LAPPEENRANTA UNIVERSITY OF TECHNOLOGY

School of Business

Finance

MUTUAL FUND PERFORMANCE EVALUATION

Bachelor's thesis

Author: Tanja Träff

Instructor: Eero Pätäri

19.12.2012

Table of contents

1. Introduction

1.1 Backround and motivation

In April 2012, there were approximately 400 mutual funds in Finland, with 60 billion euros in assets, and at the end of August 2012 the amount had increased, according to Federation of Finnish Financial Services (2012). The majority of these assets lie under actively managed funds. Although the active management offers number of services, such as check-writing and bookkeeping services, more than half of the expenses of mutual funds arise from their stock selection ability and their ability to time the market.

The investing public's interest in identifying successful fund managers is understandable given the size of their stake, but it is also interesting in academic perspective as identifying superior fund managers will challenge the efficient market hypothesis of Fama (1970). The fund managers are as well interested on quantifying their performance and the senior management has their own benefit to look after.

The massive growth in academic financial literature regarding market timing, and fund performance in general, has various reasons. The expense of producing such literature has declined. Early studies had to rely on proprietary or expensive commercial databases or pick data by hand from published paper volumes to receive their fund performance figures. In addition to low-cost databases, like Morningstar, Lipper, Reuters and Business Week, nowadays there are also regulations that make data available to researchers. Every mutual fund in U.S. must report its own performance by pricing individual securities at end of day prices and use this to calculate daily net asset values per share. In addition, the mutual funds are committed to report their investment holdings quarterly.

Regardless of the method used the vast majority of the academic finance literature finds little or no evidence of market timing ability. This might be one of the reasons that drive academics to develop new and more complex methods to explore whether market timing exists.

1.2 Objectives

This thesis is a literature review concentrating on academic literature that has investigated the ability of the mutual fund manager to time the market. Market timing in this context is referring to the ability of a fund manager to take on more exposure to the market before it goes up and to pull out of the market before it goes down. Asset allocation is a relatively known term for market timing and is used in this thesis when referring to literature that has used that term for market timing.

1.3 Limitations

As the literature on the subject is massive, the writers own interests have affected the selection. The object is to focus on the timing ability of mutual funds, and leave out methods that are used to detect other type of performance or to evaluate other fund types. Selectivity is often investigated with market timing ability, but the main focus in this thesis remains in the market timing ability. Selectivity in this context refers to the fund manager's ability to have security-spesific information, such as pick winning stock within asset classes.

1.4 Structure of the thesis

Sections 2 to 5 present a selection of these methods developed to detect market timing ability, and reviews the important problems and properties associated with these methods. The methods presented in section 2 present the Sharpe ratio and the Jensen alpha, which are the base to later academic literature on the subject. Section 3 presents the traditional methods of market timing starting with the Capital Asset Pricing -model of Sharpe (1964). Section 4 reviews conditional and weight-based performance evaluation, which was first presented by Ferson and Schadt (1996) and section 5 concentrates on attribution analysis. The conclusions are represented in section 6 at the end of the thesis.

2. The early models

2.1 Sharpe

Sharpe (1966) developed the simplest risk-adjusted performance measure which is known as the Sharpe ratio. The Sharpe ratio measures the degree to which the portfolio is able to yield a return in excess of the risk-free return to cash, per unit of risk. In performance measurement the Sharpe ratio of the portfolio is compared to the Sharpe ratio of the benchmark. If the ratio is higher for the fund, it performs better than the benchmark.

To evaluate the performance of fund $\boldsymbol{p}$ the ratio is defined as:

$$SR_p = E(r_p)/\sigma(r_p) \tag{1}$$

where $\boldsymbol{r_p \equiv R_p - R_f}$ is the return of the portfolio return $\boldsymbol{p}$, net of return, $\boldsymbol{R_f}$, to a safe asset or cash and $\boldsymbol{\sigma(r_p)}$ is the standard deviation or volatility of excess return of the portfolio.

Aragon and Ferson (2006) point out that the Sharpe ratio is inappropriate if it is used only for parts of the portfolio instead of the whole portfolio. They also state that if the fund returns are not normally distributed, the Sharpe ratio might have misleading results. Irrespective to all the deficiencies of the Sharpe ratio, it is the basis to many ratios and models later developed to predict and analyze fund performance.

The importance of the Sharpe ratio in this study is not in the market timing area but to reveal the basic assumptions of the ratio as the ratio will be the base to other models later explained.

2.2 Jensen

Jensen (1968) introduced the classical measure of investment performance called Jensen's alpha. The alpha is based on the assumptions of CAPM which will be explained in section 4.1 more closely. The CAPM is based on the assumption that (1) all investors are risk averse, and single period expected utility of terminal wealth maximizers, (2) all investors have identical decision horizons and homogenous expectations regarding investment opportunities, (3) all investors are able to choose among portfolios solely on the basis of expected returns and variance returns, (4) all transactions costs and taxes are zero, and (5) all assets are infinitely divisible. It is also assumed that the capital market is in equilibrium. The equation for the Jensen's alpha, $\boldsymbol{\propto_p}$, is as follows:

$$\propto_p = \tilde{r}_p - \left[r_f + \beta_p\left(\tilde{r}_m - r_f\right)\right] \tag{2}$$

where $\boldsymbol{\tilde{r}_p}$ is the expected total return on the portfolio

$\boldsymbol{r_f}$ is the risk free interest rate

$\boldsymbol{\beta_p}$ is the measure of systematic risk which the asset pricing model implies is crucial in determining the prices of risky assets. $\boldsymbol{\beta_p}$ is also known as the beta of the asset.

$\boldsymbol{\tilde{r}_m}$ is the expected market return.

Jensen (1968) studied the annual data of 115 mutual funds from year 1945 to 1964. He performed a regression for each of the fund to determine its alpha. The evidence showed that on average none of the funds was able to predict security prices well enough to outperform a buy-and-hold strategy. In addition there was very little evidence that any individual fund was able to do significantly better than what would be expected merely by chance.

The Jensen alpha is a relatively famous term and variations of the traditional Jensen's alpha are popular in the performance measures in academic studies. However, the measure has disadvantages. Fama (1972) noticed that the alpha does not control for nonsystematic sources of risk that could matter to investors. The Jensen ratio requires many years of performance data which was one of the disadvantages pointed out by Aragon and Ferson (2006).

3. Traditional market timing

Traditional market timing models, such as Treynor and Mazuy (1966) (hereafter referred to as TM) and Henriksson and Merton (1981) (hereafter referred to as HM), of market-timing use convexity in the relation between the fund's return and the market return to indicate timing ability. In these models the manager has (or has not) insight about the future performance of the market and adjusts the market exposure or beta of the portfolio at the beginning of the period. Successful timing implies higher betas when the market goes up, or lower betas when it goes down, leading to the convex relation. (Ferson, 2010)

3.1 Sharpe

Sharpe (1964) was one of the developers of CAPM, capital asset pricing -model. The model is the base for the forthcoming models in Section 3. CAPM refers to the equilibrium relationship among security prices which result to when investors have homogenous beliefs and choose their portfolios based on a mean-variance criterion function. The CAPM by Sharpe is probably the best known and most widely used numerical model (Goetzmann *et al.* 2007). The CAPM is as follows:

$$E(Z_s) = z_f + \beta\left[E(Z_m) - z_f\right] \tag{3}$$

where $\boldsymbol{Z_s}$ and $\boldsymbol{Z_m}$ are simple returns of the stock and the market over some specific period. $\boldsymbol{z_f}$ is the known risk-free rate and $\boldsymbol{\beta}$ is the stock's beta. $\boldsymbol{E(Z_s)}$ denotes the expected stock return. The stock's excess expected return over the risk-free rate equals its beta times the markets expected excess return over risk-free rate.

The CAPM model using conditional returns has faced criticism from at least Harvey (1989) and Roll (1978) as the inferences about performance can be sensitive to the specification of an inefficient benchmark. Conditional returns refer to benchmark returns that are constructed from different time-varying variables.

3.2 Treynor and Mazuy

Treynor and Mazuy (1966) use the CAPM as the base of their model, to detect if there is a convex relation between market beta and the return on the market. TM included a quadratic term to original CAPM, to test for market timing ability. TM regression:

$$r_{p,t} = \alpha_p + \beta_p r_{m,t} + \gamma_p r_{m,t}^2 + \varepsilon_{p,t} \,, \tag{4}$$

where $\boldsymbol{r_{p,t}}$ is the excess return on a portfolio at time $\boldsymbol{t}$, $\boldsymbol{r_{m,t}}$ is the excess return on the market, and $\boldsymbol{\gamma_p}$ measures timing ability. If a mutual fund manager increases (decreases) the portfolio's market exposure prior to a market increase (decrease) then the portfolio's return will be a convex function of the market's return, and $\boldsymbol{\gamma_p}$ will be positive. In simple terms, the mutual fund manager has the ability to change the market exposure of the portfolio in anticipation of moves in stock market. TM found, that only one of the 57 funds in their sample possessed significant timing ability on a 95% confidence level.

The convex variation may arise because of common time-variation in the fund's beta risk and the expected market risk premium, related to public information on the state of economy. This conditional method will be presented in Section 4. Ferson and Schadt (1996) develop a conditional method from the TM model that avoids the bias of public information. The method will be more closely reviewed in Section 4.4.

According to Ferson (2010), the Treynor and Mazuy (1966) model suffers from invalid intercept. The intercept in the model does not capture the return in excess of a benchmark portfolio because $\boldsymbol{r_m^2}$ in Equation 4 is not a portfolio return.

3.3 Henriksson and Merton

Henriksson and Merton (1981) developed a regression model very similar to TM model. In their model the mutual fund manager shifts the portfolio weights based on forecasts, as in the TM model, with the exception that the manager decides only between a small number of market exposure levels. HM regression:

$$r_{p,t} = \alpha_p + \beta_p r_{m,t} + \gamma_p r^*_{m,t} + \varepsilon_{p,t} \ , \tag{5}$$

where,

$$r^*_{m,t} = I\{r_{m,t} > 0\} r_{m,t} \tag{6}$$

where $\boldsymbol{r_{p,t}}$ is the excess return on a portfolio at time ***t***, $\boldsymbol{r_{m,t}}$ is the excess return on the market, and $\boldsymbol{\gamma_p}$ measures timing ability. $\boldsymbol{I\{r_{m,t} > 0\}}$ is an indicator function that equals one if $\boldsymbol{r_{m,t}}$ is positive and zero otherwise. The magnitude of $\boldsymbol{\gamma_p}$ in the Equation 5 measures the difference between the target betas, interpreted in the literature as "timing adjusted" selectivity, and is positive for a mutual fund manager that successfully times the market. The "timing adjusted" selectivity in simple terms means that a successful market timer can be seen as producing free put options. The benchmark portfolio for evaluating a market timer is therefore a combination of the market portfolio, safe assets, and an option on the market portfolio. (Aragon and Ferson, 2006; Henriksson and Merton, 1981)

The HM –model has been interpreted as a measure of timing-adjusted selectivity performance. The model only shows perfect market timing, as it assumes that the manager has the ability to obtain the option-like payoff at zero cost. As perfect timing ability does not exist, the interpretation of $\boldsymbol{\alpha_p}$ as timing adjusted selectivity breaks down. The model is also unable to separate the effects of market timing and selectivity, but is able to capture their mutual performance. The separation between timing and selection is a disadvantage among all returns-based models without making strong assumptions about at least one of the components. (Aragon and Ferson, 2006)

Henriksson (1984) estimated the HM model to 116 mutual funds and found that the average value for $\boldsymbol{\gamma_p}$ was negative. Overall he found little evidence of timing ability, only 3 of the 116 funds had significantly positive estimate of timing ability.

4. Conditional and weight-based market timing

Elton et al. (2011) states that the potential problem in both TM and HM models is, that they assume the fund manager to make timing decisions in a specific way. If and when the mutual fund manager decides to time in a more complex manner, the TM and HM models may not detect it. To avoid this bias later studies have used portfolio holdings and security betas in estimating the portfolio betas.

The strength in the traditional, return-based, models is the minimal information requirements. The requirements include only the returns of managed portfolio and the benchmark portfolio. The minimum requirements can also be seen as a disadvantage as there is plenty of useful information available. (Aragon and Ferson, 2006)

Traditional methods in Section 3 compare the average return of a fund with a benchmark portfolio designed to control for the fund's average risk. The portfolio has the same average market exposure, or "beta" risk as the fund. The returns and beta risks are typically measured as averages over the evaluation period, and these averages are taken "unconditionally," or without regard to variations in the state of financial markets or the broader economy. One weakness of this unconditional approach relates to the likelihood of changes in the state of the economy. In the Conditional Performance Evaluation (CPE) approach, fund managers' risk exposures and the related market premiums are allowed to vary over time with the state of the economy. The state of the economy is measured using predetermined, public information variables. Provided that the estimation period covers both bull and bear markets, it is possible to estimate expected risk and performance in each type of market. (Aragon and Ferson, 2006)

Traditional, unconditional, weight-based measures have also problems handling return dynamics. It is known that unconditional weight-based measures can show performance when the manager targets stocks whose expected return and risk have risen temporarily; when a manager exploits serial correlation in stock returns or return seasonalities; and when a manager gradually changes the risk of the portfolio over time. These problems might be avoided using a conditional approach. (Aragon and Ferson, 2006)

4.1 Interim trading bias

A common bias in performance evaluation is interim trading bias. The bias is related to returns data frequency that is used in the study. If returns are measured in over two periods, but the mutual fund manager trades each period, the manager might have a neutral performance but the portfolio weights for the second period can react to public information at the middle date. The higher volatility may indicate that the expected return-to-risk tradeoff for stocks has become less favorable for the second period, so the optimal portfolio is now more conservative. If only two-period returns can be measured and evaluated, the manager's strategy would appear to have partially anticipated the higher volatility. The fund's two-period market exposure would reflect some average of the before- and after-event positions. Measured from the beginning of the first period, the portfolio would appear to partially "time" the volatility-increasing event because of the move into cash. A returns-based measure over the two periods will detect this as a superior performance. (Aragon and Ferson, 2006)

Aragon and Ferson (2006) state that a weight- based measure can avoid this bias by examining the conditional covariance between the manager's weights at the beginning of the first period and the subsequent two-period returns. The ability of the fund manager to trade at the intervening period does not enter into the measure, and thus interim trading creates no bias. It is possible that managers engage in interim trading based on superior information to enhance performance, and a weight-based measure will not record these interim effects. Ferson and Khang (2002) refine a holdings-based method that evaluates these tradeoffs. The method will be presented in Section 4.5.

4.2 Grinblatt and Titman

Grinblatt and Titman (1993) (hereafter referred as GT) developed the first conditional measure for mutual funds that demean weights. Their measure is based on the assumption that, from the perspective of an uninformed investors, the vector of expected asset return is constant over time. Thus, the holdings of an uninformed investor cannot be correlated with future asset returns.

The GT derive their measure in a single-period model where the fund manager maximizes the expected utility of the terminal wealth generated by the portfolio return, conditional on the information $\boldsymbol{\Omega}$. When returns are conditionally normal given $\boldsymbol{\Omega}$, and assuming non-increasing absolute risk aversion, they show that:

$$Cov\,\{x(\Omega)'r\,\} > 0 \tag{7}$$

where $\boldsymbol{x}(\boldsymbol{\Omega})$ is the optimal weight vector. Equation 7 implies that the sum of the covariances between the weights of a manager with private information, $\boldsymbol{\Omega}$, and the returns for the securities in a portfolio is positive.

From the definition of covariance can be implemented Equation 8 by demeaning the weights or the returns:

$$Cov\,\{x(\Omega)'r\,\} = \; E\,\{[x(\Omega) - E(x(\Omega))]r\} = E\{x(\Omega)[r - E(r)]\} \tag{8}$$

The earlier studies of Copeland and Myers (1986) and the studies of Ferson and Khang (2002) (hereafter referred as FK) demean returns while GT demeans the benchmark weights. The benchmark weights introduced by GT:

$$Cov\,\{x(\Omega)'r\,\} = E\{[x(\Omega) - x_B]'r\}. \tag{9}$$

With the benchmark weights, the benchmark portfolio implied by the weight based measure is given by $\boldsymbol{r_B} = \boldsymbol{x_B}'\boldsymbol{r}$. In their study in GT define the benchmark as the fund's weights in the previous quarter. Thus, the model assumes that managers with no information holds fixed portfolio weights.

Their study showed positive abnormal investment performance. The study revealed persistency to performance, as investors with superior performance in the first half exhibited this ability persistently. The same scenario was noticed with investors who did not do well.

GT state that their study is superior to previous traditional models but it has some areas of concern. First, the new technique developed is more costly to implement. This concerns both computer time and the data collected. Second, as the quarterly holdings do not present the actual fund portfolio, the portfolio actually being evaluated is some hypothetical portfolio. If as an insider it would be possible to receive daily holdings information, this concern would be a nonissue. To outside evaluators quarterly holdings add noise to true performance. Regardless of this noise concern GT find signs of positive abnormal investment performance and therefore this is only a slight concern. The third concern GT issue is that fund managers might game the measure by selecting securities when they are riskier than usual. This can be addressed by combining traditional evaluation techniques with the evaluation processed by GT.

Daniel et al. (1997) criticize the GT model benchmarks for not fully accounting for return anomalies, such as the size, book-to-market, and momentum effects. They also point out that their characteristic timing component is more powerful than factor-based models. While using factor methods one must determine whether changes in factor loadings correspond with the realizations of the associated factors. The characteristic measure presented enables to easily look at whether shifts in the portfolio weights forecast future returns.

4.3 Daniel, Grinblatt, Titman and Wermers

Daniel *et al*. (1997) (hereafter referred as DGTW) define the benchmark portfolio based upon the securities held. This way they avoid the return anomalies that the previous studies are exposed to. The previous studies used factor models to detect abnormal performance, but DGTW develop a characteristic-matching method to construct the benchmark portfolio.

Each security in the fund’s portfolio is assigned to one of 125 characteristic groups, depending upon its size, book-to-market ratio, and momentum (lagged return), measurable with respect to the beginning of the quarter. They construct passive value-weighted portfolios across all NYSE, AMEX, and NASDAQ stocks for each characteristic group. The return on the benchmark portfolio in a given quarter is the summation of the fund's portfolio weights times the return on the characteristics-matched portfolio for that stock. The model is divided to three different components which are: average style, characteristic selectivity and characteristic timing. Because of the limitations of this study, only the latter will be construed. (Daniel et. al, 1997)

$$\mathrm{DGTV}_t = \sum_{j=1}^{N} \left(\varpi_{j,t-k} \widehat{R}_t^{b,t-k} - \varpi_{j,t} \widehat{R}_t^{b,t} \right) \tag{10}$$

where $\boldsymbol{\varpi_{j,t-1}}$, is the portfolio weight of stock $\boldsymbol{j}$ at month $\boldsymbol{t-k}$ is multiplied by $\boldsymbol{\widehat{R}_t^{b,t-k}}$, the month $\boldsymbol{t}$ return of the characteristic-based benchmark portfolio that is matched to stock $\boldsymbol{j}$ during month $\boldsymbol{t}$. Thus, if the fund increases its weight in high book-to-market stocks at the beginning of a month in which the book-to-market effect was unusually strong, then the fund would have a positive characteristic timing component for that month.

They suggest that in some cases the style of the fund can be captured by using the returns of other managed portfolios but this brings up a question whether such a comparable is possible to find. The study was based on quarterly data from 1972 to 1994. DGTW found that the component to measure timing ability was insignificant in all categories of funds examined and was never significantly positive in any sample period. Their study states that an average mutual fund is not able to effectively time the different stock characteristics.

4.4 Ferson and Schadt

Ferson and Schadt (1996) (hereafter referred as FS) developed their conditional beta to detect market timing ability. They make the assumption, that a managed portfolio strategy which can be replicated using readily available public information should not be judged as having superior performance. FS argue that unconditional measures are biased when mutual fund managers respond to information of the last period. They extend the previous studies by using time-varying factors to reduce this bias. FS include this time-varying conditional information to both Jensen's alpha, TM and MH models, for comparison.

In HM model they remove the quadratic term. Their conditional model:

$$r_{p,t+1} = \alpha_p + \beta_p r_{m,t+1} + C_p'\left(Z_t r_{m,t+1}\right) + \Lambda_p {r_{m,t+1}}^2 + \omega_{t+1}\,, \qquad (11)$$

where the term $\boldsymbol{C_p'}\left(\boldsymbol{Z_t r_{m,t+1}}\right)$ controls for predictable time-variation in the market risk premium and the fund's beta. A mutual fund manager who only uses $\boldsymbol{Z_t}$ has no conditional timing ability, and thus $\boldsymbol{\Lambda_p = 0}$. The coefficient $\boldsymbol{\Lambda_p}$ measures the market timing ability based on information beyond contained $\boldsymbol{Z_t}$. The conditional beta, $\boldsymbol{\beta_p}$, in their model consists from the lagged level of short term Treasury bill rate, the January dummy, the lagged dividend yield of the stock index, and the lagged measure of the slope of term structure to predict the market.

Their study shows that the preserve market timing performance presented in both TM and HM studies was removed with the use of time-varying conditional beta. FS refer to their model as an unconditional model which attempts to distinguish market timing based on publicly available information from market timing information that is superior to lagged information variables.

FS point out possible disadvantages in the model. The first one is the negative correlation between mutual fund conditional betas and expected market returns, which might be due to the flows of net new money into mutual funds. Secondly they point out that the conditional versions of these traditional market timing models do not solve the problem of interim trading bias presented in Section 4.1.

4.5 Ferson and Khang

Ferson and Khang (2002) use the model of Grinblatt and Titman (1993) with a different approach to the portfolio weights. Where Grinblatt and Titman (1993) use the fund's previous quarters weights as the benchmark portfolio weights, Ferson and Khang (2002) define the benchmark weights as the portfolios' actual weights lagged k periods, the same approach is used also with the economic variables or instruments (See Equations 7, 8 and 9). Each manager's position, $\boldsymbol{k}$ quarters ago defines his personal benchmark. A manager with investment ability changes the portfolio in order to beat a buy-and-hold strategy. Fund performance is measured as the average difference in raw returns over the subsequent quarter, between the fund and the benchmark portfolio defined by the weights, $\boldsymbol{x_B}$. The return of the fund is a "hypothetical" return, since it is constructed using a snapshot of the fund's actual weights at the end of a period. This hypothetical return reflects no trading within the quarter, no trading costs and no management fees.

Ferson and Khang used quarterly holdings and returns of 60 U.S. equity portfolios during time period from 1984 to 1994. Their sample implied positive and abnormal significant performance. They created a portfolio weight that can avoid interim trading bias, but is vulnerable to other biases. The portfolio weight was used with several previous measures (replacing the old weight measure), both unconditional and conditional, and the results vary depending on the measure. Therefore the results are not straightforward.

4.6 Jiang

Jiang (2003) developed a new method to reveal if market timing exists. The model leaves out the estimation of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and is therefore unique in comparison to previous studies. The simple idea behind the model is that the fund of a successful fund manager rises significantly when the market rises and falls slightly when the market falls. The model's advantages rise from the minimal information requirements (fund returns and benchmark portfolio). In addition, the test statistics are not affected by the manager's risk aversion as it separates the quality of timing information a fund manager possesses from the aggressiveness of the reaction of such information. The nonparametric model is also more robust to different information and incentive structures than previous models presented.

The model assumes that a fund manager's timing information is independent of her information about individual securities. With this assumption the model is defined as:

$$r_{i,t+1} = \alpha_i + \beta_{i,t} r_{m,t+1} + \varepsilon_{i,t+1}\,, \tag{12}$$

where $\boldsymbol{i} =$ the subscript for individual fund

$\boldsymbol{\beta_{i,t}}$ = a random variable adapted to the information available to the manager at time $\boldsymbol{t}$

$\boldsymbol{r_m}$ = the return of the relevant market in which the mutual fund invests

For triplet $\{\boldsymbol{r_{m,t_1}}, \boldsymbol{r_{m,t_2}}, \boldsymbol{r_{m,t_3}}\}$ sampled from any three periods where $\boldsymbol{r_{m,t_1}} < \boldsymbol{r_{m,t_2}} < \boldsymbol{r_{m,t_3}}$, an informed fund manager should, on average, maintain a higher average $\boldsymbol{\beta}$ in the $|\boldsymbol{r_{m,t_2}}, \boldsymbol{r_{m,t_3}}|$ range than in the $|\boldsymbol{r_{m,t_1}}, \boldsymbol{r_{m,t_2}}|$ range. The $\boldsymbol{\beta}$ estimates for both ranges (given two observations for each range) are $(\boldsymbol{r_{i,t_2}}, \boldsymbol{r_{i,t_1}})/(\boldsymbol{r_{m,t_2}}, \boldsymbol{r_{m,t_1}})$ and $(\boldsymbol{r_{i,t_3}}, \boldsymbol{r_{i,t_2}})/(\boldsymbol{r_{m,t_3}}, \boldsymbol{r_{m,t_2}})$, respectively. Therefore, he proposes the probability $\boldsymbol{\theta}$ in Equation 13 as a statistic of market timing ability.

$$\theta = 2Pr\left(\frac{\boldsymbol{r_{i,t_3}}, \boldsymbol{r_{i,t_2}}}{\boldsymbol{r_{m,t_3}}, \boldsymbol{r_{m,t_2}}} > \frac{\boldsymbol{r_{i,t_2}}, \boldsymbol{r_{i,t_1}}}{\boldsymbol{r_{m,t_2}}, \boldsymbol{r_{m,t_1}}}\right) \tag{13}$$

The fund manager's timing ability is determined by the relevance and accuracy of her information. If the timing is perfect the measure $\boldsymbol{\theta}$ would be >1 and in the case of negative market timing it would be <1.

The model of Jiang (2003) states, that with few assumptions, it is more robust to biases related to TM and HM models presented in Section 3. First, it is assumed that the fund manager's information on individual securities is independent on her information on the market. Secondly, it is assumed that the fund does not include derivatives. Regardless of these assumptions the model still faces some probable biases. Interim trading bias is named as one as there is no conditional portfolio weights used like in the model of Ferson and Khang

(2001). The return data frequency can bias the results if timing is taking place in between evaluations.

The model was tested with monthly data gathered from 1980 to 1999. There were 1827 funds that survived the whole period and 110 funds that were eliminated at some point of the evaluation period. The benchmarks for the funds were picked from the best fitting index. The suitability was evaluated by regression. Overall there were found no evidence of superior market timing ability.

4.7 Jiang, Yao and Yu

Jiang *et al.* (2007) present their holdings-based measure to predict market timing ability. The model estimates the fund's beta as the weighted average of the beta's of individual stocks held in the portfolio. It tests whether the covariance between the fund betas at beginning of the holding period and the holding period market return is significant. Fund holdings are often available for lower frequencies than fund returns but they compose their beta in a new way that does not limit the model's statistical significance. In addition to better statistical power the holdings-based model is also able to handle return dynamics. To prove the superiority of the model they apply the data to return-based models of TM and HM in addition to their new holding-based measure.

Jiang *et al.* (2007) holdings-based beta is used in the traditional models of TM and HM (presented in Section 3.2 and 3.3) and the results are compared to traditional TM and HM results. The $\widehat{\boldsymbol{\beta}}$ is constructed as follows:

$$\hat{\beta}_t = \sum_{i=1}^{N} \omega_{it}\, \hat{b}_{it}, \tag{14}$$

where $\boldsymbol{\omega_{it}}$ = the weight for stock $\boldsymbol{i}$ at the beginning of holding period $\boldsymbol{t+1}$

$\widehat{\boldsymbol{b}}_{\boldsymbol{it}}$ = the beta for stock $\boldsymbol{i}$ estimated using data prior to period $\boldsymbol{t+1}$

The coefficient $\boldsymbol{\gamma}$ from the TM and HM models, first presented in Equations 4 and 5, is estimated through regressions, respectively, as follows:

$$\hat{\beta}_t = \alpha + \gamma r_{m,t+1} + \eta_{t+1}, \tag{15}$$

$$\hat{\beta}_t = \alpha + \gamma I_{r_{m,t+1>0}} + \eta_{t+1}, \tag{16}$$

where $\widehat{\boldsymbol{\beta}}_{\boldsymbol{t}}$ = the fund beta estimated at the beginning of period $\boldsymbol{t+1}$

$\boldsymbol{\gamma}$ = market timing coefficient

Jiang *et al.* (2007) apply their measures to data set of 2294 actively managed funds over the period from 1980 to 2002. They find that, on average, the funds in the data possessed positive timing ability. They even found that some funds in the sample possessed strong market timing skills.

4.8 Cremers and Petäjistö

Cremers and Petäjistö (2009) introduce a new method called the Active Share. The Active Share describes the share of portfolio holdings that differ from the portfolio's benchmark index. To quantify the active management of the fund, the Active Share is defined as:

$$Active\ Share = \frac{1}{2}\sum_{i=1}^{N}\left|w_{fund,i} - w_{index,i}\right|, \tag{17}$$

where $\boldsymbol{w_{fund,i}}$ and $\boldsymbol{w_{index,i}}$ are the portfolio weights of asset $\boldsymbol{i}$ in the fund and in the index, and the sum is taken over the universe of all assets.

The interpretation behind the measure is that it decomposes the mutual fund portfolio in to a 100% position in the benchmark index, plus a zero-net-investment long-short portfolio. The long-short portfolio represents the actual investment decisions made by the portfolio manager. The aim of the Active Share is to measure the size of that long-short position as a fraction of the total portfolio of the fund. The portfolio weight differences are divided by 2 so that a fund with no investments in the benchmark index gets a 100% Active Share. The Active Share of a fund that never shorts stock or never buys a margin will be between 0-100%. Within hedge funds the Active Share can exceed 100% because of the leverage and net short positions in individual stocks.

Together with the Active Share, Cremers and Petäjistö (2009) use tracking error volatility to quantify active portfolio management. Tracking error volatility is the time-series standard deviation of the difference between a portfolio return ($\boldsymbol{R_{fund,t}}$) and its benchmark return index ($\boldsymbol{R_{index,t}}$) Equation 13.

$$Tracking\ error\ volatility = Stdev\left[R_{fund,t} - R_{index,t}\right] \tag{18}$$

The approaches, timing and stock selection, to add value to the portfolio contribute differently to tracking error volatility. Timing will bear a systematic risk relative to the index while stock selection may bear only idiosyncratic risk. The former will generate relatively high tracking error volatility when the latter can diversify away the idiosyncratic risk and therefore achieve relatively low tracking error volatility. A high Active Share can identify a diversified stock selector even when his tracking error is low. A diversified stock picker can be very active despite its low tracking error, because its stock selection within industries can still lead to large deviations from the index portfolio. On the other hand, a fund taking systematic factor bets can generate a large tracking error even without large deviations from index holdings. A concentrated stock picker is the combination of the two approaches taking positions in individual stocks as well as systematic risk. The last type is the closet indexer who scores low on both dimensions and yet claims to be active. (See Table 1) (Cremers and Petäjistö, 2009)

Table 1. The main types of active management by Cremers and Petäjistö. (Source Cremers and Petäjistö, 2009)

Tracking error volatility		Active Share: Low	Active Share: High
	Low	Closet indexers	Diversified stock pickers
	High	Pure factor betters	Concentrated stock pickers

Their data consisted of 25 all-equity fund returns and holdings on monthly basis from 1990 to 2003. Their results show that funds with the highest Active Share significantly outperform their benchmarks both before and after expenses. In addition they found that tracking error volatility does not predict higher returns.

4.9 Kacperczyk, Nieuwerburgh and Veldkamp

Kacperczyk *et al.* (2011) studied a different perspective to predict timing and stock picking ability of the portfolio manager. Unlike previous researches they do not expect the portfolio manager to have both value adding features at the same time. Their study includes the effect of different economical states, booms and recessions and studies the different value adding features in these different states. The model they developed gives more weight to timing ability in recessions and on the contrary to stock picking ability in booms.

The two measures of skill are as follows:

$$Timing_t^j = \frac{1}{TN^j}\sum_{i=1}^{N^j}\sum_{r=0}^{T-1}\left(w_{it+r}^j - w_{it+r}^m\right)(\beta_{it+r+1}R_{t+r+1}^m) \quad (19)$$

where $\boldsymbol{Timing_t^j}$, for fund $\boldsymbol{j}$ at time $\boldsymbol{t}$, measures how the holdings of each asset, relative to market, co-vary with the systematic component of the stock return over the next $\boldsymbol{T}$ periods. $\boldsymbol{\beta_i}$ measures the covariance of asset $\boldsymbol{i}$'s return, $\boldsymbol{R^i}$, with the market return, $\boldsymbol{R^m}$, divided by the variance of the market return. The product $\boldsymbol{\beta_i}$ and $\boldsymbol{R^m}$ measure the systematic risk component of returns of asset $\boldsymbol{i}$. The time subscripts indicate that the systematic component of the return is unknown at the time of portfolio formation. Before the market return rises, a fund with a high timing ability over weights assets that have high betas and underweights if there is anticipation of a market decline.

$$Picking_t^j = \frac{1}{N^j}\sum_{i=1}^{N^j}\left(w_{it}^j - w_{it}^m\right)\left(R_{t-1}^i - \beta_i R_{t-1}^m\right) \tag{20}$$

The $\boldsymbol{Picking_t^j}$ measures similarly how the fund's holdings of each stock, relative to market, co-vary with the idiosyncratic component of the stock return. A fund with a high picking ability overweights assets that have subsequently high idiosyncratic returns and underweights assets with low subsequent idiosyncratic returns.

Kacperczyk *et al.* (2011) studied the performance of 3477 funds and found that some portfolio managers exhibit timing ability in recessions and stock picking ability in recessions. Their study shows that these skills are based on anticipating changes in firm-specific or market fundamentals. They further suggest that this skill is conducted by research and analysis in individual firms in booms, and by the state of the aggregate economy in recessions, and to trade on that information.

5. Attribution analysis

Performance attribution analysis is a tool to understand the sources of return in a portfolio. The results of the analysis are useful to the portfolio manager as well as the senior management and clients. The aim of the analysis is to quantify the decision process of the portfolio manager. To effectively use the attribution analysis as a tool it is necessary to have a good qualitative and quantitative understanding of the portfolio. (Bacon, 2008)

The foundations of performance attribution analysis are in a model called the Brinson model. The Brinson model is the sum of articles by Brinson *et al.* (1986) and Brinson and Fachler (1985). Both articles are based on the assumption that the portfolio returns and benchmark returns can be disaggregated as follows:

$$Portfolio\,return \;\; r = \sum_{i=1}^{i=n} w_i \times r_i \tag{21}$$

where $\boldsymbol{w_i}$ = weight of the portfolio in the ***i***th asset class (note $\sum_{i=1}^{i=n} w_i = 1$)

$\boldsymbol{r_i}$ = return of the portfolio assets in the ***i***th asset class:

$$Benchmark\,return \;\; b = \sum_{i=1}^{i=n} W_i \times b_i \tag{22}$$

where $\boldsymbol{W_i}$ = weight of the benchmark in the ***i***th asset class (note also $\sum_{i=1}^{i=n} w_i = 1$)

$\boldsymbol{b_i}$ = return of the benchmark in the ***i***th asset class.

The aim of single period attribution analysis is to quantify portfolio manager's active decisions that contribute to the difference between portfolio return ***r*** and the benchmark return ***b***.

5.1 Brinson, Hood and Beebover

Brinson *et al.* (1986) suggested that the model is divided to timing, selection and other. In order for the mutual fund manager to add value, the manager will seek different weights in the portfolio in comparison to the benchmark. The aim of the manager is to overweight well performing assets and underweight poor performing assets – timing. Selection is the portfolio managers aim to add value by selecting individual securities within the asset classes. Table 2 shows the framework how to analyze portfolio returns.

Table 2. Brinson *et al.* framework to analyze portfolio returns. (Source: Brinson *et al.*, 1986)

		Selection	
		Actual	Passive
Timing	Actual	(IV) Actual Portfolio Return	(II) Policy and Timing Return
	Passive	(III) Policy and Security Selection Return	(I) Policy Return (Passive Portfolio Benchmark)

Active Returns Due to:

Timing	II - I
Selection	III - I
Other	IV - III - II + I
Total	IV - I

Quadrant I represents the funds benchmark return for the period. The Policy Return is the long-term investment policy of the portfolio. To calculate the policy benchmark return the weights of all asset classes and the benchmark return assigned to each asset class are needed. Quadrant II illustrates the return effects of Policy and Timing. Timing is undertaken to achieve incremental returns relative to the policy return. Quadrant III represents the excess returns from within the asset class compared to the benchmark return. The actual weights and returns of the portfolio are presented in Quadrant IV. Table 3 helps to understand the computational requirements for each asset class.

Table 3. Computational requirements for the Brinson *et al.* framework. (Source: Brinson *et al.*, 1986)

Selection $W_i \times (r_i - b_i)$	Interaction $(w_i - W_i) \times (r_i - b_i)$
Benchmark contribution $(W_i \times b_i)$	Allocation $(w_i - W_i) \times b_i$

where $\boldsymbol{W_i}$ = weight of the benchmark in the ***i***th asset class (note also $\sum_{i=1}^{i=n} w_i = 1$)

$\boldsymbol{b_i}$ = return of the benchmark in the ***i***th asset class.

$\boldsymbol{w_i}$ = weight of the portfolio in the ***i***th asset class (note $\sum_{i=1}^{i=n} w_i = 1$)

$\boldsymbol{r_i}$ = return of the portfolio assets in the ***i***th asset class:

The aim of the model is to differentiate the effects of investment policy and investment strategy. Investment strategy is shown to be composed of timing, security selection, and the effects of a cross-product term. The algebraic measures enable to calculate the exact effects of policy and strategy.

Brinson *et al.* (1986) tested their framework with 91 pension plans from year 1974 to 1983. The outcome on their study reveals that active management is clearly important but the largest portion of the excess return is composed of investment policy. Investment policy includes the portfolio manager's control over the portfolio. The value of timing and selectivity is small.

5.2 Brinson and Fachler

In the Brinson *et al.* (1986) model the overweight positions in positive markets generated positive attribution factors irrespective of the overall benchmark return while all overweight positions in negative markets generated negative attribution factors. If the overweight in the negative market has outperformed the overall benchmark, there should be a positive effect. (Bacon, 2008)

Brinson and Fachler (1985) solve this problem by modifying the asset allocation factor to compare returns against the overall benchmark as follows:

$$A_i = (w_i - W_i) \times (b_i - b) \tag{23}$$

where $\boldsymbol{A_i}$ = the new asset allocation factor

$\boldsymbol{w_i}$ = weight of the portfolio in the $\boldsymbol{i}$th asset class (note $\sum_{i=1}^{i=n} w_i = 1$)

$\boldsymbol{W_i}$ = weight of the benchmark in the $\boldsymbol{i}$th asset class (note also $\sum_{i=1}^{i=n} w_i = 1$)

$\boldsymbol{b_i}$ = return of the benchmark in the $\boldsymbol{i}$th asset class.

$\boldsymbol{b}$ = return of the portfolio

Brinson and Fachler (1985) made no other changes to the previous model. Bacon (2008) compared both Brinson models and the impact of the change made by Brinson and Fachler (1985) was major in the results.

Bacon (2008) criticizes both models for the inclusion of the interaction term as it is something that is very difficult to identify. Even though the interaction term is something that has an impact on the investment decision process the aim is not to add value through interaction.

Bacon (2002) and several others have created geometric excess return attribution models. The basics of the models are in the Brinson models but the models are extended to break down the geometric excess return. These models are left out of closer look in this study.

5.3 Menchero

Menchero (2004) stated that there is a need for a multi-period approach as the Brinson models are single-period models. The need for a multi-period approach rose from the need of a longer time period evaluation of fund performance. The objective is to take the components of the attribution analysis and link them over time to explain the sources of active return for the longer period of time. The challenge in remodeling the classical Brinson models were in how to link the periods together without losing the fundamental meaning and interpretation of the attribution effects.

Bacon (2008) showed that the arithmetic excess return for each finite period, $\left(\boldsymbol{R_t} - \boldsymbol{\widetilde{R}_t}\right)$, does not sum to the total arithmetic excess return for the total period, $\boldsymbol{R} - \boldsymbol{\widetilde{R}}$ (Equation 19).

$$R - \tilde{R} \neq \left(R_t - \tilde{R}_t\right) \tag{24}$$

Therefore, Menchero (2004) includes a factor $\boldsymbol{M}$ to the summation (Equation 25).

$$r - b \approx M \times \sum_{t=1}^{T} (r_t - b_t) \tag{25}$$

The product $\boldsymbol{M}$ takes into account the characteristic scaling which arises from the geometric compounding. It is the difference of the arithmetic average between portfolio and benchmark returns ($\boldsymbol{r}$ and $\boldsymbol{b}$) with the difference of the geometric average portfolio and benchmark returns ($\boldsymbol{r_t}$ and $\boldsymbol{b_t}$) and is presented as follows:

$$M = \frac{(r-b)/T}{\left[(1+r)^{1/t} - (1+b)^{1/T}\right]} \tag{26}$$

$$\text{if } r = b \qquad \text{set } M = (1+r)^{\frac{(T-1)}{T}} \tag{27}$$

To achieve $\boldsymbol{M}$ the corrective term, $\boldsymbol{\alpha_t}$, needs to be calculated (Equation 28).

$$\alpha_t = \left(\frac{r - b - M \times \sum_{t=1}^{t=T}(r_t - b_t)}{\sum_{t=1}^{t=T}(r_t - b_t)^2}\right) \tag{28}$$

Menchero (2004) does not test the timing or selection abilities of the funds on his data but the investigates different approaches how to link the single-period method to a multi-period method. Even though the Menchero (2004) receives complements from Bacon (2008) compared to Carino (1999), it also faces criticism. Hsu *et al.* (2010) criticize the model for not being able to explicitly measure the manager's ability to allocate dynamically in the factor domain.

5.4 Hsu, Kalesnik and Myers

Hsu *et al.* (2010) criticize the Brinson models and Menchero (2004) for not being able to measure the portfolio manager's ability to allocate dynamically in the factor domain. They divide the allocation ability to static factor allocation and dynamic factor allocation. The static factor allocation refers to when the portfolio manager wishes to add value by increasing the portfolio weights in value stocks. Dynamic factor allocation refers to when the portfolio wishes to add value by increasing the weights in value stocks compared to growth stocks when she thinks value will perform better than growth. The importance of distinguishing the difference between these two is important for various reasons. From the view of the customer it is valuable for the justification of the portfolio manager's fees to be able to point out if the performance is due to dynamic decisions rather than replicating some style indices.

The model they represent to distinguish to point out this difference in timing is as follows:

$$E(w_t R_t) = E(w_t)E(R_t) + cov(w_t, R_t) \tag{29}$$

where the hypothetical return of the portfolio manager, $\boldsymbol{E(w_t R_t)}$, is decomposed in to two terms. The first term, $\boldsymbol{E(w_t)E(R_t)}$, is the static allocation effect. Any weight on an asset that has a positive expected return will generate a positive return. The second term, $\boldsymbol{cov(w_t, R_t)}$, represents the dynamic allocation effect. The term captures the portion of the portfolio manager's ability to time the equity market. If the portfolio manager's weight in stocks is large when the market return is high and small when the market returns are low $\boldsymbol{cov(w_t, R_t) > 0}$. A portfolio manager with no meaningful timing abilility would exploit $\boldsymbol{cov(w_t, R_t) = 0}$. Thus, $\boldsymbol{cov(w_t, R_t) < 0}$ would imply that the portfolio manager is actively destroying value.

Hsu *et al.* (2010) studied the performance of 10 mutual funds with data from year 2008. The mutual funds were hand-picked to present the elite of mutual funds in 2008. The result of their study reveals that the 10 mutual funds exhibit significant skill in security selection and in timing industry sectors in their fund management. Modest skill in timing growth- and value-style stocks is shown and the static allocation skill is negative. The aim of their paper was to represent the new model to distinguish the two different types of timing ability rather than make comprehensive analysis on the subject.

6. Conclusions

This thesis was a relatively narrow review of the performance measures developed so far. All the models presented have their basics in the models developed in between 1960's and 1980's. The massive growth in the academic literature on the subject is obvious as different approaches have been developed since 1980's.

I started with presenting the early models to give the reader the basic knowledge where the performance measurement is based on. The traditional market timing methods in Section 3 were not able to find market timing ability in the sense that it would have existed. The problem of these traditional methods was the exclusion of the variations in the state of the financial markets or the broader economy. Other problems among these traditional methods were the interim trading bias and the ability to handle return dynamics. Both problems were solved in models presented in Section 4.

The conditional and weight-based market timing methods presented were able to show market timing ability among funds investigated. These models use the available data of the fund holdings to have more insight to the fund performance. The models were free of the bias regarding the economic variable and interim trading but have faced criticism of some level on other matters. The evolvement of the methods is shown in more complex approaches to construct the benchmark portfolio and different variations of the way to calculate the weights of the benchmark portfolio.

Section 5 presented the attribution analysis which combines the use of both holdings and returns. The models presented were able to show positive timing ability. The roots of the attribution analysis are in 1980's and the analysis has faced similar developments as the conditional market timing methods since. To better answer to questions if the fund is performing efficiently Menchero (2004) presented the multi-period approach on the side of the single-period approach. Recent developments have included the inclusion of the dynamic and static allocation factor to the side of traditional allocation (market timing).

This thesis gives a good starting point on reviewing the methods developed to detect market timing ability. It was interesting to notice the enthusiasm of certain authors to develop new models and aspects year after year. Given the stake of the fund manager, clients and the senior management it is understandable that the motivation to achieve models that could point out fund managers with timing ability is high. Going through the sea of articles, it seems that the motivation among the academics is and has been high as well. If there would have been more time to work among the subject, it would have been interesting to test these models with my own data. Comparable results could give interesting results as the results seen in this thesis are based on different fund data sets.

References

Aragon, G. and Ferson, W. (2006) Portfolio performance evaluation. *Foundations and Trends® in Finance* **2**(2), 83-190.

Bacon, C. (2008) Portfolio performace measurement attribution. *Wiley finance series*, 2nd edition, pp. 117-128.

Bacon, C. (2002) Excess returns – arithmetic or geometric. *Journal of Performance Measurement,* Spring, 23-31.

Brinson, G. and Fachler, N. (1985) Measuring non-US equity portfolio performance. *Journal of Portfolio Management* Spring, 73-76.

Brinson, G., Hood, R. and Beebower, G. (1986) Determinants of Portfolio Performance. *Financial Analysts Journal* July-August, 39-44.

Carino, D. (1999) Combining attribution effects over time. *The Journal of Performance Measurement,* Summer, 5-14.

Copeland, T. and Mayers, D. (1982) The value line enigma (1965-1978): a case study of performance evaluation issues. *Journal of Financial Economics* **10**, 289-321.

Cremers, M. and Petäjistö, A. (2009) How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies* **22**, 3329-3365.

Daniel, K., Grinblatt, M, Titman, S. and Wermers, R. (1997) Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* **52**, 377-411.

Elton, E., Gruber, M. and Blake, C. (2011) An examination of mutual fund timing ability using monthly holdings data. *Review Of Finance*, 1-27.

Fama, E. (1970) Efficient capital markets: A review of theory and empirical work. *Journal of Finance* **25**, 383-417.

Fama, E. (1972) Components of investment performance. *Journal of Finance* **27**, 551-567.

Ferson, W. (2010) Investment performance evaluation. Working paper. Center for Financial Innovation and Stability.

Ferson, W. and Khang, K. (2002) Conditional performance measurement using portfolio weights: Evidence for pension funds. *Journal of Financial Economics* **65**, 249-282.

Ferson, W. E. and Schadt, R. W. (1996) Measuring fund strategy and performance in changing economic conditions. *Journal of Finance* **51**, 425-462.

Goetzmann, W., Ingersoll, J., Spiegel, M. and Welch, I. (2007) Portfolio performance manipulation and manipulation-proof measures. *Review of Financial Studies* **29**(5), 1503-1546.

Grinblatt, M. and Titman, S. (1993) Performance measurement without benchmarks: an examination of mutual fund returns. *Journal of Business* **60**, 97-112.

Harvey, C. (1989) Time-varying conditional covariance in tests of asset pricing model. *Journal of Financial Economics* **24**, 289-317.

Henriksson, R. (1984) Market Timing and Mutual Fund Performance: An Empirical Investigation. *Journal of Business* **57**, 73-96.

Henriksson, R. and Merton, R. (1981) On market timing and investment performance. *Journal of Finance* **54**, 513-533.

Hsu, J., Kalesnik, V. and Myers, B. (2010) Performance Attribution: Measuring Dynamic Allocation Skill. *Financial Analysts Journal* **66**:6, 17-26.

Jensen, M. (1968) The performance of mutual funds in the period 1945-1964. *Journal of Finance* **23**, 389-416.

Jiang,W. (2003) A nonparametric test of market timing. *Journal of Emipirical Finance* **10**, 399-425.

Jiang, W., Yao, T. and Yu, T. (2007) Do mutual funds time the market? Evidence from portfolio holdings. *Journal of Financial Economics,* **86**, 724-758.

Kacperczyk, M., Nieuwerburgh, S. and Veldkamp, L. (2011) Time-Varying Fund Manager Skill. Working Paper. NBER.

Menchero, J. (2004) Multiperiod Arithmetic Attribution. *Financial Analysts Journal* **60**, 76-91.

Roll, R. (1978) Ambiquity when performance is measured by the securities market line. *Journal of Finance* **33**, 1051-1069.

Sharpe, W. (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* **19**, 425-442.

Sharpe, W. (1966) Mutual fund performance. *Journal of Business* **1**(part II), 119-138.

The Federation of Finnish Financial Services. (13.4.2011) ”*Sijoitusrahastot on suosittu sijoitusmuoto*”. Sited 13.9.2012 [http://www.fkl.fi/tietoa_meista/toimiala/sijoitusrahastot/Sivut/default.aspx]

Treynor, J. and Mazuy, M. (1966) Can mutual funds outguess the market? *Harvard Business Review* **44**, 131-136.