follows (Elton et al. 2003):

$$\beta_i = \frac{\operatorname{cov}(r_i, r_m)}{\operatorname{var}(r_m)} \tag{2}$$

where $cov(r_i,r_m)$ is the covariance of market return and return on investment and $var(r_m)$ is the variance of market return.

When it comes to the market portfolio, Sharpe et al. (1999, 232) suggest that it does not only consists of common stocks but also of other kind of investments such as real estate, bonds and preferred stocks. However, generally investors restrict the market portfolio to just common stocks. Actually, the definition of the true market portfolio has been a controversial topic among academics for years. For example, Roll (1977) argues that the true market portfolio is difficult to determine. According to him, this means that therefore it is not possible to test the Capital Asset Pricing Model. Furthermore, Roll (1977) claims that the employment of different proxies for the market portfolio may cause some measurement errors. For example, different proxies, even if their returns are highly correlated, may lead to different beta estimates for the same security. The Capital Asset Pricing Model has also been criticized that it reduces the situation to very extreme case. Even if the model explains the behavior of security returns, it does not necessarily explain the behavior of individual investors. For example, investors may analyze and process the information in a different way and therefore they might have different expectations about securities future performance. (Elton et al. 2003)

However, despite the criticism directed to the CAPM it is widely used in finance. Obviously, it describes the reality in a quite reliable way. Another reason for its employment is its mathematical simplicity. Therefore, the CAPM is generally used e.g. in project and portfolio evaluations, in a firm’s capital budgeting, portfolio construction and even measuring the effect of policy change on risk. (Chen, 2003)

2.2.1 The Jensen Alpha

Based highly on the Capital Asset Pricing Model, Jensen (1968) derives a measure for portfolio performance called Jensen alpha. It measures the average return on the portfolio over and above that predicted by CAPM. The Jensen alpha can be given as follows:

$$r_p - r_f = \alpha_p + \beta_p \left(r_m - r_f \right) \qquad (3)$$

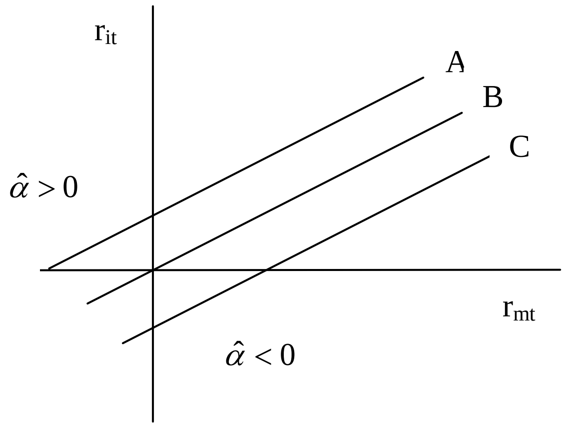
where r_p return for portfolio *p*, α_p is the Jensen Alpha of portfolio *p*, r_f is return of the risk-free asset, r_m is the return for the market portfolio and β_p is the beta coefficient for portfolio *p*.

The Jensen alpha can be interpreted so that if the α_p is positive the portfolio has performed better than the CAPM has predicted. Moreover, the higher the alpha the better performance the portfolio has obtained. On the other hand if the α_p is below zero it indicates that the portfolio has underperformed compared with predicted by the Capital Asset Pricing Model. Jensen (1969) suggests that the alpha measures the forecasting ability of a fund manager. Therefore, if the manager has the ability to forecast security prices (or perhaps some insider information not available to others) it should lead to a positive abnormal return compared with CAPM.

To clarify this more, the Jensen alpha can also be graphically described. It can be demonstrated as the vertical distance of the investment's characteristic line from the origin where market excess return is presented on the horizontal axis and excess return on investment is on vertical axis. Figure 1 gives an example of three portfolio's characteristic lines. Clearly, each portfolio has an equal beta coefficient. On the other hand, their intercepts differ. Portfolio B has an intercept of zero, but the intercepts of the portfolios A and C are different from zero. This means that these portfolios have earned abnormal return different from what was predicted by CAPM. Obviously, the portfolio A has a positive abnormal return when the abnormal return on the portfolio C is negative.

Figure 1. The Jensen Alpha in the r_i, r_m-space (Pätäri 2000, 41)

Figure 1 displays the interpretation of the Jensen Alpha. On the horizontal axis is presented the market return (r_{mt}) and on the vertical axis is presented the return on the investment (r_{it}). A, B and C describe different portfolios.



If α_p and β_p are assumed to be constant over the evaluation period, they can easily be estimated using the simple linear regression. Therefore, this equation can be presented as follows (Pätäri 2000, 40-41; Sharpe et al. 1999, 841):

$$r_{pt} - r_{ft} = \alpha_p + \beta_p \left(r_{mt} - r_{ft} \right) + \varepsilon_{pt} \tag{4}$$

where r_{pt} is return for portfolio p at time t, α_p is the Jensen Alpha of portfolio p, r_{ft} means the return of the risk-free asset at time t, r_{mt} is the return for the market portfolio at time *t*, β_p is the beta coefficient for portfolio *p* and ε_{pt} is the error term of portfolio *p* at time *t*.

2.2.2 The Fama-French 3-factor model

The Capital Asset Pricing Model assumes that only one risk factor affects the expected return. This is the covariance between the return on security and the return on the market portfolio i.e. the beta coefficient (Cuthbertson 2000, 61). However, the employment and development of multi-factor models in the security selection, in the investment management and in the evaluation of portfolio performance has grown rapidly. These multi-factor models have become popular since empirical results have suggested that there are more factors that may affect the expected asset returns than just one, like the CAPM assumes. (Elton et al. 2003, 383)

In developing one of the most popular multi-factor model, Fama and French (1992) try to identify additional factors that might explain stock returns. Using data from the US equity market, they test the joint roles of market β, size, E/P, leverage and book-to-market equity in the cross section of average stock returns. Their main findings indicate that relation between average returns and the β is not strong. However, used alone, size, book-to-market equity, E/P and leverage implicate strong relation with the cross-section of average stock returns. Moreover, Fama and French (1993) further extend and develop their previous study. They e.g. use time-series regressions to study asset pricing and form portfolios to mimic the risk factors related to size and book-to-market equity. Based on the results they conclude that three factors are for the most part able to capture strong variation in returns, no matter what other factors are used in the same regression.[4] These factors are:

[4] As e.g. Prigent (2007, 150) points out, it is worthwhile to note that the Fama-French model assumes that the market is efficient, but more than one factor is needed to explain asset returns.

1. The excess return on a market portfolio. *(r_m-r_f)*
2. The difference between the return on a small stock portfolio and the return on a large stock portfolio. *(SMB)*
3. The difference between the return on a high book-to-market stock portfolio and the return on a low book-to-market portfolio. *(HML)*

Finally, generalized in the equation form the three-factor model suggested by Fama and French (1993), which expands the Capital Asset Pricing model, can be given as follows:

$$E(r_i) - r_f = b_i[E(r_m) - r_f] + s_i E(SMB) + h_i E(HML) \tag{5}$$

where $E(r_m)$-r_f , *E(SMB)* and *E(HML)* denote the expected premiums for each factor described before. The b_i, s_i and h_i in the equation 5 measure the sensitivity of each factor (i.e. factor beta or factor loading) in the expected return and they can be estimated through time series regression as follows:

$$r_i - r_f = \alpha_i + \beta_i(r_m - r_f) + s_i SMB + h_i HML + \varepsilon_i \tag{6}$$

On the other hand, the intercept of the previous regression i.e. α_i can be also interpreted as a performance measure. However, following this approach, it is possible to capture excess returns generated by tactical asset allocation strategies that try to exploit inconsistencies of the Capital Asset Pricing Model. To be more specific, fund excess returns are decomposed into three components; excess market returns, returns generated based on well known strategies of buying small-cap stocks and selling large-cap stocks (SMB), and finally returns generated by buying stocks with high book-to-market ratios and selling stocks with low book-to-market ratios (HML). Therefore, the intercept in the Equation 6 represents the value that the manager has added to the portfolio over and above what could be justified by market risk and generated by these known strategies.

Hence, at least in theory, statistically significant positive alpha would implicate some managerial skill. (Babalos et al. 2007)

2.3 The Sharpe ratio

The Sharpe Ratio developed by William Sharpe (1966) is one of the most commonly used performance measures due to its simplicity. [5] The ratio is calculated by dividing the excess return on the portfolio by the standard deviation of the return.[6] Therefore, it takes a different approach to performance measurement than the two previous models. Mathematically the Sharpe ratio can be given as follows:

$$S = \frac{r_p - r_f}{\sigma_p} \tag{7}$$

where r_p is the return for the portfolio p, r_f is the return of the risk-free asset and σ_p describes the standard deviation of the returns of portfolio p.

Respectively, the standard deviation σ_p of the portfolio p needed in the previous formula can be given as follows:

$$\sigma_p = \sqrt{\frac{\sum_{i=1}^{N}\left(r_{pt} - \overline{r_p}\right)}{n-1}} \tag{8}$$

where r_{it} is the return of portfolio p at time t, $\overline{r_p}$ is the mean return of portfolio p and n is the total number of observations.

---

[5] In addition to discussion of the ratio, Sharpe (1994) provides broader range of the applications for the original measure.

[6] Existing literature suggests alternative names for the Sharpe ratio like the Sharpe index, the Sharpe measure and reward-to-variability ratio. Pätäri (2000)

Investors can interpret the Sharpe ratio to denote, how much excess return they are receiving for the extra volatility they take for holding a riskier asset. To be more specific, the Sharpe ratio shows investors, if the return on portfolio is due to smart investment decisions or due to extra risk. According to Sharpe (1966, 120), fund performance might vary in two respects. Firstly, funds might exhibit different variability in returns due to selection of different degrees of risk or due to erroneous prediction of the risk related to particular portfolio. Secondly, funds with similar risks might show variability in returns due to inability of some managers to select under priced securities or to diversify properly their holdings. Hence, the Sharpe ratio measures also the managerial skill. When comparing e.g. two different funds one can be seen as a good investment if these higher returns are not due to too much additional risk. Therefore, investors are often advised to pick portfolios with high ratios.

When using the Sharpe ratio it is reasonable to note that standard deviation measures the total risk of the investment and therefore including also the unsystematic risk. However, because the risk is measured this way, the Sharpe ratio is independent from the asset pricing models such as the CAPM. It does not take into account e.g. the correlation structure of the returns with the investor's other holdings. On the other hand, Elton et al. (2003) propose that the Sharpe ratio looks the investment decision from the investor's point of view. Therefore, it assumes investors to choose mutual funds to represent majority of their investments. If it is so, investors are only concerned with the full risk of the fund and the standard deviation is a reasonable measure for that risk. Hence, the employment of the standard deviation as a risk component makes the Sharpe ratio most useful in situations where the investor has only one risky investment. (Elton et al. 2003)

In practice, there may be situations when funds have underperformed the risk-free interest rate on average and hence have negative excess returns. To be more specific, when sorting funds based on the Sharpe ratio in descending order the funds will be ordered correctly if the excess return is positive. On the other hand, if the return is negative sorting funds in descending order will lead to unreliable rankings. For example, in a case of two funds with equal positive excess return, the one with the lower standard deviation will receive the highest score. However, if the average excess returns are equal but negative, the fund with the higher standard deviation receives the highest Sharpe ratio score (less negative). Therefore, comparing the Sharpe ratios e.g. when analyzing different funds can cause problems. (Israelsen, 2003)

2.4 Downside risk-based performance measures

The performance measures presented above are based on the mean-variance framework. This means that investors try to maximize the expected return (i.e. average return) and try to minimize the expected risk (i.e. variance). However, the employment of the variance (or the standard deviation) as a risk measure has been controversial topic among the academics. The main criticism is directed to the approach which gives an equal weight to upside and downside fluctuations of the returns.

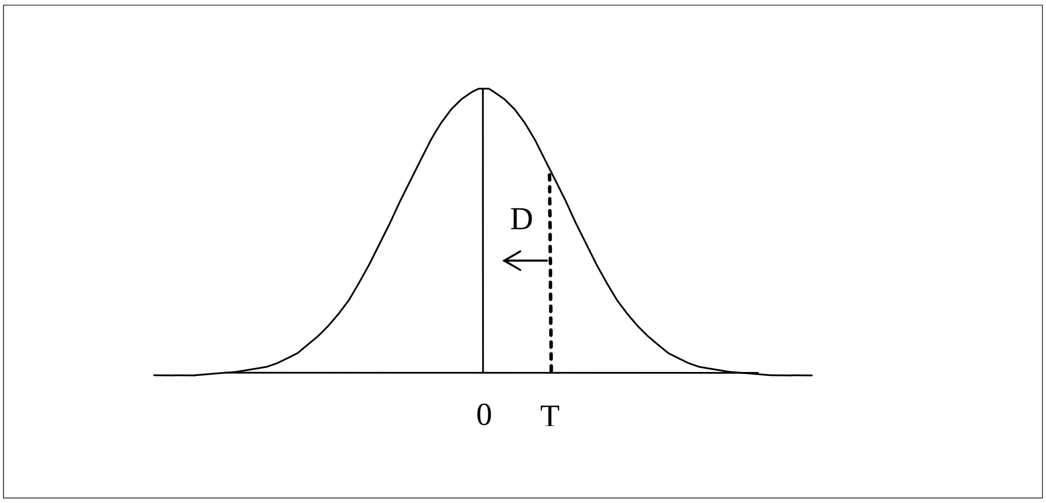
For example DiMarzio et al. (1993) suggest that rational investors do not necessarily view risk in this way where positive and negative deviations are treated equally. Firstly, investors do not normally worry if the value of their portfolio suddenly increases because they perceive positive volatility as a good outcome. Secondly, investors consider risk as the possibility of a bad outcome i.e. when the rate of return falls below some minimal acceptable return.[7] Therefore, investors are risk averse since they desire to avoid shortfalls below their minimal acceptable return. This leads us to another risk concept in finance, better known as the downside risk. The generalized idea of the downside risk is that the left-hand side of a return

[7] In financial literature and in performance measurement the minimum acceptable return is often defined as MAR.

distribution involves the risk since the right-hand side includes the better investment opportunities. Figure 2 illustrates this situation. *T* is the target return for an investor and the *D* describes the downside risk i.e. the returns that are below the target.

Figure 2. Graphical presentation of the downside risk

Figure 2 gives a graphical presentation of the downside risk. T describes the target return or minimum acceptable return for investor and D denotes the returns falling below this target return.



As a consequence, among academics semi-deviation is one of the commonly accepted measures for the downside risk. For example, DiMarzio et al. (1993) and Grootveld and Hallerbach (1999, 306) propose semi-deviation as a better measure of risk instead of standard deviation. They claim that a risk concept in which undesirable downside fluctuations are captured in this way, better matches investors' intuition about the risk than the standard deviation. Following e.g. Estrada (2004; 2006) the semi-deviation with respect to a specified benchmark *T* can be defined as follows:

$$\sum_T = \sqrt{(1/N)\cdot\sum_{t=1}^{N}\{Min(r_t - T, 0)\}^2} \qquad (9)$$

where, *N* is number of observations, *T* stands for the benchmark return or target return and *t* is time. When it comes to relevant target return e.g. Hwang and Pedersen (2004, 112) suggest that one possibility is to use a risk-free interest rate. Eftekhari et al. (2000, 21-22) propose the sample mean as a critical return. Moreover, Prigent (2007, 363) propose that in order to control the loss risk, target return could be set at zero. However, one should note that the relevant target or benchmark return depends always on investor's preferences. Therefore, there is no unanimous answer for this.

2.4.1 Downside deviation-based Sharpe ratio

One performance measure that is able to capture the downside risk of a portfolio is so-called downside deviation-based Sharpe ratio (DDSR) and it can be derived from the traditional Sharpe Ratio suggested before. However, the risk measure i.e. standard deviation is replaced with target semi-standard deviation (TSD), which can be firstly given as follows (Pätäri 2000, 94):

$$TSD = \sqrt{\frac{\sum_{i=1}^{n}(r_i - r_t)^2}{n}} \quad \text{for all } r_i < r_t \tag{10}$$

where r_i is return on the portfolio *i* for each period, r_t is the target return below which outcomes are considered risky and *n* denotes the number of outcomes in the whole distribution. As we can see, the formula for semi-standard deviation shows some similarity with the standard deviation used in the Sharpe ratio. However, there are two important differences. At first, in calculating the numerator, the target return is used instead of the mean return. Secondly, only negative deviations are included in the sum of the subtractions.[8]

[8] In this study we use zero as a target rate of return.

Finally the downside deviation-based Sharpe ratio (DDSR) can be given as follows (Pätäri 2008b, 77):

$$DDSR = \frac{r_i - r_f}{TSD_i} \tag{11}$$

where r_i is the return on portfolio *i* during the observation period, r_f is the risk-free rate of return and TSD_i is the downside deviation (target semi-standard deviation) for portfolio *i.* Despite the adjusted risk factor, the interpretation of this measure remains the same compared with traditional Sharpe ratio.

2.4.2 Modified Sharpe ratio

Besides characteristics presented before, other additional parameters could be crucial for investors when evaluating performance of a mutual fund. For example, for risk-averse investors negative skewness is an unwelcome characteristic. To be more specific, rational investors prefer positive skewness, which offers better protection against losses and provides higher profit opportunities in form of higher returns. Moreover, fund returns can also show excess kurtosis, which is also known as fat tails. This implies that there is higher probability of big positive and big negative returns than indicated by normal probability distribution. (Favre and Signer, 2002)

Therefore, incorporating these additional characteristics leads us to the third (skweness) and fourth (kurtosis) order moments of the return distribution. However, mean-variance based performance measures do not go far enough to capture these moments. Therefore, one potential approach would be to adjust the risk measure so that also the third and the fourth order moments of the return distribution can be taken into account. (Favre and Signer, 2002)

One option to adjust the risk in terms of skewness and kurtosis is to employ so-called Cornish-Fisher expansion, which can be given as follows:

$$Z_{CF} = Z_C + \frac{1}{6}\left(Z_C^{\,2} - 1\right)S + \frac{1}{24}\left(Z_C^{\,3} - 3Z_C\right)K - \frac{1}{36}\left(2Z_C^{\,3} - 5Z_C\right)S^2 \qquad (12)$$

Where, Z_c is the critical value for the probability with a standard normal distribution.[9] S denotes the skewness and K stands for the kurtosis of the return distribution. Respectively, the skewness and kurtosis in the previous formula are defined as follows:

$$S = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_t - r}{\sigma}\right)^3 \qquad (13)$$

$$K = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_t - \bar{r}}{\sigma}\right)^4 - 3 \qquad (14)$$

Where, T is number of observations, r is the return of a portfolio and σ is the standard deviation of a portfolio. Next, using the Cornish-Fisher expansion we formulate the modified deviation (MD) for the portfolio risk. It can be given as follows:

$$MD = \frac{Z_{CF_i}}{Z_C} \times \sigma_i \qquad (15)$$

Where, Z_{CF} is the Cornish-Fisher expansion presented above, Z_C is the critical value from the standard normal distribution and σ is the standard deviation of the portfolio *i*. The Cornish-Fisher expansion means that the portfolio risk can be now calculated for asymmetric distributions since the modified deviation scales the standard deviation according to skewness and kurtosis. Finally, the performance measure, the modified Sharpe ratio can be derived as follows:

---

[9] Z_c is equal to -2.33 for a 99% probability or to -1.96 for a 95% probability.

$$Mod.Sharpe = \frac{r_i - r_f}{MD} \tag{16}$$

where r_i is the return on portfolio i and r_f denotes the risk-free rate of return. The previous measure is similar to the traditional Sharpe ratio. However, the advantage of this measure is that it is able to incorporate possible non-normalities of the return distribution through the modified risk term. Therefore, this measure may lead to the choice of different portfolios and provide interesting results when measuring the fund performance.

2.4.3 Downside risk-based measures and emerging markets

Although there has been some debate concerning suitable risk measures, some authors have proposed the usefulness of the downside risk-based measures especially when evaluating the emerging market investments. As far as unsymmetrical return distributions are concerned, e.g. Bekaert et al. (1998) argue that emerging market returns may show significant kurtosis and skewness and hence e.g. Plantiga et al. (2001) and Estrada (2000; 2002; 2007) propose that downside risk measures are maybe able to better overcome these unsymmetrical return distributions.

Raj et al. (2001, 3) suggest that emerging markets may respond more rapidly to the negative news and hence increase the downside risk. On the other hand, in the case of positive news, the market may be more skeptical and react slowly. Hence, the authors argue that it would be relevant to use semideviation in an environment such as the emerging markets. Moreover, e.g. Stevenson (2001) analyses the use of downside risk measures in the construction of an optimal international portfolio covering 15 emerging markets and 23 developed markets. The results suggest that for risk-averse investors the employment of downside risk measures can result in significant improvements in performance, particularly in the context of minimum risk portfolios. Finally, e.g. Hwang and Pedersen (2004) go even further concluding that risk and asset management in emerging markets would require even customized approaches to risk quantification depending on the regions.

Therefore, being motivated by these suggestions and the previous empirical results, we consider it interesting and relevant to apply performance metrics that capture the downside risk and also pay attention to asymmetrical return distribution also in this thesis.

2.5 Portfolio management

One factor why some fund managers outperform others may be due to different investment strategies. Therefore, depending on the fund management style, a distinction is often made between passive and active management in the investment industry. Advocates of passive management believe that the markets behave according to the efficient market hypothesis. Therefore, a simplest case of passive management strategy seeks to match the return and risk of a market segment or an index by replicating exactly its composition. (Elton et al. 2003, 676-677)

A fund replicating some index buys each stock in the index in exactly the same amount it represents of the index. Although replicating some certain index is the simplest technique for constructing an index fund, many of index funds are not constructed this way. (Elton et al. 2003, 676-677) This is mainly because passive managers face various decisions and problems when trying to replicate an index. These involve e.g. dealing with transaction costs and the trade-off between accuracy in replicating the index.[10] Therefore, the manager has to decide if it is necessary to buy some stocks with smallest market weight or exclude them in order to lower the transaction costs. (Elton et al. 2003, 676-677)

[10] This trade off is often called tracking error.

Elton et al. (2003, 677) suggest a couple of approaches to construct an index fund. Each of these approaches makes a distinction between accuracy in replicating the index and transaction costs. These three approaches are as follows:

1. Holding each stock in the proportion that it represents of the index.
2. Mathematically forming a portfolio with specified number of stocks (e.g. 300), which best tracks the index historically.
3. Finding a smaller set of stocks that matches the index in the percent invested in a specified set of characteristics (e.g. same percent in industrial, utility and financial stocks). Commonly used characteristics have been sector, industry, quality and size of capitalization.

Moreover, also combinations of all three approaches can be used to construct an index fund. Active management instead, takes a different position from the passive management. Active portfolio managers do not follow the efficient market hypothesis. They believe that it is possible to profit from the stock markets through various strategies that aim to identify mispriced securities. Therefore, active management is based strongly on a forecast about the future. Existing literature has normally classified active management styles into three classes; market timers, sector selectors and security selectors. Market timers change the beta of the portfolio according to their forecasts on the market. If an active manager assumes the bull market, he will increase the beta of the portfolio and when the bear market is assumed then the manager will lower the beta of the portfolio, respectively.[11] For example, the portfolio composition toward higher beta can be implemented as follows (Pätäri 2000, 47):

1. By increasing the ratio of stocks to other assets (e.g. bonds)

[11] The term “bull market” is often used to describe a stock market that is rising or is expected to rise. Respectively, the term “bear market” refers to declining markets.

2. By investing in stock whose market risk (beta) is greater compared with the stock that was previously included in the portfolio
3. The beta of the portfolio can be raised by investing in offensive derivatives e.g. buying forward contracts or options.

When the bull market is assumed the implementation is done the other way around. Another active management style is security selection. This means searching for undervalued securities. Managers who are practicing security selection are making bets that the market weights on securities are not the optimum amount to hold in each security. Therefore, managers increase the weight of undervalued stocks (i.e. make a positive bet) and decrease the weight of overvalued stock. A third frequently used method in active portfolio management is sector or industry selection. This is similar to security selection with the exception that the unit of interest is a certain sector or an industry. Based on the analysis, a positive or negative bet will be made on a sector. Managers who practise sector selection will rotate their portfolios' overweighting and underweighting sectors over time as their forecasts change. (Elton et al. 2003, 677-678)

However, despite the various strategies discussed above especially managers managing the emerging market funds may face some problems when trying to implement these strategies. For example, transaction costs may be higher in the emerging markets. Moreover, some trading barriers may exist. Therefore, these problems may prevent fund managers from followings some certain investment strategies.

3 PREVIOUS RESEARCH

During the last decade there has been a general shift in persistence studies to use shorter selection and holding periods compared with the very first studies. Another interesting feature has been the employment and development of several multi-index models in performance evaluation. Moreover, recently several authors have proposed even more innovative methods, especially the Bayesian estimation, for performance measurement.

In presenting the previous literature related to the topic we will focus more on the post 1992s and the more recent studies. As can be later seen, lots of studies have been conducted using data based on the US and the UK mutual funds. We will report the methodologies used in these studies and we will also report the most important findings. However, despite the recent developments and innovations, some of the pioneer studies have had a great influence on this topic. Therefore, we consider important to start with presenting them first.

3.1 Early persistence studies

In one of the very first studies Sharpe (1966) examines performance persistence of 34 US mutual funds. The data covers a period from 1944 to 1963. Using both 10-year holding and selection periods Sharpe ranks funds based on the Sharpe ratio and the Treynor ratio.[12] The results show weak positive correlation between fund rankings although not statistically significant.

[12] The Treynor Ratio can be calculated dividing the excess return beyond risk-free return by the systematic risk of investment i.e. as follows:

$T = \frac{r_i - r_f}{\beta_i}$, where r_i denotes return on the investment *i,* r_f is the risk free rate of return and β_i stands for the beta coefficient of the investment *i*.

Jensen (1968) investigates the performance of 115 mutual funds and their ability to predict the future performance during the period of 1945-1964 using the Jensen alpha method. The author employs also selection and holding periods of 10 years such as Sharpe (1966). The results show positive correlation in the performance between holding and selection period meaning that some fund may consistently outperform the other funds. However, Jensen concludes that these findings about possible managerial skill should be interpreted carefully since he suggests that the persistence might be mainly due to persistence of inferior performers.

Using raw returns and the Sharpe ratio, Carlson (1970) explores the performance of 57 mutual funds over a period from 1948 to 1967. When Carlson compares two consecutive ten-year periods, the results show no persistence in performance. Moreover, the author examines a smaller sample including 33 common stock funds in a same way. However, the results remain the same. In addition, the author further divides the data into two five-year periods, which changes the results significantly. The results indicate a greater degree of performance persistence, since the funds seem to remain in the top or bottom groupings on the holding period.

3.2 The studies from the 1990s

In the early nineties, Grinblatt and Titman (1992) analyze monthly mutual fund data of 279 US funds from 1975 to 1984. Performance is measure using an extension of the Jensen's Alpha.[13] In order to check the reliability of the persistence tests, the authors construct a control sample of 109 passive portfolios. To study persistence in fund returns the authors run a cross-sectional regression of abnormal returns where the five-year holding

[13] The performance measure is computed relative to the eight-portfolio benchmark. The idea behind formation this benchmark is that various firm's characteristics are correlated with their stock's factors loadings. Therefore, portfolios formed from stocks grouped by different securities characteristics can be used as proxies for the factors.

period returns are explained by the previous five-year selection period returns. The results indicate positive persistence in mutual fund performance. The authors conclude that these results cannot be explained by inefficiencies in the benchmark that are related to various firm's characteristics such as size, dividend yield or past returns.

In one of the very first studies concentrating on the short-run performance persistence, Hendricks et al. (1993) explore quarterly returns of 165 open-end, no-load, growth-orientated mutual funds during 1974-1988. The authors use a selection period of one-year when the holding periods range from 3 months to 2 years. The funds are ranked based on several methods including e.g. the Sharpe ratio and multi-factor regressions. The results indicate that top performing funds tend to continue superior performance in the near future. In addition, results also show that funds that have performed poorly in the recent years tend to perform poorly in the near future as well. Actually the persistence of poor performance seems to be stronger than the persistence of superior performance.

Using absolute and relative benchmarks, Brown and Goetzmann (1995) study to what extent the previous year performance of a fund can predict the performance in the following year. The sample ranges from 372 common stock funds in 1976 to 829 funds in 1988. To evaluate fund performance the authors employ several methods including the traditional Jensen alpha, the three-index model, the appraisal ratio and the three-index appraisal ratio.[14] In order to test the performance persistence the authors apply a nonparametric methodology based upon contingency tables.[15] The findings indicate significant persistence for some certain periods. On the other hand, also performance reversal occurs. These results are also parallel with the results obtained by Malkiel (1995) who

[14] The appraisal ratio is calculated dividing the CAPM alpha by residual standard deviation and the three index appraisal ratio is calculated dividing the three index alpha by residual standard deviation respectively.
[15] Contingency table identifies a fund as a winner in the current year if it is above or equal to the median of all funds with returns reported that year. The same criterion is used to identify it as a winner or loser for the following period.

also suggests that the pattern of persistence may be dependent on the evaluation period employed.

Elton et al. (1996) study a survivorship-bias free sample of 188 mutual funds from 1977 to 1993. They measure fund performance using raw returns, the one-, three and four-factor alpha. Based on these performance measures, they rank funds into deciles and then investigate how these deciles perform in subsequent periods using one- and four-index alpha. The results indicate that for the one-year performance period, all measurement techniques show statistically significant correlation with the future performance. Moreover, for the three-year evaluation period all methods except the raw returns show again statistically significant rank correlation. Therefore, the authors conclude that the past return data carries useful information about the future. Moreover, using nearly same data and similar methods, Sauer (1997) documents parallel results. However, after portioning the data by investment object the pattern of performance persistence disappears.

Using CAPM, three and four-factor models, Carhart (1997) examine performance persistence of diversified equity funds during a period from 1962 to 1993.[16] The data consists of the monthly returns of 1892 funds. To investigate the short run persistence funds are sorted into portfolios based on lagged one-year returns and these portfolios are sorted every year into deciles. Then the performance persistence is evaluated based on these portfolios. The results for one-year lagged returns indicate strong performance persistence. To examine long-run persistence the author forms again portfolios based on lagged two- and five-year returns but the results show no persistence for longer intervals. However, maybe the most

[16] In addition to size and style factor portfolios of Fama-French (1993), a momentum portfolio *(PR1YR)*, which tries to catch the attributable return created by momentum and contrarian stocks is added in the four-factor model. Therefore, the model takes the following form:

$$r_i - r_f = \alpha_i + b_i\left(r_m - r_f\right) + s_i SMB + h_i HML + p_i PR1YR$$

where, the last term describes the momentum portfolio.

important finding of the study is that most of the short run persistence seems to be explained by the common factors of the four-factor model i.e. size, book-to-market and momentum factors.

When it comes to studies based on the UK markets, Allen and Tan (1999) investigate the persistence of investment trust company managers on UK funds during 1989-1995. The data covers monthly returns of 131 funds. To measure performance the authors employ both raw returns such as the Jensen alphas and apply evaluation periods of two-year, one-year, six month and one month. They find that both the raw returns and the risk adjusted returns exhibit persistence over one-year and two-year intervals. For shorter periods their results show only reversal pattern in performance. Interestingly, the results obtained by Allen and Tan differ from other studies concentrating on the UK market. For example, Fletcher (1999) uses similar methods to examine the UK mutual funds, but he reports no evidence of performance persistence. Moreover, e.g. Quigley and Sinquefield (2000) conclude that based on their study the UK mutual funds do not exhibit persistence in performance. On the other hand, the results obtained by Quigley and Sinquefield may differ due to methodological reasons since they employ also the three-factor alpha for performance estimation.

3.3 The studies of the 2000s

Blake and Morey (2000) conduct an interesting study when comparing the Morningstar rating system as a predictor of fund performance with traditional performance measures including Sharpe ratio, total returns, the Jensen alpha and the four-factor alpha. The prediction power of the past performance is evaluated through regression analysis and the nonparametric Spearman rank correlation test. The data covers two different sample groups depending on the length of period, during roughly 1983-1997. The main findings are that based on a 10-year selection period only the Sharpe ratio does better than the Morningstar ratings. On the other hand, results also indicate that total returns and the four-index alpha do considerably worse. Interestingly, for shorter selection periods (for 3-and 5-years), results show that the Morningstar rating system has a better prediction power than the traditional performance measures. Especially the rating system seems to be able to detect well the bad-performing funds.

ter Horst et al. (2001) investigate how survivorship bias and look-ahead-bias affect on mutual fund performance. The sample covers 2678 US equity funds during a period from 1989 to 1994 and performance is examined employing the raw returns, the one-factor alpha and the four-factor alpha. To study short run persistence the authors use 1-year selection and holding periods, and for medium-term performance 3-year selection and holding periods, respectively. Results indicate that without any risk adjustments funds with growth investment style exhibit short-term persistence. The authors propose that a strategy of buying last year winners from the sample would have outperformed a strategy of buying last year losers with almost 6.76 % on an annual basis. However, the results show that the persistence disappears when the factor models are used but on the other hand the results implicate some performance reversal. Moreover, the authors show that the look-ahead-bias can cause spurious pattern in performance persistence although not in this particular study.

To extend the existing evidence in mutual fund studies Deaves (2004) examines performance persistence of Canadian equity funds. The survivorship-bias free sample covers time period from 1988 to 1998. To rank the funds the author uses the Jensen alpha, the conditional CAPM alpha and five-factor alpha. The author measure persistence by the methodology based on contingency tables. The results show significant short term persistence when a one-year selection period is used to evaluate future performance. In contrast, for longer periods funds do not show any sign of performance persistence. However, Deaves concludes that it is likely that the Canadian fund managers have some sort of stock picking ability.

Busse & Irvine (2006) use daily returns to compare the performance predictability of the Bayesian estimates of fund performance with the standard measures (the Jensen alpha, Fama-French alpha and four-factor alpha). The data covers 230 equity funds over a period from 1985 to 1995. Interestingly, when they formulate the Bayesian estimates the results show that they predict future performance better than the standard alphas.[17] During the sample period the Bayesian results indicate the strongest persistence. Moreover, the strength in persistence increases when it is evaluated based one-year selection period. Therefore, the authors conclude that the Bayesian measures are particularly useful for predicting future alphas.

Using monthly returns of around 6400 US equity mutual funds Huij and Verbeek (2007) examine short-run performance persistence during the period from 1984 to 2003. Using the four-factor model developed by Carhart (1997) the funds are sorted into decile portfolios. To evaluate the performance Huij and Verbeek use both the traditional alphas obtained through OLS procedure as well as the Bayesian alphas.[18] The results clearly indicate the validity of the prior information in predicting the future performance. For both 36 month and 12 month horizons, top decile funds significantly outperform the bottom decile funds. Moreover, more interesting finding is that the predictive accuracy of the Bayesian alphas is significantly higher compared to OLS alphas. They find that on average the Bayesian alphas are around 40% more accurate compared with the standard OLS alphas. Therefore, the results are parallel with ones

[17] When estimating the Bayesian alphas Busse and Irvine (2006) combine e.g. investors' beliefs on managerial skills such as returns on passive assets.
[18] The Bayesian method employed by Huij and Verbeek (2007) is different from the one employed by Busse and Irvine (2006) since the investors' inference is based solely on the monthly returns of the whole cross-section of mutual funds.

documented by Busse and Irvine (2006), supporting the advantage of the Bayesian estimation.

When it comes to studies related to emerging markets, Huij and Post (2008) analyze monthly returns of US mutual funds investing in multiple emerging countries during 1967-2006. Using ranking periods of 12 months, three months and one month the funds are ranked into 3-quartiles based on the single-index alpha. Then, equally weighted time series of returns are calculated for each 3-quartile on the following period and holding period performance of these portfolios is evaluated using the Sharpe ratio and the single-index alpha. The results clearly indicate performance persistence, with strongest evidence for three-month and one-month selection periods. Moreover, using several tests the authors show that the persistence can not be attributed to fund characteristics (e.g. expenses and load fees) or to exposures to common factors such as currency or commodity exposures. Only one-third of persistence can be explained by the momentum strategies. Therefore, the authors conclude that the fund managers may possess some investment skills.

4 DATA AND METHODOLOGY

4.1 Data description

This study explores European equity funds investing in the Russian stock market. The data for this Master's thesis consists of weekly returns on a sample of the equity funds that existed during the period from 2002 to 2007. The data is provided by the Morningstar. The minimum length of return history is set to two years, since we employ one-year selection period and one-year holding period. Therefore, we exclude those funds that do not meet this criterion.

Some data conditioning issues have received considerable attention in performance persistence studies. One of these is so called survivorship bias, firstly documented by Brown et al. (1992). More specifically, this means the problem of how to deal with the failed or merged funds during the period under study. Survivorship bias rises if only the funds that exist at the end of the observation period are included in the sample instead of including all funds.

Our sample is basically free of the survivorship bias since it contains all the existing funds from 2002 to 2007, except the funds we had to remove for the reason documented above. Obviously, this raises some bias in our sample, but it can not be avoided by any means. On the other hand, as can be noticed from the studies authored e.g. by Malkiel (1995), Blake and Timmermann (1998), Quickley and Sinquefield (2000) or Carhart et al. (2002), the real economical impact of the survivorship bias on fund performance and especially performance persistence is less certain and a rather controversial topic.

Moreover, it is possible that our data may suffer from some look-ahead bias. It means the bias stemming from the employment of data or information that would not have been available during the period being analyzed. For example, in persistence studies it is common to rank funds based on the selection period performance and put them into portfolios. Then the average performance of these portfolios is determined for the funds that survived the holding period. Look-ahead bias stems if the number of funds is not stable within these portfolios between the selection and holding period. However, this bias can not be eliminated even if a survivorship bias free sample is used. (Pätäri, 2008a)

Table 1 reports the descriptive statistics of the sample employed in this study. As can be seen, average weekly return of the sample varies from -0.50 percent in 2002 to 1.290 percent in 2005. Weekly standard deviation varies from 3.450 percent in 2002 to 2.153 percent in 2007. Converted to annual volatility this means the variation between 24.7 percent and 15.5 percent. High volatility shows clearly that the funds' returns can change dramatically in either direction over a short period of time. In general, the funds seem to exhibit negative skewness. It varies from -0.975 in 2006 to -1.06 in 2005. Negative skewness implies that the returns are more concentrated on the right hand side of the return distribution. Kurtosis varies from -0.169 in 2004 to 2.962 in 2006. However, maybe the most interesting characteristics are presented in the two last columns. When we look at the maximum return in each year we can see that some of the funds have been able reward investors considerably well. The maximum return varies from 9.05 percent in 2007 to 14.59 percent in 2004. On the other hand, the column of maximum returns reveal an interesting feature as well, indicating that there has been a declining trend in maximum returns during the last three years.

The minimum returns show that not all funds have been successful. The minimum returns vary from -7.95 percent in 2005 to -20.15 percent in 2006. Therefore, if investors had chosen wrong funds, they might have suffered big losses. In general, the descriptive statistics show the profit possibilities of these emerging market funds but on the other hand, also the risks that are related to these funds.

Table 1. Fund return characteristics from 2002 to 2007

Table 1 shows fund's return characteristics in each year based on the weekly excess return of each fund. The table presents the number of funds in the beginning of the year, mean return, standard deviation, skewness, kurtosis and the maximum and minimum fund return in each year. Standard deviation, skewness and kurtosis are presented in average values.

Year	Num. of funds	Average	Std. dev	Skew.	Kurt.	Max.	Min.
2002	24	-0.050%	3.450%	-0.505	1.270	12.07%	-15.56%
2003	29	0.660%	3.347%	-0.508	0.564	10.21%	-12.70%
2004	31	0.040%	3.143%	-0.289	-0.169	14.59%	-12.89%
2005	38	1.290%	2.548%	-0.106	0.511	10.97%	-7.95%
2006	44	0.530%	2.960%	-0.975	2.962	10.28%	-20.15%
2007	54	0.280%	2.153%	-0.461	0.841	9.05%	-10.58%

The time series for the stock index returns likewise the interest rate used in this study are collected through several sources. They include Thomson's Datastream database, MSCIBarra and Thomson Banker One data source. As a proxy for the risk-free interest rate we use the one-month Euribor. Since the Euribor is quoted on a yearly basis and our other data is quoted on weekly basis, we convert the annual interest rates into weekly returns. To do this we follow the conversion formula of Vaihekoski (2007) as follows:

$$r_t^d = m^{-1} \times \ln\left(1 + R_t\right) \tag{17}$$

where, R_t is the one month Euribor, d stands for the length of the period in days over which the risk-free rate of return is calculated and m means the number of periods with length of d in a year.

As a proxy for the market portfolio we use the return on RTS Index. That is calculated based on the 50 most liquid Russian companies. In addition, the weight of each company is limited to 15 percent. Therefore, we consider it to be the most comprehensive index to describe the whole performance of the Russian stock market.

When it comes to size and value factors needed in the Fama-French (1993) model, we use the MSCI Russia Growth and MSCI Russia Value indices provided by Morgan Stanley to form the value factor. For size factor we use the MSCI Russia Large Cap and MSCI Russia Small Cap indices, respectively. Both size and value factor mimicking portfolios are then constructed as we discussed in the previous sections.

Table 2 presents the summary statistics for the factor portfolios. As can be seen, the average weekly returns for each factor mimicking portfolio are quite low. The standard deviations vary from 4.33 percent for the market portfolio to 2.47 percent for the size portfolio. The results also indicate that the regression models should be able to explain considerably well variation in returns. This is due to relatively low cross-correlations between each factor portfolio. The highest positive correlation seems to be 0.159. Table 2 also reveals an interesting relationship between the portfolios. In two cases the correlation seems to be negative.

On the other hand, low cross-correlations denote that our results obtained through regression analysis are not necessarily affected by multicollinearity. Multicollinearity may arise if there is strong correlation between explanatory variables in a regression model and therefore multicollinearity makes it difficult to isolate separate the effects of the individual explanatory variables. For example, Hill et al. (2001, 189-190) suggest that multicollinearity may cause estimation errors when using the ordinary least squares method. However, the regression analysis, the ordinary least squares method and its assumptions will be discussed more detailed in the next section.

Table 2. Summary statistics for the factor portfolios

Table 2 shows the summary statistics for the factor mimicking portfolios. r_m-r_f is the RTS Index minus risk-free interest rate and denotes the market proxy. HML and SMB are the Fama-French model's (1993) factor mimicking portfolios for book-to-market equity and size. The second and the third column show average weekly excess return and standard deviation for each factor mimicking portfolio. The last column presents the cross-sectional correlations between each factor mimicking portfolio.

Factor Portfolio	Average Weekly Return	Std. dev	Cross-Correlations		
			r_m-r_f	SMB	HML
r_m-r_f	0.019%	4.33%	1		
SMB	0.150%	2.47%	-0.480	1	
HML	0.003%	3.27%	-0.031	0.159	1

4.2 The ordinary least squares

Alpha-based performance measures employed in this study are estimated using the regression analysis and the ordinary least squares (OLS henceforth) method. In this case mutual fund performance can be measured by a fund alpha, which can be defined as the intercept term in a regression of the fund's excess returns on the (excess) returns of one or more benchmark factors. In practice, single or multiple regression models can be given as follows (Hendricks et al. 1993):

$$r_{it} = \alpha_i + \sum_{k=1}^{K} \beta_{ik} F_{k,t} + \varepsilon_{it} \tag{18}$$

where r_{it} is the excess return of fund *I* in period *t*, α_i is expected return of a fund *i* in excess of a factor mimicking portfolio or portfolios, β_{ik} denotes the sensitivity of fund *i* to factor *k*, *F* is the return of factor *k* at time *t* and ε_{it} is the error term. As a result of the above formula, alpha can be interpreted as the portion that can be attributed to a fund manager and his or her skills.

As briefly mentioned before, besides the multicollinearity there are a couple of assumptions for the ordinary least squares method to be valid. In particular, it is important to test that the residuals from the regressions follow the normal distribution:

$$\varepsilon_t \sim N\left(0, \sigma^2\right) \tag{19}$$

Secondly, it is important to investigate that the residuals do not exhibit autocorrelation:

$$Cor\left(\varepsilon_i \varepsilon_j\right) = 0 \tag{20}$$

Third important issue to study is that the residuals in the regression model have constant variance i.e. they are homoscedastic:

$$\mathrm{var}\left(\varepsilon_t\right) = \sigma_t^2 \tag{21}$$

Violation of these assumptions may cause problems related to the model. For example, it may lead to biased results in prediction interval estimation or in hypothesis testing. This may cause that the regression coefficients become unreliable. (Watsham and Parramore, 1997)

To study the regression diagnostics we perform several tests. When it comes to multicollinearity, the results in Table 2 already indicated low correlation between the explanatory variables. In addition, we run so-called auxiliary regressions.[19] The results obtained through these regressions are presented in Appendix 1. The results show that at its highest, the R^2 is around 0.25 when the size factor is explained by the other factors. This indicates that multicollinearity does not really affect our model. To determine the normality in the residuals we employ the Jarque-

---

[19] For example Hill et al. (2000, 190-191) suggest the usefulness of these regressions in detecting the multicollinearity. In these regressions, one of the explanatory variables is regressed on all the remaining explanatory variables. If the R^2 of these regressions is above 0.8, it indicates the multicollinearity between the variables.

Bera test. To investigate the homoscedasticity we employ the White's test. Finally, for detecting the autocorrelation we perform the Breusch-Godfrey test, which is considered as a general test for testing the autocorrelation. We run these tests separately for each fund using the residuals from the Fama-French three-factor model.[20]

The Jarque-Bera tests show (not presented) that in general the regression residuals are normally distributed. However, in 2006 the residuals exhibit rather strong non-normality. This is mainly due to high kurtosis of fund returns in that particular year (see Table 1). The other tests for the regression residuals showed some heteroscedasticity and autocorrelation (not presented). Hence, we performed additional regressions using so-called Newey-West autocorrelation and heteroscedasticity consistent standard errors. Then we compared these results separately with the original regressions. When it comes to heteroscedasticity, after adjusting the errors, in three cases the alpha became statistically significant and in one case insignificant compared with the original regression. As far as the autocorrelation is concerned, after adjusting the errors with the Newey-West regression, 8 alphas became statistically significant and three alphas became statistically insignificant compared with the original regressions that exhibited autocorrelation in residuals. Therefore, we consider that our results are not badly affected neither by the heteroscedasticity nor the autocorrelation.

4.3 The Bayesian method for fund performance

4.3.1 A short introduction to the Bayesian estimation

Obviously, the ordinary least squares method is widely used in mutual fund performance evaluation. However, the Bayesian estimation has gained a lot of attention during the recent years. One interesting reason

[20] The diagnostic tests were performed using the EViews program.

motivating the employment of the Bayesian estimation is that additional prior information can be exploited. According to Koop (2003, 2) the Bayesian approach is based on subjective view if probability, which argues that uncertainty about anything unknown can be expressed using the rules of probability.[21] More specifically, the Bayesians take it given that econometrics involves learning about something unknown θ (e.g. parameters of a model) given something known y (e.g. data) and conditional probability. This can be given as follows:

$$p(\theta|y) \propto p(y|\theta)p(\theta) \tag{22}$$

The term $p(\theta|y)$ can be defined as a posterior distribution for the data given the parameters of the model, $p(y|\theta)$ is the likelihood function and $p(\theta)$ is the prior distribution. However, the prior $p(\theta)$ does not depend on the data. In contrast, it can contain any non-data information about θ before seeing the data. For example, one could consider that all the funds are index trackers and their betas could be approximately close to one. Hence, one could have prior information related to θ before seeing the data.[22] However, normally the prior is chosen so that it has the same functional form than the likelihood function. The likelihood function $p(y|\theta)$ refers to the distribution of data conditional on the parameters of the model. Hence, it is often regarded as the data generating process. For example, the OLS assumes that the error terms follow the normal distribution. Therefore, this would mean that $p(y|\theta)$ is the normal distribution, which depends on the regression parameters. Finally, the posterior $p(\theta|y)$ is the distribution, which is our fundamental interest. Basically, it summarizes all we know about θ after seeing the data. Therefore, Equation 22 can be understood as an updating rule. In other

[21] An extensive overwiev of the Bayesian econometrics is given e.g. by Koop (2003)

[22] Moreover, e.g. Karoui (2008), Busse and Irvine (2006) such as Baks et al. (2001) suggest that in performance measurement prior could include issues such as managerial skills, funds' expenses or returns on other assets.

words, the data allows us to update our prior views about θ and the result is the posterior distribution which combines the both data and non-data prior information. (Koop 2003, 1-3)

4.3.2 The iterative empirical Bayesian procedure

One variation of the Bayesian approach is shrinkage estimators. However, the shrinkage estimators exploit the cross-sectional data based information to choose the prior. Therefore, this partly violates the basic Bayesian premise discussed before. On the other hand, this makes them useful e.g. in situations where the time series may be limited so that only a small number of observations are available. This short sample problem may lead to inaccurate and biased OLS alpha estimates. Therefore, shrinking them may improve accuracy. (Huij and Verbeek, 2003; Koop, 2003)

Huij and Verbeek (2003, 4) explain the shrinkage estimation that if the funds are similar, the data can be pooled and each fund can be characterized as by an overall estimate. On the other hand, if there is no similarity between the funds, a pooled estimate is uninformative and each fund should be estimated separately. However, it there is some similarity between the funds, both the pooled estimate and the time series estimate contain information and with shrinkage estimation the resulting estimate is a weighted average of both. In the case of inaccurate OLS estimates, large negative alphas may be underestimated, while high positive alphas may be overestimated. Therefore, shrinking them towards the common mean reduces the positive and negative estimation errors and increases the accuracy. This is simply due to incorporating the fact that funds' alphas tend to fluctuate around a common mean, close to zero. (Huij and Verbeek, 2003)

The iterative empirical Bayes is a shrinkage procedure where the degree of shrinkage varies depending on the variables. Some of the variables are less probable to occur than others and hence they are shrunk more towards the pooled estimate. Using this approach, the prior distribution can be specified for the unknown parameters, specifying the degree of prior uncertainty about parameter $\theta_{i.}$ Then, the posterior distribution can be derived conditional on the data.

A usual choice for the prior distribution is a normal distribution (Huij and Verbeek, 2003):

$$\theta_i \sim N(\mu, \Sigma) \tag{23}$$

If we assume that the error terms received from the ordinary leas squares are independent and identically distributed (i.i.d), the posterior distribution of θ_i can be assumed normal with the following expectation:

$$E\left(\boldsymbol{\theta}_i^*\right) = \left(\frac{\mathbf{1}}{\boldsymbol{\sigma}_i^2}\mathbf{X}_i'\mathbf{X}_i + \boldsymbol{\Sigma}^{-1}\right)^{-1}\left(\frac{\mathbf{1}}{\boldsymbol{\sigma}_i^2}\mathbf{X}_i'\mathbf{X}_i\hat{\boldsymbol{\theta}}_i + \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right) \tag{24}$$

where $\left(\frac{\mathbf{1}}{\boldsymbol{\sigma}_1^2}\mathbf{X}_i'\mathbf{X}_i + \boldsymbol{\Sigma}^{-1}\right)^{-1}$ is the covariance matrix, $\boldsymbol{\sigma}_i^2$ is the variance of the error term, $\boldsymbol{X}_i$ is the excess return matrix of the benchmark factors, $\boldsymbol{\Sigma}$ stands for the *(k+1)* by *(k+1)* covariance matrix of the OLS estimates $\hat{\boldsymbol{\theta}}_i$ and $\boldsymbol{\mu}$ is the *(k+1)*-dimensional vector of cross sectional means of the alphas and the factor sensitivities.

The equation 24, which determines the Bayesian alphas such as the factor sensitivities, shows that the posterior estimates of alpha and betas are a matrix-weighted average of the ordinary least squares estimates $\hat{\theta}_i$ and the prior μ. Huij and Verbeek (2007, 977) interpret this equation as shrinkage formula. It shrinks the raw estimates for alpha and beta obtained through OLS towards a common mean. The precision of the OLS estimates and the cross-sectional dispersion defines the degree of the shrinkage procedure. For example, the more similar the funds are, the more information the pooled estimate contains compared with the individual OLS estimate for each fund, and the more they are shrunk towards the overall pooled estimate.

One should also note that for derivation the posterior distribution requires that so-called hyperparameters *μ*, σ_i^2 and *Σ* are known. However, it raises some problems to estimation. For example, to estimate θ_i^**,* we have to know the parameters *μ*, σ_i^2 and *Σ.* On the other hand, to estimate the parameters *μ*, σ_i^2 and *Σ,* we have to know θ_i^*. Instead of fixing these hyperparameters at some priori values, Huij and Verbeek (2003, 3-4) suggest a couple of iterative Bayesian equations where $\hat{\theta}_i$ (the OLS estimate) is used as an initial estimate of posterior θ_i^*. The iteration process can be described as follows:

First, the cross-sectional means of alphas and factor sensitivities can be defined as:

$$\mu = \frac{1}{N}\sum_{i=N}^{N}\theta_i^* \tag{25}$$

Secondly, the parameter $\boldsymbol{\sigma}_i^2$ can be calculated as follows:

$$\boldsymbol{\sigma}_i^2 = \frac{1}{T_i - k - 1}\left(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\theta}_i^*\right)'\left(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\theta}_i^*\right) \tag{26}$$

And finally ***Σ*** can be given as follows:

$$\Sigma = \frac{1}{N-1}\left[D + \sum_{i=1}^{N}\left(\theta_i^* - \mu\right)\left(\theta_i^* - \mu\right)'\right] \tag{27}$$

where N is the number of funds, T_i is number of observations of fund i, k is the number of factors, y_i denotes the vector of the excess return of fund i and D is the additional diagonal matrix with small values on the diagonal.[23]

After defining the hyperparameters μ, σ_i^2 and Σ, the posterior θ_i^* can be estimated using the formula 21. Moreover, afterwards this process can be repeated using re-estimates for the hyperparameters and re-estimate the posterior θ_i^* until the desired degree of convergence of the posterior parameters is reached. This simulation process is called as the iterative empirical Bayesian process. Hence, when talking about the iterations, we refer to the number of re-estimations of posterior θ_i^* used in the Bayesian process.

Using the iterative Bayesian shrinkage estimators described above, we also formulate the Bayesian alphas in this thesis. We form the Bayesian alphas for both the selection and the holding period. Throughout the Bayesian shrinkage estimation, we employ the Fama-French three-factor model estimates obtained through OLS regression as initial estimates.

4.3.3 Efficiency of the Bayesian alphas

It is of our interest to study if and to what extend the Bayesian alphas provide better and more accurate estimates for fund performance than the OLS three-factor alphas. To extent that the estimates are more accurate, they may provide more information about future. For example, using the root mean square error test (RMSE), Huij and Verbeek (2003) show that in their sample, the Bayesian alphas were on average 40% more accurate compared with the OLS estimates. To investigate the accuracy of the

[23] In our study, the additional matrix has a size of 4 x 4. Hu and Maddala (1994) propose the employment of the additional matrix to improve the iterative procedure and to provide better estimates. Therefore, we replicate their study using value of 0.0001 along the diagonal.

shrinkage estimators we employ the mean squared error test (MSE) suggested e.g. by Mincer and Zarnowitz (1969) and Klemkosky and Maness (1978).[24] However, the MSE is normalized by dividing each component by the average of the selection period measure as follows:

$$MSE = \frac{\left(\overline{AB} - \overline{PB}\right)^2}{\overline{X}} + \frac{(1-B)^2 \sigma_{pb}^2}{\overline{X}} + \frac{\left(1-R^2\right)\sigma_{ab}^2}{\overline{X}} \tag{28}$$

where, $\overline{AB}$ is the average of the holding period measure, $\overline{PB}$ is the average of the selection period measure, $\overline{X}$ is the average of the selection period measure, *B* is the slope coefficient of the regression when $\overline{PB}$ is regressed on $\overline{AB}$, σ_{pb}^2 stands for the variance of the selection period measure, R^2 is the coefficient of the determination of the regression and finally σ_{ab}^2 is the variance of the holding period measure.

According to Mincer and Zarnowitz (1969), the first determinant is the bias, which arises of forecasting holding period measure with the one from selection period. The second term measures the inefficiency or tendency of the holding period measure to be larger (smaller) than the selection period measure at low (high) values of selection period measure. The last term measures the residual variance of the whole test. An MSE of zero would indicate that the selection period measure predicts the holding period measure with perfect accuracy.

4.4 Analysis methods for performance persistence

Finally, measures of performance persistence try to identify to what extent fund performance during one period continues during the following period. Persistence in performance can be studied e.g. as follows (e.g. Jan and Hung, 2004):

[24] Basically this is the same test than employed by Huij and Verbeek (2003).

1. Group funds based on the previous year performance (selection period).
2. Hold the funds over the subsequent period (holding period).
3. Compare the funds for performance over that subsequent period.

If the funds show persistence in performance, active fund selection based on past performance may be of interest to individual investors. On the other hand, if there is no sign of persistence past information would have no value for investors.

In this study, we explore performance persistence at one-year frequency meaning that both our selection and holding periods equals one year. This can be justified with a couple of arguments. Firstly, investors tend to evaluate fund performance over annual periods. Secondly, our sample is relatively short so that lengthening the periods e.g. to two years would not make sense. On the other hand, also the previous empirical studies suggest that persistence in performance is stronger and more prevalent for shorter measurement horizons.

4.4.1 Spearman rank correlation test

To investigate the persistence in fund performance we first follow the methodology employed, e.g. by Allen and Tan (1999) and apply the Spearman rank correlation test. The correlation coefficient can be given as follows:

$$\rho_s = 1 - \frac{6\sum D^2}{n\left(n^2 - 1\right)} \tag{29}$$

where ρ_s is the rank correlation coefficient, D is the difference between fund's selection period and holding period ranks and n denotes the number of funds.

The statistical significance of the correlation coefficient can be measured using the t-test since the correlation coefficient follows asymptotically a t-Student distribution. It can be given as follows:

$$t = \rho \left[\frac{(n-2)}{(1-\rho)^2} \right]^{1/2} \tag{30}$$

where ρ is the correlation coefficient and n is the number of funds. The correlation coefficient ρ_s always assumes values between -1 and 1. The zero hypotheses assumes that there is no persistence in performance and the past return information has no value in predicting the future performance. Therefore, the correlation coefficient should be close to zero. If the fund rankings show persistence, the correlation coefficient should be positive and statistically significant. On the other hand, statistically significant negative correlation coefficient would indicate performance reversal and would support investors to employ contrarian investment strategy. In this case investors would sell mutual funds which had performed well in the past and replace them with the funds that had performed poorly in the past. (Casarin et al. 2007; Sauer 1997)

4.4.2 Cross-sectional regression

An alternative methodology to study performance persistence is employed e.g. by Bollen and Busse (2005). Using the cross-sectional regression funds' holding period performance estimates are regressed on the performance estimates from selection period as follows:

$$Perf_{i,t} = \alpha + \beta Perf_{i,t-1} + \varepsilon_{i,t}, i = 1,2,..., N \tag{31}$$

where $Perf_{i,t}$ is the holding period performance estimate of fund *i*, α is the intercept, β is the slope coefficient, $Perf_{i,\ t-1}$ is the selection period performance estimate of fund *i* and $\varepsilon_{i,t}$ is the error term.

Again stated formally, the null hypothesis is that performance on the subsequent period is independent of the prior period performance. It can be presented as follows:

$$H_0 : \beta = 0$$

And the alternative hypothesis that posits existence of some relationship between past and future performance can be stated as follows:

$$H_1 : \beta \neq 0$$

Positive estimates for slope coefficients with significant t-statistic would implicate that past performance predicts the performance on the following period. On the other hand, negative coefficients would implicate again rather performance reversal. Besides the statistically significant slope coefficient, high adjusted R^2 would implicate strong explanatory power of the future performance.

4.4.3 Stacked return method

Compared with the two previous tests, the stacked return method can be considered more an investment strategy than a statistical test for detecting the pattern of performance persistence. To employ the stacked return-method we can follow a three-step procedure described for example as follows:

1. We form top and bottom quartile portfolios such as top-5/bottom-5 portfolios based on the rankings of the each selection period OLS three-factor alpha.

2. Then, we calculate the holding period return for these portfolios every week until the whole investigation period is covered, creating a weekly stacked-return time series.
3. Finally we estimate the performance of each portfolio throughout the holding period.

Moreover, following the three-step procedure described above we also form identical stacked return portfolios based on the on the Bayesian alpha rankings such as Sharpe ratio and the modified Sharpe ratio rankings. This allows us to study how the performance of these portfolios changes when different performance measures for the selection period are used.

To examine the performance persistence we compare the alphas created by the top and bottom portfolios and statistical significance of their difference. The statistical significance of differences between the top and bottom portfolio alphas can be tested employing the Welch's t-test suggested e.g. by Dixon and Massey (1968). However, in this case the test statistic is calculated for the top and bottom three-factor alpha spread, which can be given as follows:

$$t = \frac{\alpha_{top} - \alpha_{bottom}}{\sqrt{SE_{top}^2 + SE_{bottom}^2}} \tag{32}$$

where, α_{top} is the alpha of the top portfolio, α_{bottom} is the bottom portfolio alpha and *SE* stands for the standard errors of the stacked-return.

Respectively, the degrees of freedom for the previous test can be obtained using the following formula:

$$v = \frac{\left(SE_{top}^2 + SE_{bottom}^2\right)^2}{\frac{SE_{top}^4}{v_1} + \frac{SE_{bottom}^4}{v_2}} \tag{33}$$

where, SE is the standard error for the top and bottom portfolios, V_1 and V_2 are the degrees of freedom determined on the basis of the number of time-series returns for top and bottom portfolios *(v = n-1)*, *n* stands for the number of observations. If the funds exhibit persistence performance, the top (bottom) funds from the selection period should remain top (bottom) performers on the holding period as well. This kind of performance should lead to positive and statistically significant alpha spread. In addition, it would give some implications where in particular possible persistence is concentrated.

Moreover, in order to prevent spurious results rising from possible model misspecification, we also measure the holding period performance of these portfolios with the Sharpe ratio. In addition, rather similar statistical test with the one presented in Equation 32 to investigate the statistical significance in difference of two Sharpe ratios is suggested by Jobson and Korkie (1981) and modified by Memmel (2003). The test statistic between two portfolios (*i,n*) can be given as follows:

$$Z = \frac{\hat{\sigma}_n \hat{\mu}_i - \hat{\sigma}_i \hat{\mu}_n}{\sqrt{\hat{\theta}}} \tag{34}$$

where, $\hat{\sigma}_n$ and $\hat{\sigma}_i$ are the standard deviations of the portfolios *i* and *n*. $\hat{\mu}_i$ and $\hat{\mu}_n$ denote the mean return of the same portfolios. Finally, $\hat{\theta}$ stands for the asymptotic variance, which can be given as follows:

$$\theta = \frac{1}{T}\left(2\sigma_i^2\sigma_n^2 - 2\sigma_i\sigma_n\sigma_{in} + \frac{1}{2}\mu_i^2\sigma_n^2 + \frac{1}{2}\mu_n^2\sigma_i^2 - \frac{\mu_i\mu_n}{\sigma_i\sigma_n}\sigma_{in}^2\right) \tag{35}$$

where, T is the number of observations, σ_i and σ_n are the standard deviations of the portfolios i and n, σ_{in} is the covariance between the portfolios and μ_i and μ_n are the average return of the portfolios i and n. A statistically significant Z-statistic would implicate the rejection of the equal risk-adjusted performance on the holding period and would suggest that the portfolio with the higher Sharpe ratio outperforms the other portfolio. For example, in this case, the top portfolio wins the bottom portfolio.

5 EMPIRICAL RESULTS

5.1 Accuracy of the Bayesian alphas

To estimate the Bayesian alphas we replicate the methodology suggested by Huij and Verbeek (2003; 2007). Each year we formulate the priors for each fund using the whole cross-section of the funds in each year. Then, using the shrinkage procedure, we estimate the posterior parameters for each fund for holding and selection periods. For each fund we run 6 iterations for each time period. First, to investigate the suitability and degree of convergence of the iterative empirical Bayesian process, we calculate the coefficient of variation for the OLS three-factor alphas and the Bayesian alphas for both the selection and holding periods.[25] These results are presented in Appendix 3 and 4.

In general, the results indicate the functionality of the Bayesian shrinkage procedure. The coefficient of the variation seems to be relatively high for the OLS three-factor alphas for both on the selection and on the holding periods. However, already after the first iteration the coefficient of variation decreases significantly in each case, implying the convergence of estimates. This suggests that the shrinkage procedure works well in practice.

Before proceeding any further with the persistence analysis, it is of our interest to study to what extend the Bayesian alphas offer better and more accurate estimates for fund performance compared with the OLS three-factor alphas. To measure accuracy we employ the normalized mean squared error test (MSE) presented in Equation 28. Table 3 presents the results of the accuracy test. As can be seen from the Panel A, the MSE of the OLS three-factor estimates seem to be higher compared with those of

[25] The coefficient can be calculated dividing the standard deviation of the observations by the mean of the observation. Basically it measures how much the observations vary around the mean. However, we use the absolute values of the estimates so that positive and negative estimates would not neutralize each others.

the Bayesian estimates. Even though the difference does not seem to be so high, the results indicate better accuracy of the Bayesian alphas. Another interesting feature is that the MSE of the Bayesian alphas seem to decrease when more iterations are run. For example, the mean squared error for the 3rd iteration alphas is around 22% smaller compared with the OLS alphas. This implies that the Bayesian selection period alphas estimate around 22% more accurate the holding period OLS alphas than the OLS selection period alphas. On the other hand, the results show that accuracy does not increase between the 3rd and 5th iteration alphas. Moreover, the results also show that the bias component is zero in each case and the inefficiency component reduces gradually for the Bayesian estimates.

Furthermore, Panel B in table 3 shows the mean squared errors, when the Bayesian alphas are used for both on the selection and on the holding periods. Interestingly, now the MSE suggests even better accuracy for the Bayesian alphas. For example, when the Bayesian 3rd iteration alphas from both selection and the holding period are used, the MSE is around 60% smaller than the MSE for the OLS three-factor alphas. These results suggest that employment of the Bayesian alphas on both on the selection and on the holding periods would increase the measurement accuracy. Therefore, based on the results, it seems that the Bayesian estimates may have some advantage over the OLS estimates in fund performance evaluation.

Table 3. Accuracy of the OLS alphas and the Bayesian alphas.

Table 3 shows the normalized MSEs for the OLS three-factor alphas and the Bayesian alphas after 1st, 3rd, and 5th iterations. Panel A shows the MSEs when both the OLS and the Bayesian selection period alphas are used to explain the OLS holding period alpha. Panel B shows the MSEs when the OLS selection period alphas are used to explain the holding period OLS alphas and the Bayesian selection period alphas are used to explain the holding period Bayesian alphas. The normalized MSEs are also divided into three components. The percentual proportion of each component is presented below the absolute value.

Panel A

	OLS alphas	1st iteration alphas	3rd iteration alphas	5th iteration alphas
MSE	0.0089 *(100%)*	0.0076 *(100%)*	0.0070 *(100%)*	0.0070 *(100%)*
Bias	0.0000 *(0%)*	0.0000 *(0%)*	0.0000 *(0%)*	0.0000 *(0%)*
Inefficiency	0.0063 *(70%)*	0.0052 *(68%)*	0.0048 *(69%)*	0.0049 *(70%)*
Standard error	0.0026 *(30%)*	0.0024 *(32%)*	0.0022 *(31%)*	0.0021 *(30%)*

Panel B

	OLS alphas	1st iteration alphas	3rd iteration alphas	5th iteration alphas
MSE	0.0089 *(100%)*	0.0065 *(100%)*	0.0054 *(100%)*	0.0022 *(100%)*
Bias	0.0000 *(0%)*	0.0000 *(0%)*	0.0000 *(0%)*	0.0000 *(0%)*
Inefficiency	0.0063 *(70%)*	0.0051 *(79%)*	0.0048 *(87%)*	0.0017 *(78%)*
Standard error	0.0026 *(30%)*	0.0014 *(21%)*	0.0006 *(13%)*	0.0005 *(22%)*

On the other hand, it would be interesting to know how the Bayesian alphas predict future performance compared with OLS alphas. To provide information about the prediction power of the Bayesian alphas we run a couple of regressions where the OLS holding period alphas are regressed on the 1st, 3rd and 5th iteration Bayesian alphas. In other words, the Bayesian alphas from the selection period are used as independent variables and the OLS three-factor alphas from holding periods as dependent variables. Table 4 presents the results from these regressions.

Similarly to the results from the normalized MSE-test, the results based on the regression analysis suggest that the Bayesian alphas may have better prediction power for the holding period performance than the corresponding OLS 3-factor alphas. Interestingly, in each case, the slope is negative. However, the Bayesian alphas seem to explain better the variations in the holding period measure. This can be easily seen again from the last row, which presents the adjusted coefficient of determination of the regression models. The results show that the explanatory power arises when the Bayesian alphas are used in the regression. For example, for the OLS alphas the adjusted R^2 is around 8.5 percent when in the case of 1st iteration Bayesian alphas the adjusted R^2 increases to 15.6 percent. In addition, again the prediction power of the Bayesian alphas seems to increase when more iterations are run. For example, again when 3rd iterations are used in regression, the adjusted R^2 rises to around 23 per cent, being almost three times higher than in the case of the OLS regression.

Based on the previous analysis and the results from the normalized MSE test it seems that the Bayesian alphas provide better and more accurate estimates of future performance than the OLS estimates do. We consider that the 3rd iteration Bayesian alphas could provide the best estimate for the fund performance. On the other hand, one could argue that the 5th iteration alphas may provide even more accurate and better estimates. However, based on the MSE test there is no difference in holding period accuracy between the 3rd and the 5th iteration selection period Bayesian alphas. Moreover, based on the regression analysis, their prediction power does not significantly increase compared with the 3rd iteration alphas. Therefore, we believe that the 3rd iteration alphas would be suitable estimates for fund performance. This means that we will use the 3rd iteration Bayesian alphas for both the selection and the holding period when we estimate persistence in performance. Hence, from this moment one when we mention the Bayesian alphas we will always refer to the 3rd iteration Bayesian alphas.

Table 4. Prediction power of the alphas

Table 4 presents the results of the regression analysis where the OLS three-factor alphas from holding period are regressed on different selection period alphas. The first column shows the slope coefficient, its p-value and the adjusted coefficient of determination for each regression. The second column shows the results when OLS three-factor alphas from the holding period are regressed on the equivalent OLS three-factor alphas from the selection period. The third, fourth and fifth column shows the results when the OLS three-factor holding period alphas are regressed on the Bayesian selection period alphas after 1st, 3rd and 5th iterations.

	OLS alphas	1st iteration alphas	3rd iteration alphas	5th iteration alphas
Slope	-0.2607	-0.4261	-0.5942	-0.6740
p-value	0.0000***	0.0000***	0.0000***	0.0000***
Adj. R^2	0.0862	0.1559	0.2275	0.2686

*** Statistically significant at 1% level.

5.2 Rank correlations

Finally, to study the pattern of performance persistence we start with the Spearman rank correlation test. To apply the test, we use two-year subperiods. We first rank the funds based on their performance from the preceding one-year period (i.e. selection period), and then we rank them again based on the subsequent one-year period (i.e. holding period) performance. This process is then repeated so that the full sample period (2002-2007) is covered.

The results are presented in Table 5. First, the Spearman rank correlation-test based on the Bayesian alpha rankings show strong negative correlation (reversal pattern) for the first two subperiods. For example, for the first subperiod (2002-2003) the correlation coefficient is -0.620 and for the second subperiod (2003-2004) -0.717, respectively. In both cases, the correlation coefficients are also highly statistically significant at 1 % risk level. Moreover, the corresponding test for the Fama-French alphas show some slight negative correlation on the second subperiod (2003-2004), but statistical significance of the coefficient is close to zero. On the other hand, it is interesting that the Sharpe ratio and the downside risk-based Sharpe ratio indicate rather performance persistence than performance reversal, especially for the first subperiod. However, none of the rank correlations based on these measures reaches the level of statistical significance.

Finally, when we take a look at the two most previous subperiods (2005-2006 and 2006-2007), in general, the results show positive correlation in fund rankings. Again, the strongest implications are found using the rankings based on the Bayesian alpha. For both subperiods the correlation coefficients are substantially positive and statistically significant at 1 % risk level. In addition, the downside-risk based measures and the Jensen alpha indicate positive and statistically significant correlation for the last two subperiods. Moreover, the Sharpe Ratio based rankings show positive correlation for the last subperiod. However, all these measures indicate significantly lower rank correlation compared with Bayesian alphas. In general, it seems that the Sharpe ratio and its variations produce rather similar rankings. Another interesting feature is that the Fama-French alphas do not show any signs of persistence on the last two subperiods. Interestingly, the Fama-French rankings produce rather different results compared to the Bayesian alphas although the Bayesian alphas are based on them.

Using solely statistical arguments, it seems that in general the Bayesian alphas are able to detect performance persistence or performance reversal quite well. In four cases out of five they indicate statistically significant correlation coefficient. Moreover, when comparing the significance levels of the correlation coefficients the results show that the

Bayesian alphas may have some advantage over other measures. In four cases out of five, the Bayesian alphas get statistically most significant coefficients.

Based on the Spearman correlation test, five measures out of six implicate statistically significant correlation between fund rankings for some of the subperiods. The findings are strongest during the last two intervals, when four measures suggest statistically significant positive correlation. This suggests performance persistence among the funds. On the other hand, it is good to keep in mind that its degree seems to vary depending on the performance measure employed. It seems that prior information may have some value for investors in predicting performance on the subsequent period. However, it may be difficult to use it since it seems that e.g. persistence appears randomly.

Table 5. Spearman rank correlation-test

Table 5 presents the results of the Spearman correlation test. The Spearman correlation coefficient is calculated for the Sharpe ratio, the downside deviation-based Sharpe ratio (D-Sharpe), the modified Sharpe, the Jensen alpha, the Fama-French alpha and for the Bayesian alpha. The Spearman correlation coefficient and its t-value are presented for each performance measure between each selection and holding period.

		Sharpe ratio		**D-Sharpe**		**Modf. Sharpe**		**Jensen alpha**		**Fama-French alpha**		**Bayesian alpha**	
		Rank correlation	t-stat	Rank correlation	t-stat	Rank correlation	t-stat	Rank correlation	t-stat	Rank correlation	t-stat	Rank correlation	t-stat
Selection period	***Holding period***												
2002	2003	0.323	1.564	0.334	1.629	0.252	1.198	0.089	0.413	0.148	0.686	-0.620	-3.626***
2003	2004	0.164	0.863	0.198	1.030	0.133	0.685	0.030	0.157	-0.024	-0.127	-0.717	-5.239***
2004	2005	0.270	1.485	0.292	1.518	0.145	0.776	-0.048	-0.259	0.002	0.011	0.103	0.547
2005	2006	0.253	1.548	0.306	1.903**	0.287	1.778*	0.279	1.725*	0.023	0.140	0.782	7.432***
2006	2007	0.277	1.851*	0.258	1.711*	0.340	2.232**	0.293	1.963*	0.200	1.312	0.639	5.317***

* Statistically significant at 10% level.
** Statistically significant at 5% level.
*** Statistically significant at 1% level.

5.3 Cross-sectional regression

To further test the pattern of performance persistence we run cross-sectional regressions where the selection period performance is regressed on the holding period performance using each performance measure used in this study. First, we run the regressions using the observations from the full sample period. Afterwards, we divide the data into shorter subperiods.

5.3.1 Full sample period

The results of the cross-sectional regression for the full sample period are presented in Table 6. Interestingly, the results indicate statistically significant negative slope coefficient for each performance measure, implying no evidence of performance persistence. However, the results also reveal that past information may have some value. The negative slope coefficients suggest that the funds, which were winners (losers) in the selection period did not remain as winners (losers) on the holding period i.e. performance reversal.

Again, when comparing the different performance measures, the results show that the Bayesian alphas seem have the strongest explanatory power. The adjusted R^2 in the Bayesian regression (0.485) is significantly higher compared with the other regressions. This means that the holding period performance is highly dependent on the selection period performance. Moreover, the Bayesian alphas seem to reach the highest slope.

Fama-French alphas seem to also have some sort of explanatory power of the holding period performance since the adjusted R^2 is around 8.5 percent. However, there is still a huge difference compared with the Bayesian alphas. Also the slope coefficient remains significantly lower. When it comes to the other four measures, the results indicate that they are not so sensitive to detect persistence. Despite the fact that each of the remaining four measures produce statistically significant slope coefficient, in each case the adjusted R^2 remains really low. This implies that they are not able to predict the holding period performance so well.

Table 6. Cross-sectional regression for the full period

Table 6 presents the results from the cross-sectional regression analysis for the full sample period. In table are shown the slope coefficient, its p-value and the adjusted coefficient of determination for each performance measure regression. N denotes the number of observations in each regression. Each regression is run so that each holding period measure is regressed on the equivalent selection period measure.

Performance Measure	N	Slope	p-value	Adjusted R^2
Sharpe ratio	166	-0.193	0.004***	0.043
D-Sharpe	166	-0.108	0.078*	0.013
Modified Sharpe	166	-0.193	0.004***	0.044
Jensen alpha	166	-0.127	0.056**	0.016
Fama-French alpha	166	-0.260	0.000***	0.086
Bayesian alpha	166	-0.588	0.000***	0.485

* Statistically significant at 10% level.
** Statistically significant at 5% level.
*** Statistically significant at 1% level.

5.3.2 Subperiod analysis

After the analysis for the full sample period, we divide the data into shorter periods and repeat the cross-sectional regression. The results are presented in Table 7. Interestingly, now the results are somewhat similar with the ones from the Spearman rank correlation test. Again, the Bayesian alphas exhibit negative slopes for the first two subperiods (2002-2003 and 2003-2004) indicating reversal pattern in fund performance. Especially on the first subperiod, the slope coefficient -12.44 show extremely high reversal pattern between selection and holding period performance. In both cases, the slope coefficients are also highly significant in statistical sense, since they reach 1 percent risk level. The explanatory power of the Bayesian alphas is also really high, since the adjusted coefficient of determination is at least 0.49.

When it comes to the last two subperiods the Bayesian alphas show again positive and statistically significant slope coefficients implicating performance persistence. Moreover, based on the adjusted coefficients of determination the Bayesian alpha cross-sectional regression models seem to have the greatest explanatory power. On the last two subperiods the adjusted R^2 is at least 0.45 for the Bayesian alphas. The other measures indicate also performance persistence with statistically significant slope. In general, the modified Sharpe ratio seems to have good explanatory power, since the adjusted R^2 is at least 0.23 in both cases. Despite one exception, the modified Sharpe measure seems to detect the holding period performance better than the Sharpe ratio and its variations on the last two subperiods. In addition, the selection period Jensen alphas seem to explain their holding period counterparts fairly well, especially on the last subperiod, when the adjusted R^2 is almost 0.26. Moreover, the Fama-French alphas seem to have good explanatory power of the holding period performance on the last subperiod since the adjusted R^2 is around 0.17.

Besides the analysis presented before, Kosowski et al. (2007), propose the employment of the t-statistic of the alpha estimates as a performance measure. They argue that the t-statistic has some statistical advantages, since it scales the alpha by its standard error. Therefore, we run these additional regressions (not presented) for the Jensen alpha and for the Fama-French alpha for both the full sample period and for the subperiods. In general, the results from these regressions were somewhat similar compared with the regressions where the alphas were used as input variables. However, an interesting finding was that when the t-values were used as input variables, the Fama-French alpha exhibited a positive and statistically significant slope coefficient also for the second-last subperiod (2005-2006) and the adjusted coefficient of determination was around

10%. This finding is contrast to that when the Fama-French alphas are used as regression input variables.

On the other hand, such as in the Spearman correlation test, the Sharpe ratio and the downside risk-based Sharpe ratio indicates rather performance persistence than reversal for the first subperiod. Moreover, the Fama-French alphas do not detect any reversal although the Bayesian alphas are based on the Fama-French alphas. To analyze possible reasons for differing results we calculate the Spearman rank correlations between these performance measures on the same selection and the holding period. The results are presented in Appendix 5 and 6. Based on the results it seems that different performance measures just lead to different rankings on the same period (selection or holding). For example, the rank correlation is slightly negative although not statistically significant between the Sharpe ratios and the Bayesian alphas in two cases (selection period 2003 and holding period 2003). Moreover, although the rank correlation between the Fama-French rankings and the Bayesian rankings is positive, it is not so strong in each case. Obviously, this leads to differing results between these measures.

Interestingly, the results of the cross-sectional regression analysis show again that the Bayesian alphas seem to be the most sensitive measure to detect fund performance on the subsequent period. For the full sample period they indicate the strongest evidence of performance reversal and compared with the other performance metrics, the Bayesian alpha-based regressions get also the highest adjusted R^2 on four out of five subperiods. Similarly, the statistical significance of the slope coefficients is significantly higher on four subperiods compared with the other performance metrics.

On the other hand, the overall results show that the time period employed in the cross-sectional regression may lead to very differing results. The full sample regression suggests rather performance reversal for each measure employed. On the other hand, after dividing the data into subperiods, the results are quite consistent with the results obtained from the Spearman rank correlation test, indicating some persistence in performance during the most recent years. Actually, based on the cross-sectional regression tests the persistence seems to be stronger during the last years since more methods suggest persistence. In general, one factor explaining the possible performance persistence during the recent years might be good overall performance of the Russian stock market over the last years (see Appendix 2). Obviously, it is easier for the funds to maintain their performance when the market is going up instead if the market is going down. On the other hand, it would be interesting to know what might have caused the performance reversal or is it just due to poor investment decisions.

Table 7. Cross-sectional regressions for subperiods

Table 7 presents the results from the cross-sectional regression when the whole observation period is divided to one-year selection (SP) and holding periods (HP). Each regression is run so that the holding period measure is regressed on the equivalent selection period measure. For each regression are shown the slope coefficient and adjusted coefficient of determination. The p-value for the slope coefficient is presented in the parenthesis below coefficient. N denotes the number of observations in each regression.

			Sharpe ratio		D-Sharpe		Modf. Sharpe		Jensen alpha		Fama-French alpha		Bayesian alpha	
		N	Slope	Adj. R^2	Slope	Adj. R^2	Slope	Adj. R^2	Slope	Adj. R^2	Slope	Adj. R^2	Slope	Adj. R^2
SP	***HP***													
2002	2003	24	0.401 *(0.113)*	0.058	0.507 *(0.112)*	0.073	0.272 *(0.327)*	0.000	-0.003 *(0.985)*	-0.048	0.008 *(0.973)*	-0.047	-12.44 *(0.000***)*	0.635
2003	2004	29	0.079 *(0.749)*	-0.034	0.076 *(0.713)*	-0.003	0.009 *(0.972)*	-0.003	-0.137 *(0.614)*	-0.028	-0.131 *(0.570)*	-0.025	-0.289 *(0.000***)*	0.491
2004	2005	31	0.232 *(0.516)*	-0.020	0.121 *(0.086*)*	-0.003	-0.006 *(0.986)*	-0.003	-0.106 *(0.781)*	-0.032	-0.105 *(0.771)*	-0.033	0.454 *(0.589)*	-0.024
2005	2006	38	0.197 *(0.013**)*	0.136	0.134 *(0.000***)*	0.361	0.244 *(0.000***)*	0.256	0.197 *(0.005***)*	0.178	0.105 *(0.135)*	0.036	0.501 *(0.000***)*	0.700
2006	2007	44	0.338 *(0.014**)*	0.116	0.418 *(0.008***)*	0.134	0.645 *(0.000***)*	0.235	0.339 *(0.000***)*	0.257	0.311 *(0.003***)*	0.174	0.496 *(0.000***)*	0.453

* Statistically significant at 10% level. ** Statistically significant at 5% level. *** Statistically significant at 1% level.

5.4 Stacked return analysis

Our last test for performance persistence is so-called stacked return time series analysis. Compared with the previous tests this test can be considered rather an investment strategy than traditional test for performance persistence. To be more specific, we study whether there is difference in holding period return if an investor keeps on investing in the past top performers compared with strategy of investing in the past poor performers. First, we study the returns of these portfolios using the top and bottom quartile funds on the selection period. Later we narrow our investment universe to include only the top five and bottom five funds. This allows us to study if the performance is more related to the few extreme funds.

5.4.1 Results for the OLS and the Bayesian portfolios

Panel A of Table 8 displays results for portfolios based on the OLS three-factor alpha rankings. As can be seen, both quartile portfolios earn basically equal positive annualized excess return. Around 30 percent annual returns indicate good performance for both portfolios. In addition, the annualized volatility is almost equal. When we look at the three-factor holding period alphas, for both portfolios the alpha is positive and statistically significant. However, the alphas indicate slightly better performance for the top quartile portfolio.

Equal excess return and almost equal alphas suggest that the past losers performed as well as the top funds on the holding period. However, holding period performance of these portfolios can be further studied using Equations 32 and 34. When we first evaluate the holding period performance with the Sharpe ratio, the Jobson-Korkie test statistic shows equal performance for both portfolios meaning no evidence of performance persistence. Moreover, when the three-factor alpha is used to evaluate the holding period performance, the alpha-spread test indicates that the top and bottom portfolio alphas are equal at the 95 % confidence level. Therefore, it seems that investing only in the best performing funds does not outperform the strategy of investing in the poorly performing funds.

Panel B presents the same results when the Bayesian alphas are used as selection period criterion. The bottom quartile portfolio produces slightly higher excess return compared with the top quartile portfolio but in both cases the returns are again substantially high. However, the variation in returns seems to considerably larger for the bottom portfolio. Moreover, both portfolios seem to earn positive and statistically significant three-factor alphas, the one being now higher for the top portfolio.

However, there is no difference in holding period performance of these portfolios. Firstly, when the performance is measured by means of the Sharpe ratio, the test statistic from the Jobson-Korkie test indicates equal performance for both portfolios. In addition, the alpha spread between the portfolios is far away of being statistically significant. Therefore, neither Bayesian alpha based rankings seem to be able to detect persistence among these funds.

Table 8. Performance comparison of top and bottom quartile portfolios based OLS and Bayesian rankings

Table 8 presents average annual return, volatility, the Sharpe ratio, alphas for quartile portfolios (Q1 indicates top quartile and Q4 bottom quartile, respectively) from the 2003-2007 holding period. Moreover, the performance differences are shown by z-statistic of the Jobson-Korkie test and by alpha spread test (significance levels in parenthesis). The holding period alphas are estimated using the Fama-French 3-factor model. Panel A shows the results based on the use of Fama-French 3-factor as the selection criterion for quartile portfolios. Correspondingly, Panel B shows the results based on the Bayesian alphas as a selection criterion.

Panel A

Portfolio	Annual excess return	Annual volatility	Sharpe ratio	z-stat.	Annual alpha	Alpha spread	Adj. R^2
OLS Q1	30.48%	17.84%	0.236	0.049 (0.960)	18.21% (0.004***)	0.55% (0.960)	0.415
OLS Q4	30.50%	18.00%	0.234		17.66% (0.003***)		0.481

Panel B

Portfolio	Annual excess return	Annual volatility	Sharpe ratio	z-stat.	Annual alpha	Alpha spread	Adj. R^2
Bayes Q1	28.28%	15.72%	0.249	0.927 (0.353)	17.16% (0.002***)	1.54% (0.892)	0.387
Bayes Q4	29.40%	20.39%	0.199		15.60% (0.021**)		0.469

** Statistically significant at 5% level
*** Statistically significant at 1% level

Table 9 presents the results after narrowing the investment universe to the Top 5 and the Bottom 5 funds from the selection periods. Again, when first looking at the OLS portfolios we can see the results are almost identical with the ones from the quartile portfolios. The excess returns and the standard deviations are basically the same. Neither the three-factor alphas for the holding period change dramatically. Therefore, it seems that the holding period return on the top portfolio does not change when only the top 5 funds from the ranking period are selected in the portfolio. This indicates that there is no real difference in holding period performance between the top 5 funds and the top quartile funds from the selection period. Obviously, the case is same with the bottom performers from the ranking period. It seems that despite the poor performance on the ranking period, the poor performers are able to improve their performance on the holding period. Therefore, neither the Jobson-Korkie test and nor the alpha spread test indicate performance persistence for the top 5 or the bottom 5 funds. Both test statistics are insignificant, implying that persistence does not exist.

In the case of the Bayesian alphas, the results are also very comparable with ones based on the quartile portfolios. However, now the excess return spread is even higher in favour of the bottom 5 portfolio. Interestingly, both portfolios produce exactly an equal alpha on the holding period. Finally, when we look at the statistical tests they show that there is no difference in performance of these portfolios on the holding period. Again, the Jobson-Korkie test statistic for the Sharpe ratios is insignificant. Moreover, the alpha spread test implicates equal performance. Almost at 100 percent probability we can say that the holding period alpha spread between the top 5 and bottom 5 portfolios is zero.

The overall results show that in each case (Bayes or OLS) the bottom portfolio alphas are quite close or equal to the one for the top portfolios. This suggests performance reversal among the bottom funds. After poor performance on the ranking period, they can improve significantly their performance on the holding period. This strong reversal effect may also cause that none of the statistical tests detected performance persistence among the top or bottom funds. These findings show also some similarity with results from the cross-sectional regression for the full sample period since it also implicated performance reversal among the full cross-section. On the other hand, the Spearman correlation test and the cross-sectional regression for the subperiods suggested performance persistence for the last two years of observation period. Therefore, it seems again that using subperiod analysis may lead to very differing results compared with the results when the full sample period of data is used.

Table 9. Performance comparison of top and bottom 5 portfolios based OLS and Bayesian rankings

Table 9 presents average annual return, volatility, the Sharpe ratio, alphas for top and bottom 5 portfolios from the 2003-2007 holding period. Moreover, the performance differences are shown by z-statistic of the Jobson-Korkie test and by alpha spread test (significance levels in parenthesis). The holding period alphas are estimated using the Fama-French 3-factor model. Panel A shows the results based on the use of Fama-French 3-factor as the selection criterion for the portfolios. Correspondingly, Panel B shows the results based on the Bayesian alphas as a selection criterion.

Panel A

Portfolio	Annual excess return	Annual volatility	Sharpe ratio	z-stat.	Annual alpha	Alpha spread	Adj. R^2
OLS Top5	30.87%	17.68%	0.242	0.290 (0.771)	19.16% (0.002***)	0.93% (0.935)	0.368
OLS Bot5	30.85%	18.61%	0.230		18.23% (0.004***)		0.449

Panel B

Portfolio	Annual excess return	Annual volatility	Sharpe ratio	z-stat.	Annual alpha	Alpha spread	Adj. R^2
Bayes Top5	27.44%	15.86%	0.240	0.606 (0.544)	16.63% (0.005***)	0.00% (0.999)	0.335
Bayes Bot5	30.41%	20.82%	0.202		16.63% (0.017**)		0.463

** Statistically significant at 5% level
*** Statistically significant at 1% level

5.4.2 Results for the Sharpe and modified Sharpe portfolios

Moreover, we construct the similar stacked return series using the selection period rankings based on the Sharpe ratio and the modified Sharpe ratio. This allows us to study e.g. how these extreme portfolios perform or how the pattern of performance persistence (or reversal) changes when the funds are ranked based on performance metrics employing the total and the downside risk of the funds. As Panel A in Table 10 reveals, both portfolios constructed using the Sharpe ratio rankings have earned high returns. As one could expect, the top portfolio has earned slightly higher holding period return compared with significantly more volatile bottom portfolio. In addition, the OLS three-factor holding period alpha is positive and statistically significant for both portfolios being higher for the top quartile portfolio. Moreover, when holding period performance of these portfolios is evaluated with the Sharpe ratio, the Jobson-Korkie test statistic is statistically significant at 5 percent level, indicating performance persistence for top and bottom quartile funds. On the other hand, based on the alpha spread test it seems that there is no difference in holding period performance of these portfolios since the test statistic is insignificant.

The results are almost identical for the modified Sharpe rankings presented in Panel B. As we can see, the average return and the standard deviation of the top and bottom portfolios are almost equal compared with the results using the Sharpe ratio rankings. Also the three-factor alphas are at the same level. Again, when the modified Sharpe ratio is used to measure performance on the selection and the holding period, the Jobson-Korkie test statistic shows statistically significant persistence at 5 percent level. In contrast, the alpha spread is statistically insignificant implying no evidence of persistence.

Table 10. Performance comparison of top and bottom quartile portfolios based on Sharpe and modified Sharpe rankings

Table 10 presents average annual return, volatility, the Sharpe ratio, alphas for quartile portfolios (Q1 indicates top quartile and Q4 bottom quartile, respectively) from the 2003-2007 holding period. Moreover, the performance differences are shown by z-statistic of the Jobson-Korkie test and by alpha spread test (significance levels in parenthesis). The holding period alphas are estimated using the Fama-French 3-factor model. Panel A shows the results based on the use of the Sharpe ratio as the selection criterion for quartile portfolios. Correspondingly, Panel B shows the results based on the modified Sharpe ratio as a selection criterion.

Panel A

Portfolio	Annual excess return	Annual volatility	Sharpe ratio	z-stat.	Annual alpha	Alpha spread	Adj. R^2
Sharpe Q1	30.50%	14.74%	0.287	2.562 (0.010**)	19.02% (0.000***)	3.61% (0.751)	0.459
Sharpe Q4	29.88%	20.83%	0.199		15.41% (0.017**)		0.535

Panel B

Portfolio	Annual excess return	Annual volatility	Mod. Sharpe ratio	z-stat.	Annual alpha	Alpha spread	Adj. R^2
Modf.Sh. Q1	30.58%	15.11%	0.263	2.569 (0.010**)	19.35% (0.000***)	4.00% (0.729)	0.449
Modf.Sh. Q4	29.92%	21.07%	0.183		15.35% (0.019**)		0.529

** Statistically significant at 5% level
*** Statistically significant at 1% level

Table 11 presents again the results when both portfolios are constructed using the top and bottom five funds on the selection period. Panel A for the Sharpe ratio portfolios indicate that the results do not significantly change when only the few extreme funds are included in the return series. However, now the return on bottom portfolio increases and it beats the corresponding top portfolio. On the other hand, the top portfolio earns a higher three-factor alpha. When the holding period performance is evaluated on the basis of the Sharpe ratio, the Jobson-Korkie test statistic is again statistically significant at 5 percent level, indicating performance persistence. However, the results are again mixed since the alpha spread test statistic remains insignificant. Finally, Panel B of Table 11 displays the same results for the modified Sharpe based portfolios. As can be seen, the results are almost parallel with the ones from quartile portfolios. When it comes to performance persistence, the Jobson-Korkie test statistic is statistically significant at 5 percent level. This implicates that it pays off to invest in the top portfolio. On the other hand, again the alpha spread test finds no evidence of persistence since the alpha spread remains statistically insignificant.

Table 11. Performance comparison of top and bottom 5 portfolios based on Sharpe and modified Sharpe rankings

Table 11 presents average annual return, volatility, the Sharpe ratio, alphas for top and bottom 5 portfolios from the 2003-2007 holding period. Moreover, the performance differences are shown by z-statistic of the Jobson-Korkie test and by alpha spread test (significance levels in parenthesis). The holding period alphas are estimated using the Fama-French 3-factor model. Panel A shows the results based on the use of the Sharpe ratio as the selection criterion for the portfolios. Correspondingly, Panel B shows the results based on the modified Sharpe ratio as a selection criterion.

Panel A

Portfolio	Annual excess return	Annual volatility	Sharpe ratio	z-stat.	Annual alpha	Alpha spread	Adj. R^2
Sharpe Top5	31.54%	14.71%	0.297	2.017 (0.043**)	21.46% (0.000***)	3.29% (0.772)	0.355
Sharpe Bot5	32.09%	20.74%	0.214		18.17% (0.007***)		0.478

Panel B

Portfolio	Annual excess return	Annual volatility	Mod. Sharpe ratio	z-stat.	Annual alpha	Alpha spread	Adj. R^2
Modf.Sh. Top5	30.47%	14.77%	0.275	2.148 (0.031**)	20.38% (0.000***)	3.99% (0.728)	0.348
Modf.Sh. Bot5	30.59%	21.16%	0.187		16.39% (0.017**)		0.494

** Statistically significant at 5% level
*** Statistically significant at 1% level

Interestingly, the results for the Sharpe ratio and the modified Sharpe ratio-based portfolios seem to differ from the OLS alpha and the Bayesian alpha-based portfolios. The results indicate performance persistence among the top and bottom funds when Sharpe ratio and the modified Sharpe ratio are used to measure performance on the selection and on the holding period. On the other hand, when holding period performance is evaluated by means of the three-factor alpha and the alpha spread test no evidence of persistence is found. Therefore, part of the results would implicate that performance persistence would concentrate on the top and bottom funds. On the other hand, part of the results would implicate that persistence would rather exist among the middle performers and reversal concentrates on the bottom funds. Hence, it seems that partly the degree of persistence and its existence varies depending on the model employed.

Although based on the Stacked return analysis it is difficult to make any unanimous conclusion about performance persistence the overall results clearly show that the equity funds investing in Russia have produced good profits during the last years. For example in each case the portfolio excess return seems to be considerably high. Hence, we conduct an additional analysis comparing the performance of all these portfolios with the overall performance of the Russian stock market. The results are reported in Table 12. As can be seen, in 11 cases out of 16 the stacked return portfolios outperform the market portfolio measured by the means of the Sharpe ratio and the Jobson-Korkie test. Moreover, in remaining five cases the test statistic is quite close of being statistically significant. The market portfolio seems to have significantly higher volatility but also lower annual excess return. This leads to better performance of the stacked return portfolios and also to statistically significant test statistics. Therefore, although these funds have not necessarily exhibited persistence in performance, in general they have been able to perform better compared with the market performance.

Table 12. Stacked return portfolio performance comparison with the RTS Index

Table 12 shows the performance comparison between the stacked return portfolios and the RTS Index. Table presents average annual return, volatility and the Sharpe ratio for each portfolio from the 2003-2007 holding period. Moreover, the performance differences are shown by z-statistic of the Jobson-Korkie test and its significance level. In each case, stacked return portfolio performance is compared with the RTS Index. Panel A shows the performance comparison with the top and bottom quartile portfolios. Panel B shows the performance comparison with the top 5 and bottom 5 portfolios, respectively.

Panel A

Portfolio	Annual excess return	Annual Volatility	Sharpe ratio	Z-stat.	Sign.
RTS Index	27.56%	31.02%	0.123		
OLS Q1	30.48%	17.84%	0.236	2.12	0.033**
OLS Q4	30.50%	18.00%	0.234	2.25	0.024**
Bayes Q1	28.28%	15.72%	0.249	2.25	0.024**
Bayes Q4	29.40%	20.39%	0.199	1.54	0.123
Sharpe Q1	30.50%	14.74%	0.287	3.06	0.002***
Sharpe Q4	29.88%	20.83%	0.199	1.64	0.101
M. Sharp. Q1	30.58%	15.11%	0.263	2.98	0.002***
M. Sharp. Q1	29.92%	21.07%	0.183	1.59	0.110

Panel B

Portfolio	Annual excess return	Annual Volatility	Sharpe ratio	Z-stat.	Sign.
RTS Index	27.56%	31.02%	0.123		
OLS Top5	30.87%	17.68%	0.242	2.10	0.034**
OLS Bot5	30.85%	18.61%	0.230	2.08	0.036**
Bayes Top5	27.44%	15.86%	0.240	1.95	0.050**
Bayes Bot5	30.41%	20.82%	0.202	1.57	0.114
Sharpe Top5	31.54%	14.71%	0.297	2.95	0.003***
Sharpe Bot5	32.09%	20.74%	0.214	1.85	0.063*
M. Sharp.Top5	30.47%	14.77%	0.275	2.73	0.006***
M. Sharp. Bot5	30.59%	21.16%	0.187	1.60	0.107

* Statistically significant at 10% level.
** Statistically significant at 5% level.
*** Statistically significant at 1% level.

6 CONCLUSIONS

This thesis examined the performance persistence of European equity funds investing in the Russian stock markets during the time period from 2002 to 2007. Firstly, the general purpose of this study was to fill the existing gap in the financial literature between studies concentrating on the developed and on the emerging mutual fund markets. Secondly, using traditional and innovative measures, our empirical objective was to compare the results obtained by using different performance metrics and methodologies to find out whether the performance persistence truly exists.

We evaluated fund performance using several performance metrics. We employed the Sharpe ratio, the downside risk Sharpe ratio, modified Sharpe ratio, the Jensen alpha and the Fama-French 3-factor alpha. Moreover, we applied the iterative empirical Bayesian estimation for fund performance and we estimated the Bayesian alphas for each fund for the selection and the holding periods. Finally, to detect the relation between the selection and the holding period performance, we applied several methodologies.

First, we employed the Spearman correlation test. Second, we performed the cross-sectional regression to determine the prediction power of the selection period performance on the holding period performance. Third, we applied so-called stacked return portfolio analysis, which can also be understood as an investment strategy. Based on the selection period performance, we formed top and bottom portfolios of the funds and evaluated the performance difference of these extreme portfolios on the holding period by means of the Jobson-Korkie test and the three-factor alpha spread test. As a contribution to earlier analysis, this method allowed us to investigate where the performance persistence is concentrated in our sample.

We found evidence that the Bayesian alphas may have some advantage over the OLS three-factor alphas in performance estimation. Firstly, using the normalized mean squared error test we found that the 3rd iteration Bayesian alphas were around 22 percent more accurate than the OLS alphas in estimating the holding period performance. Secondly, our results suggested that the when the Bayesian alphas are used both on the selection and on the holding period, the estimation accuracy increases further.

When it comes to performance persistence, the results from the Spearman correlation test showed strong reversal pattern in performance during the first two subperiods (2002-2003 and 2003-2004) when the Bayesian alphas were used as a ranking criterion. For the last two subperiods (2005-2006 and 2006-2007) the Bayesian alphas, the downside deviation-based Sharpe and the modified Sharpe ratio, likewise the Jensen alpha indicated positive relation between the prior and the subsequent period performance. Moreover, the Sharpe ratio implicated persistence for the last subperiod. When persistence was detected by the means of the cross-sectional regression, each performance metrics indicated performance reversal for the full sample period. On the other hand, after dividing the data into subperiods, the results were somewhat identical with the ones from the Spearman correlation test. However, the cross-sectional analysis indicated slightly stronger persistence for the last two subperiods since severe methods were able to detect persistence. In general, based on the Spearman rank correlation test and the cross-sectional analysis the Bayesian alphas seemed to be the most sensitive measure to detect persistence or performance reversal.

Finally, when it comes to stacked return approach, our results were somewhat mixed. We found no evidence of performance persistence among the top and bottom funds when the portfolios were formed on the basis of the OLS three-factor alpha or the Bayesian alpha. However, our results suggested some performance reversal of the bottom funds.

Interestingly, when the same portfolios were formed using the Sharpe ratio and the modified Sharpe ratio as a selection period criterion and when the holding period performance of the same portfolios was examined by the means of same methods, the results showed statistically significant performance persistence for the top and the bottom funds. In contrast, when the holding period performance of the Sharpe and modified Sharpe-based portfolios was evaluated by the means of the three-factor alpha, the alpha spread test found no evidence of persistence.

Clearly, the results show that in general the equity funds investing in Russia have exhibited strong performance compared with overall market performance during the observation period. However, based on the findings it is difficult to make any unanimous conclusions whether performance persists among these funds or not. Obviously, we found some evidence of persistence especially during the most recent years. However, the results also suggested strong signs of performance reversal for the full sample period. Therefore, it seems that the degree and existence of persistence is dependent on time period used in the analysis and on the other hand, partly dependent on the methodology employed. Past information may have some value for investors but by picking e.g. only the top performers from the selection period would not necessarily lead to significantly superior investment strategy.

In general, our results show similarity with the previous studies. Firstly, our findings parallel with the ones obtained by Huij and Verbeek (2007). We also found that the Bayesian estimates are more accurate to estimate fund performance compared with the OLS estimates. On the other hand, due to limited studies on emerging market funds it is difficult to compare our results with the previous findings. However, our results show some slight similarity with ones obtained by Huij and Post (2008).

Despite the relatively extensive analysis, there are plenty of possibilities to further expand this thesis. When it comes to performance measurement, it would be interesting to consider additional explaining factors to the three-factor model. For example, to investigate how the commodity and currency exposures are related to fund performance. Obviously, also the Bayesian performance estimation needs more development and examination. For instance, what would be the optimal number of iterations and in respect to previous, what would be the optimal level of convergence of the estimates. Moreover, it would be of interest to perform comparative analysis using different shrinkage estimators such as Huij and Verbeek (2003) propose. As far as the performance persistence measurement is concerned, it would be interesting to study e.g. the relation of the fund age to performance persistence. In addition, it would be interesting to know to what extent the performance persistence can be explained by the managerial skill. This could be estimated using e.g. bootstrapping analysis.

REFERENCES

Allen, D. E. – Tan, M. L.: “A test of the persistence in the performance of UK managed funds”. *Journal of Business, Finance and Accounting*, 1999, vol. 26, no. 5, 559-593.

Babalos, W. – Caporale, M. G. – Kostakis, A. – Phillippas, N.: ”Testing for persistence in mutual fund performance and the ex post variation problem: Evidence from Greek market”. University of Piraeus, Greece, Working Paper, 2007, 1-41.

Baks, K. P. – Metric, A. – Watcher, J.: “Should investors avoid all actively managed mutual funds? A study in Bayesian performance evaluation”. *The Journal of Finance*, 2001, vol. 56, no. 1, 45-86.

Bekaert, G. – Erb, B. C. – Harvey, C. R. – Viskanta, T. E.: ”Distributional characteristics of emerging market returns and asset allocation”. *Journal of Portfolio Management*, 1998, vol. 24, no. 2, 102-116.

Blake, D. – Timmermann, A.: “Mutual fund performance: Evidence from the U.K”. *European Finance Review*, 1998, vol.2, no.1, 57-77.

Blake, C. R. – Morey, M. R.: “Morningstar ratings and mutual fund performance”. *The Journal of Financial and Quantitative Analysis*, 2000, vol. 35, no. 3, 451-483.

Bodie, Z. – Kane, A. – Marcus, A. J.: Investments (6th edition). USA: The McGraw-Hill Companies, Inc. 2005.

Bollen, N. P. – Busse, J. A.: “Short-term persistence in mutual fund performance”. *The Review of Financial Studies*, 2005, vol. 18, no. 2, 569-597.

Brown, S. J. – Goetzmann, W. N.: “Performance persistence”. *The Journal of Finance*, 1995, vol.1, no. 2, 679-698.

Brown, S. J. – Goetzmann, W. – Ibbotson, R. G. – Ross, S. A. : “Survivorship bias in performance studies”. *The Review of Financial Studies*, 1992, vol. 5, no. 4, 553-580.

Busse, J. A. – Irvine, P. J.: “Bayesian alphas and mutual fund persistence”. *The Journal of Finance*, 2006, vol. 61, no. 5, 2251-2288.

Carhart, M. M.: “On persistence in mutual fund performance”. *The Journal of Finance*, 1997, vol. 52, no. 1, 57-82.

Carhart, M. M. – Carpenter, J. N. – Lynch, A. W. – Musto, D. K.: “Mutual fund survivorship”. *The Review of Financial Studies*, vol. 15, no. 5, 1439-1463.

Carlson, R. S.: “Aggregate performance of mutual funds”. *The Journal of Financial and Quantitative Analysis*, 1970, vol. 5, no. 1, 1-32.

Casarin, R. – Pelizzon, L. – Loriana, A.: “Italian equity funds: Efficiency and performance persistence”. University of Venice, Italy, Working Paper, 2007, 1-24.

Chen, M. H.: “Risk and return: CAPM and CCAPM”. *The Quarterly Review of Economics and Finance*, 2003, vol. 43, no. 2, 369-393.

Cuthbertson, K.: Quantitative financial economics: Stocks, bonds and foreign exchange (4th edition). USA: John Wiley & Sons Inc. 2000.

Deaves, R.: “Data-Conditioning biases, performance, persistence and flows: The Case of Canadian Equity Funds”. *Journal of Banking & Finance*, 2004, vol. 28, 673-694.

DiMarzio, A. A. – Haire, D. B. – Ritter, T. J.: "An alternative perspective on investment performance". *Journal of Financial Planning*, 1993, 129-133.

Dixon W.J. – Massey, F. J.: Introduction to statistical analysis, international student edition. USA: McGraw-Hill Inc. 1969.

Droms, W.: "Hot hands, cold hands: Does past performance predict future returns. *Journal of Financial Planning*, 2006, 61-69.

Eftekhari, B. – Pedersen, C. – Satchell, S. E.: "On the volatility of measures of financial risk: an investigation using returns from European markets". *The European Journal of Finance*, 2000, vol. 6, no. 1, 18-38.

Elton, E. J. – Gruber, M. J. – Brown, S. J. – Goetzmann, W. N.: Modern Portfolio Theory and Investment Analysis (6th edition). USA: John Wiley & Sons, Inc. 2003.

Estrada, J.: "The Cost of Equity in Emerging Markets: A downside risk approach". *Emerging Markets Quarterly*, 2000, vol. 4, no. 3, 1-12.

Estrada, J.: "Systematic risk in the emerging markets: the D-CAPM". *Emerging Markets Review*, 2002, vol.3, no.4, 365-379.

Estrada, J.: "The cost of equity of Internet stocks: a downside risk approach". *The European Journal of Finance*, vol. 10, no. 4, 239-254.

Estrada, J.: "Downside risk in practise". *Journal of Applied Corporate Finance*, 2006, vol. 18, no. 1, 117-125.

Estrada, J.: "Mean-semivariance behaviour: Downside risk and capital asset pricing". *International Review of Economics and Finance*, 2007, vol. 16, no.2, 169-185.

Fama, E. F.: “Efficient capital markets: A review of theory and empirical work”. *The Journal of Finance*, 1970, vol. 25, no. 2, 383-417.

Fama, E. F. – French, K. R.: “The cross section of expected stock returns”. *The Journal of Finance*, 1992 vol. 47, no. 2, 427-465.

Fama, E. F. – French, K. R.: “Common risk factors in the returns on stocks and bonds”. *The Journal of Financial Economics*, 1993, vol. 33, no. 1, 3-56.

Favre, L. – Signer, A.: “The difficulties of measuring the benefits of the hedge funds”. *The Journal of Alternative Investment*, 2002, vol. 5, no. 1, 1-17.

Fletcher, J.: “The evaluation of the performance of UK American unit trusts”. *International Review of Economics and Finance,* 1999, no. 8, 455-466.

Grootveld, H. – Hallerbach, W.: “Variance vs. downside risk: Is there really that much difference?”. *European Journal of Operational Research*, 1999, vol. 114, 304-319.

Grinblatt, M. – Titman, S.: ”The persistence of mutual fund performance”. *The Journal of Finance*, 1992, vol. 47, no. 5, 1977-1984.

Hendricks, D. – Patel, J. – Zeckhauser, R.: “Hot hands in mutual funds: Short-run persistence of relative performance, 1974-1988“. *The Journal of Finance*, 1993, vol. 48, no. 1, 93-130.

Hill, R. C. – Griffiths, W. E. – Judge, G. G.: Undergraduate Econometrics (2nd edition). USA: John Wiley & Sons, Inc. 2001.

ter Horst, J. – Verbeek, M.: “Estimating short-run persistence in mutual fund performance”. *The Review of Economics and Statistics*, 2000, vol. 82, no. 4, 646-655.

ter Horst, J. – Nijman, T. E. – Verbeek, M.: “Eliminating look-ahead bias in evaluating persistence in mutual fund performance”. Journal of Empirical Finance, 2001, vol. 8, 345-373.

Hwang, S. – Pedersen, C. S.: “Asymmetric risk measures when modelling emerging market equities: evidence for regional and timing effects”. *Emerging Markets Review*, 2004, vol. 5, no. 1, 109-128.

Hu, W. – Maddala, G.: ”Estimation and prediction problems in dynamic heterogeneous panel data models”. Ohio State University, USA, Working Paper.

Huij, J. – Post, T.: “On persistence in the performance of emerging market equity mutual funds”. Rotterdam School of Management, The Netherlands, Working Paper, 2008, 1-33.

Huij, J. – Verbeek, M.: “Evaluating mutual fund performance and its persistence using shrinkage estimators”. Erasmus University Rotterdam, The Netherlands, Working Paper, 2003, 1-20.

Huij, J. – Verbeek, M.: “Cross-sectional learning and short-run Persistence in mutual fund performance”. *Journal of Banking & Finance*, 2007, vol. 31, 973-997.

Israelsen, G. L.: “Sharpening the Sharpe ratio”. Financial Planning Magazine, 2003.

Jan, Y. C. – Hung, M. W.: “Short-run and long-run persistence in mutual funds”. The Journal of Investing, 2004, spring, 67-71.

Jensen, M. C.: “The performance of mutual funds in the period 1945-1964”. *The Journal of Finance*, 1968, vol. 23, no. 2, 389-416.

Jensen, M. C.: “Risk, the pricing of capital assets and the evaluation of investment portfolios”. *Journal of Business,* 1969, vol. 42, no. 2, 167-247.

Jobson, J. D. - Korkie, B. M.: “Performance hypothesis testing with the Sharpe and Treynor Measures”. *The Journal of Finance*, 1981, vol. 36, no. 4, 889-908.

Karoui, A.: “Performance analysis of new mutual funds: a Bayesian approach”. HEC, Montreal, Canada, Working Paper, 2008, 1-49.

Kauppalehti, Finland: Kustannus Oy Aamulehti, Tampere. 2008, no. 122.

Klemkosky, R. C. – Maness, T. S.: “The predictability of real portfolio risk levels”. *The Journal of Finance*, 1978, vol. 33, no. 2, 631-639.

Koop, G.: Bayesian Econometrics. USA: John Wiley & Sons, Inc. 2003.

Kosowski, R. – Timmermann, A.G. – Wermers, R. – Halbert, L. W. JR.: “Can mutual fund stars really pick stocks? New evidence from a bootstrap analysis”. *The Journal of Finance*, 2007, vol. 6, no. 6, 2551-2595.

Malkiel, B. G.: “Returns from investing in equity mutual funds 1971 to 1991”. *The Journal of Finance*, 1995, vol. 4 no. 2, 549-572.

Memmel, C.: “Performance hypothesis testing with the Sharpe ratio”. *Finance Letters*, 2003, vol. 1, no. 1, 21-23.

Mincer, J. – Zarnowitz, V.: “The evaluation of economic forecasts”. Part one of the publication “Economic forecasts and expectations: Analysis of forecasting behaviour and performance”. Studies in Business Cycles, 1969, no. 19, 1-46.

Mossin, J.: “Equilibrium in a capital asset market”. *Econometrica*, 1966, vol. 34, 768-783.

Plantiga, A. – van der Meer, R. – Sortino, F.: “The impact of downside risk on risk adjusted performance of mutual funds in the Euronext markets”. University of Groningen, The Netherlands, Working Paper, 2001, 1-14.

Prigent, J. L.: Portfolio optimization and performance analysis. USA: Taylor & Francis Group, LLC, 2007.

Pätäri, E.: “*Essays on portfolio performance measurement*”. Lappeenranta University of Technology, Acta Universitatis Lappeentantaensis, nro. 106, 2000.

Pätäri, E.: a “A closer look at the performance persistence in mutual funds”. *The Journal of Performance Measurement*, 2008, vol. 12, no. 3, 38-46.

Pätäri, E.: b “Comparative analysis of total-risk-based performance measures”. *The Journal of Risk*, 2008, vol. 10, no. 4, 69-112.

Quigley, G. – Sinquefield, R. A.: “Performance of UK equity unit trusts”. *Journal of Asset Management*, 2000, vol. 1, no. 1, 72-92.

Raj, M. – Forsyth, M. – Tomini, O.: “Fund performance in a downside context”. *Journal of Investing*, 2003, vol. 12, no.2, 1-41.

Roll, R.: “A critique of the asset pricing theory tests: Part I: On past and potential testability of the theory”. *Journal of Financial Economics*, 1977, vol. 4, no. 2, 129-176.

Sandvall, T.: “Essays on mutual fund performance evaluation”. Swedish School of Economics and Business Administration, Research Reports 45, 1999, 1-88.

Sauer, D. A.: “Information content of prior period mutual fund Performance rankings”. *Journal of Economics and Business*, 1997, vol. 49, 549-567.

Sharpe, W. F.: “Capital asset prices: A theory of market equilibrium under conditions of risk”. *Journal of Finance*, 1964, vol. 19, no. 3, 425-442.

Sharpe, W. F.: “Mutual fund performance”. *The Journal of Finance*, 1966, vol. 39, no. 1, 119-138.

Sharpe, W. F.: “The Sharpe ratio”. *The Journal of Portfolio Management*, Fall 1994, 49-58.

Sharpe, W. F. – Alexander, G. J. – Bailey, J. V.: Investments (6th edition). USA: Prentice Hall, Inc. 1999.

Stevenson, S.: “Emerging markets, downside risk and the asset allocation decision”. *Emerging Markets Review*, 2001, vol. 2, no.1, 50-66.

Vaihekoski, M.: “On the calculation of the risk-free rate for tests of asset pricing models”. Lappeenranta School of Business, Lappeenranta University of Technology, Finland, Working Paper, 2007, 1-16.

Watsham, T. J – Parramore, K.: Quantitative methods in finance (1st edition). The UK: Thomson, 1997.

APPENDICES

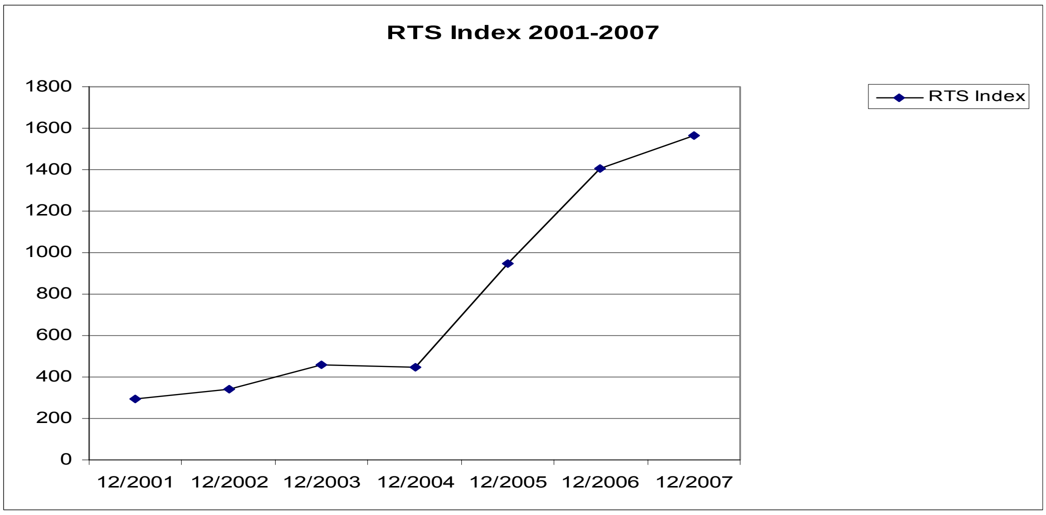
Appendix 1: Auxiliary regression results

Appendix 1 presents the results form the auxiliary regressions. For each regression are shown the dependent variable (Y), the independent variables (x) and the adjusted coefficient of determination.

Y variable	X variables	Adjusted R^2
r_m-r_f	HML and SMB	0.234
SMB	r_m-r_f and HML	0.252
HML	r_m-r_f and SMB	0.028

Appendix 2: Development of the RTS Index during 2001-2007

Appendix 2 presents the development of the RTS Index during 2001-2007. The RTS Index is denominated in euros.



Appendix 3: Coefficient of variation from the selection periods

Appendix 3 presents the coefficient of variation for each selection period for the OLS alphas and for the Bayesian alphas after 1st, 2nd, 3rd, 4th, 5th, 6th iterations.

	OLS Alpha	1st Iteration Bayes	2nd Iteration Bayes	3rd Iteration Bayes	4th Iteration Bayes	5th Iteration Bayes	6th Iteration Bayes
2002	354.4%	80.87%	18.57%	9.85%	9.46%	3.81%	1.31%
2003	26.34%	17.56%	14.46%	12.03%	9.33%	6.13%	2.71%
2004	262.9%	132.9%	60.54%	23.59%	12.04%	9.79%	9.13%
2005	69.00%	48.10%	32.42%	20.52%	9.76%	1.83%	0.21%
2006	190.20%	76.11%	39.40%	32.09%	27.65%	23.25%	18.52%

Appendix 4: Coefficient of variation from the holding periods

Appendix 3 presents the coefficient of variation for each holding period for the OLS alphas and for the Bayesian alphas after 1st, 2nd, 3rd, 4th, 5th, 6th iterations.

	OLS Alpha	1st Iteration Bayes	2nd Iteration Bayes	3rd Iteration Bayes	4th Iteration Bayes	5th Iteration Bayes	6th Iteration Bayes
2003	23.98%	15.11%	11.68%	9.21%	6.60%	3.63%	1.22%
2004	274.78%	138.53%	65.89%	36.61%	29.19%	26.11%	23.01%
2005	68.11%	46.08%	29.36%	17.18%	7.34%	1.03%	0.60%
2006	91.37%	54.72%	42.92%	37.52%	32.80%	27.97%	22.75%
2007	76.34%	42.18%	23.21%	14.66%	10.75%	8.57%	6.98

Appendix 5: Cross-rank correlations between the performance measures on the selection periods

Appendix 5 presents the cross-rank correlations between the Sharpe ratio, Fama-French alpha and the Bayesian alpha during the on the selection periods 2002 and 2003. The statistical significance of the correlation coefficient is measured with the t-statistic.

Selection period	Performance measures	Rank correlation	t-statistic
2002	Sharpe ratio vs. Fama-French	0.693	4.413***
2003	Sharpe ratio vs. Fama-French	0.362	1.982**
2002	Sharpe ratio vs. Bayesian alpha	0.811	6.358***
2003	Sharpe ratio vs. Bayesian alpha	-0.201	-0.104
2002	Fama-French vs. Bayesian alpha	0.693	4.413***
2003	Fama-French vs. Bayesian alpha	0.362	1.982**

** Statistically significant at 5% level.

*** Statistically significant at 1% level.

Appendix 6: Cross-rank correlations between the performance measures on the holding periods

Appendix 5 presents the cross-rank correlations between the Sharpe ratio, Fama-French alpha and the Bayesian alpha during the on the holding periods 2003 and 2004. The statistical significance of the correlation coefficient is measured with the t-statistic.

Holding period	Performance measures	Rank correlation	t-statistic
2003	Sharpe ratio vs. Fama-French	0.401	2.007**
2004	Sharpe ratio vs. Fama-French	0.892	10.071***
2003	Sharpe ratio vs. Bayesian alpha	-0.177	-0.828
2004	Sharpe ratio vs. Bayesian alpha	0.726	5.396***
2003	Fama-French vs. Bayesian alpha	0.606	3.497***
2004	Fama-French vs. Bayesian alpha	0.605	3.883***

** Statistically significant at 5% level.

*** Statistically significant at 1% level.