A RESEARCH ON INDIVIDUAL PENSION SYSTEM: AS A LONG TERM INVESTMENT ALTERNATIVE

ÖZGE ELVAN BALKAN

112626007

İSTANBUL BİLGİ ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

FİNANSAL EKONOMİ YÜKSEK LİSANS PROGRAMI

YRD. DOÇ. DR. DENİZ İKİZLERLİ

2017

A Research on Individual Pension System: As a Long Term Investment Alternative

Uzun Vadeli Yatırım Alternatifi Olarak Bireysel Emeklilik Sistemi Üzerine Bir Araştırma

Özge Elvan Balkan

112626007

Yrd. Doç. Dr. Deniz İkizlerli
(Tez Danışmanı)

Doç. Dr. Cenktan ÖZYILDIRIM

Yrd. Doç. Dr. Bora ERDAMAR

Onay Tarihi : 19.06.2017

Toplam Sayfa Sayısı : 47

Keywords

1) Individual Pension System
2) Pension Mutual Funds
3) Performance Evaluation
4) Benchmark
5) Hypothesis Test

Anahtar Kelimeler

1) Bireysel Emeklilik Sistemi
2) Emeklilik Yatırım Fonları
3) Performans Değerlemesi
4) Karşılaştırma Ölçütü
5) Hipotez Testi

Table of Contents

List of Tables

Abstract

In this study, the performances of pension mutual funds, which are operating in the Individual Pension System, are compared with their benchmark. Since the issue of which pension company performs better than others, is beyond the scope of the study, all funds are considered as a single fund under their own fund group. When we looked at the actualized returns of pension funds between 2007 and 2016, it was seen that pension funds performed better than their benchmarks. However, when the fund returns were tested at the 95% confidence interval, only 2 out of the 10 fund groups performed significantly better but in the 90% confidence interval this number rises to 4 out of 10.

Keywords: Individual Pension System, pension mutual funds, performance evaluation, benchmark, hypothesis test

Özet

Bu çalışmada, Bireysel Emeklilik Sistemi'nde faaliyet gösteren emeklilik yatırım fonlarının performansları, karşılaştırma ölçütlerinin getirileri ile karşılaştırılmıştır. Hangi emeklilik şirketinin diğerlerine göre daha iyi performans gösterdiği konusu, çalışmanın amacı dışında olduğu için 2007-2016 yılları arasında kesintisiz faaliyet gösteren tüm fonlar, kendi fon grubunun altında tek bir fon gibi değerlendirilmiştir. Türk emeklilik fonlarının 2007-2016 tarihleri arasında gerçekleşen getirilerine bakıldığında, genel olarak piyasadan daha iyi performans sergilediği görülmüştür. Ancak fon getirileri %95 güven aralığında test edildiğinde, 10 fon grubundan sadece 2 tanesinin istatiksel olarak daha iyi performans gösterdiği, %90 güven aralığında ise bu sayının 4'e çıktığı görülmüştür.

Anahtar Kelimeler: Bireysel Emeklilik Sistemi, emeklilik yatırım fonları, performans değerlemesi, karşılaştırma ölçütü, hipotez testi

1. INTRODUCTION

With the beginning of Individual Pension System in October 2003, the number of participants and the total asset value of funds have grown rapidly. The system, which had 1.99 million participants at the beginning of 2010, reached 2.64 million at the end of 2011 with a growth of 32.7%. In the same time frame, 9.12 billion TL total amount of fund size increased by 57% and reached to 14.33 billion TL.

With the "State Contribution Incentive" coming into force as of January 2013, the growth rate increasingly continued. According to the number of participants, an increase of 32.8% was achieved in only one year, which was narrowly achieved in 2 years between 2010 and 2011.

System was based on the voluntary participation until the end of 2016. By the beginning of 2017, automatic enrollment process has also begun. By the end of 2016, a total of 6.62 million volunteer participants had reached a fund size of 53.4 billion TL. The system, which already has a substantial growth potential, covenants a much greater potential, along with the automatic participation process. With the automatic enrollment, there is an expectation of similar growth rates, which had realized by the government contribution.

While industry expectation is in the positive direction, potential participants have concerns about the system. The first concern is, whether it would be better to invest their money in another instrument rather than entering Individual Pension System.

From this point of view, the Individual Pension System will be examined in this study as a long-term investment alternative. Our purpose is to evaluate the performance of Turkish pension funds and to show whether it is a good investment alternative.

The purpose of this study is not to determine which pension company is performing better but to see the performance of pension system as a whole. For this reason, the funds have not been examined on a company basis, but examined in an integrated manner. To calculate the weighted average returns, we used the proportion of each

fund's asset value to the total asset value of their own fund group. For the period of 2007-2016, the performance of 196 pension mutual funds' are examined on the basis of fund groups.

There are two main reasons for data selection. First one is, the system was started at the end of 2003 and funds showed too much instability for entering and leaving the system in the first years. The other reason is the stipulation criteria for entitlement to retirement benefits on Individual Pension System, which is to be a participant for at least 10 years. In other words, it is necessary that savings would be accumulated along with their returns for at least 10 years in the system. For this reason, it would be appropriate to examine the data of at least 120 months.

We used two methods to determine the benchmarks. First method is value-weighted benchmark determination which has been examined to see what has already actualized in the whole system. In this method, each fund's asset value in a specific fund group were proportioned to the total asset value of their own fund group. Then, the value-weighted benchmarks were calculated as a vectorial multiplication of asset value ratios and portfolio allocations of each funds. The other method is equally-weighted benchmark determination which has been examined in order to eliminate any neglected situation because of the dominant leader companies the market.

The generally accepted idea is that actively managed portfolios have a lower performance than passive index funds. Null and alternative hypotheses which were used in the test phase were also determined in this direction and the tests were completed.

According to the results of tests, 2 out of 10 fund groups showed statistically significant performances on their benchmarks at 95% confidence level, while 4 fund groups performed better at 90% confidence level.

On the other hand, the fact that benchmarks are equally-weighted or value-weighted has not had much effect on test results. It indicates that there is no remarkable difference in terms of pension companies' decision on portfolio allocations.

2. REVIEW OF LITERATURE

There are many studies which aim to find whether active portfolio management or passive portfolio management is more preferable. These studies are divided in two main parts. Some studies suggest that active management is more convenient, and others advocate the opposite.

One of the first studies for performance evaluation of mutual funds, was done by **Treynor (1965)**. He developed an index factor, namely Treynor ratio, which is based on the CAPM model and measured the performance of mutual funds relative to its systematic risk. He identified a characteristic line (Security Market Line – SML), which shows the relation between the fund return and its beta. The slope of this line gives us reward to volatility ratio (Treynor ratio) of the fund.

In consideration of these assumptions, the Treynor ratio explains the excess return of a portfolio on each unit of systematic risk taken.

$$T_p = (R_p - R_f) / \beta_p \ ,$$

where T_p, is the Treynor ratio of the portfolio, $(R_p - R_f)$ is the risk premium of the portfolio and β_p is the systematic risk of the portfolio.

When comparing two portfolios with same betas, the portfolio with higher Treynor's ratio is more effective, that is, it obtains higher return for one unit of systematic risk.

The second study about performance evaluation of mutual funds was made by **Sharpe (1966)**. He completed his study which was made under some assumptions of modern portfolio theory. His point was, if a fund manager can select the mispriced securities and do an effective diversification, then he/she can achieve a higher return than CAPM's estimation in a given risk class. To evaluate the performance of funds, he improved reward to variability ratio, which is called

"Sharpe Ratio". Sharpe ratio, unlike Treynor ratio, is not just about systemic risk, but about the total risk of portfolio.

$$S_p = (R_p - R_f) / \sigma_p \ ,$$

where S_p, is the Sharpe ratio of the portfolio, $(R_p - R_f)$ is the risk premium of the portfolio and σ_p is the total risk of the portfolio.

Sharpe completed his study with 34 open-end mutual funds, during the period 1954-1963. When the Sharpe ratios of the funds are examined, it is seen that the best performing fund's ratio is 0.78 and the worst performing fund's ratio is 0.43.

Then, he compared the Sharpe Ratios of the funds after expenses and benchmark (Dow Jones Industrial Average - DJIA) and saw that the reward to variability ratios of mutual funds are lower than the ratio of Dow Jones index. It was a sign that mutual funds cannot outperform the market.

Investors' expect managers to be able to predict market movements and whether they will enter the market in time. Investors' expectation from fund managers is to predict the market movements and to take the right position at the right time. For testing this ability of fund managers, **Treynor and Mazuy (1966)** completed a paper, which used 57 funds during the period 1953-1962. Moreover, they produced a quadratic model, which is based on CAPM, to measure mutual fund's performance.

$$R_p = \alpha_p + R_{f+} \beta_p (R_m - R_f) + \gamma_p (R_m - R_f)^2 + \varepsilon_p \ ,$$

where $(R_m - R_f)$ shows the excess return of market above risk-free rate, α_p represents the selectivity ability and γ_p represents market timing ability of a fund manager.

Positive γ_p indicates that the fund manager has market timing capability and negative γ_p indicates that fund manager is insufficient for market timing.

Even if their study cannot show any statistical evidence that any of the 57 funds have outperformed the market, they claimed that investors can beat the market. The results showed that investors were highly correlated with market fluctuations. So,

since the market is bullish, if a fund manager hold a portfolio with high beta (β) and diversified the individual risk of his/her portfolio, the portfolio would be hard-charging.

Jensen (1968) measured the performance of mutual funds in order to determine whether the fund manager could use his/her own skills to get abnormal returns. While exploring the performance of mutual funds, he used an improved version of Capital Asset Pricing Model (CAPM).

$$R_p = \alpha_p + R_{f+} \beta_p (R_m - R_f) + \varepsilon_p ,$$

Above formulation, except from a_p, calculates the portfolio's rate of return under its systematic risk and ignores the individual risk of the portfolio P. So, if fund managers can find a portfolio with positive α, they can get higher returns than CAPM's prediction.

Jensen worked with 115 mutual funds during the period 1945-1964 and S&P 500 index as benchmark. The results shows that the majority of these mutual funds do not have positive alphas (α), in other words cannot beat the market. Only 39 of these funds had average return, so there is a weak evidence that portfolio managers can gain abnormal returns.

In 1965, Fama became the eponym researcher of the term "Efficient Market". Roberts (1967) declared 'Efficient Market Hypothesis' for the first time and reviewed this approach into two distinct forms, weak form and strong form. **Fama (1970)** added another form, namely 'semi-strong form', and completed three classicized forms of Efficient Market Hypothesis (EMH). EMH claimed that, security prices "fully reflect" all available information. If there is appropriate information that will affect securities' prices, this effect would be reflected in prices quickly and correcty. Fama tested the efficiency for three subsets, to find out which level of information will refute the EMH, if there is any. Weak form of EMH implies that past price cannot predict future price, in other words, no one can gain excess returns with the help of historical data or technical analysis. On the other side, semi-strong form implies that any of public information can cause higher

returns in addition to historical data. According to assumptions of EMH, a public information already reflect the price, so investors cannot gain higher returns on this kind of information.

The third and last form of EMH is strong form. This form argues that any abnormal profit can be obtained by using historical data, public information or insider information. In another words, current prices already reflected by all kinds of available information. If market is strong form efficient, even insider information would not be beneficial for an investor who wants higher returns than the market. Fama's study supported the hypothesis of weak and semi-strong form. He also acknowledged that the strong form is not the true counterpart of the real world. Except for a few exceptional cases, he concluded that the efficient market model hold.

On the other hand, **Carlson (1970)** thought that Jensen's study might be biased because of the time period and benchmark. So he simulated Jensen's study for 82 mutual funds in between 1948-1967. First of all, he generated three types of fund groups with mutual funds; such as stock funds group, balanced funds group and income funds group. He also used three different benchmarks; NYSE, S&P500, NYSE composite or Dow Jones Industrial Average (DJIA). His regression shows us that average net alpha of funds is 60 basis points per annum, so mutual funds can beat the market. Carlson rebutted Jensen's (1968) result and he denoted that the time period and benchmark have an effect upon the mutual funds' performance.

McDonald (1974) queried if there is a relationship between a fund's objective and its risk-return attitude. To find the answer, he examined 123 mutual funds for the period 1960-1969. He found a significant positive relationships between funds' objective and risk, and also return and risk. He used Sharpe ratio, Treynor ratio and Jensen's alpha for evaluating the performance of funds. 67 funds showed higher performance than the market according to Treynor ratio, while 39 funds performed better than the market according to Sharpe ratio.

Statistical tests have been conducted by a number of researchers on whether CAPM is sufficient to explain the return of assets. CAPM began to be criticized, because some of the result of these tests contradict to CAPM's assumptions and formula. In consequence of negative criticisms, Arbitrage Pricing Theory (APT) was developed by **Ross (1976)**. He derived the below formula;

$$R_X = R_f + \lambda_1\beta x_1 + \lambda_2\beta x_2 + \ldots + \lambda_K\beta x_K \qquad | \; x=1, 2, 3 \ldots K \;,$$

where λ_i is the risk premium of the i^{th} risk factor, βx_i is the factor-dependent sensitivity of Asset X.

In some cases, APT is similar to CAPM. APT stands behind the assumption that firm specific risk is easily broken away with a well-diversified portfolio. It considers only the systematic risk and firm specific risk is not a factor that affects the asset's return. However, APT claims that there are several non-diversifiable risk factors that affect the return of a security or portfolio. APT gives us more meaningful results in diversified portfolios, because it takes into account the macroeconomic and systematic factors.

Kon and Jen (1979) used the Sharp-Lintner-Mossine (SLM) and Black models, to show that systematic risk of a mutual fund may not be stationary, but nonstationary. They examined 49 mutual funds and saw that non-linear regression models better fit the sample of data than a unique linear model. Their findings explained that, 37 funds had different level of systematic risk and they decided that most of the funds involved in market timing activity. Later, **Kon (1983)** completed a study himself, testing not only timing but also selectivity for 37 mutual funds in 1960-1976. He found timing ability for 14 funds, but the results were statistically insignificant. At the same time, he found selectivity ability for 23 funds, but only 5 of them were significant.

Another approach for evaluating timing performance, which was completed in two steps, was introduced by Merton (1981) and derived by **Merton and Henriksson (1981)**. The model is similar to the model of Treynor and Mazuy (1966). In despite of a linear beta, Henriksson and Merton's model assumed that the beta of a portfolio

is variable and switch between two values that are based on market movement. They presumed that fund managers can only predict the trend of market (bullish or bearish) but not length of the trend. To set up the model, they used a dummy variable "D", which is equal to 1 when the market is bullish and 0 when market is bearish.

$$R_p = \alpha_p + R_f + \beta_p (R_m - R_f) + \gamma_p (D (R_m - R_f)) + \varepsilon_p \ ,$$

γ measures market timing ability, like Treynor and Mazuy formula.

Most of the researchers investigated market timing and security selection abilities of funds' managers separately. The minority of researchers thought that it was not an obligation to test these abilities individually. **Bhattacharya and Pfleiderer (1983)** derived a model to test the individual abilities of assets managers. Contrary to Jensen (1972), a simple regression technique could be used to measure stock selection ability and market timing ability, like Treynor and Mazuy (1966) suggested before. They eliminated a shortage in Jensen's work and showed how to use a simple regression technique to get accurate measure of timing and selection ability. In the developed model, timing ability can be measured with a quadratic correlation between funds managers' prediction and realized excess market return, and selection ability can be directly measured from realized excess market return.

Change and Lewellen (1984) used Merton and Henriksson's (1981) method to investigate the performance of 67 mutual funds between 1971 and 1979. They divided the date range into two parts, as up-market and down-market, and examined these 67 funds' returns based on market movements, monthly and quarterly. There were 52 up-periods and 56 down-periods between 1971 and 1979.

The result of market timing ability of fund managers, only 4 funds were statistically significant, although 42 of the 67 funds obtained higher returns than the regression's estimation. The result of security selection skill of fund managers was also parallel to the first test results, only 5 funds were statistically significant, although 41 of the 67 funds obtained higher returns than the regression's estimation.

Their results were consistent with the predictions of the Merton and Henriksson's model. Findings indicated that there is no statistical evidence of the existence of fund managers' market timing skills or asset selection skills.

Lehmann and Modest (1987) tried to ascertain whether the rankings of managed funds are sensitive to the chosen benchmark. They emphasized that numerous investigators, like Sharpe, Jensen, Treynor, Treynor and McDonald had employed the standard CAPM benchmarks. And they protested that the use of the CAPM is insufficient because of its assumptions. The insistence of CAPM is that abnormal returns can only occur if the market is inefficient. The second theory which was used in studies is Arbitrage Pricing Theory. Unlikely CAPM, the Theory defended that the risk of a security, should be represented by a multiple factor. So they employed the standard CAPM benchmark and two varieties (5-factors and 10-factors) of APT benchmarks. They used 130 mutual funds over the period 1968-1982.

The authors concluded that the choice of benchmark has significantly impact on performance evaluation. The results shows us that Jensen measure is very sensitive to APT benchmarks, while it is insensitive to the number of common factors (5-10-15) in APT benchmark.

Grinblatt and Titman (1989) measured the net and gross performance of 274 mutual funds between 1975 and 1984 on the basis of quarterly returns and Jensen measure. They also used 4 different benchmarks, CRSP equal-weighted, CRSP value-weighted, Lehmann & Modest ten factor portfolios and eight-portfolio benchmark based on firm size, yield, and past returns. According to the findings of the study, the returns of aggressive growth funds, growth funds and small sized funds were significantly higher than the average return of benchmarks and somehow there is an inverse proportion between fund size and its gross return. After the costs were deducted, funds could not outperform the market in net returns, which support the Efficient Market Hypothesis.

There were also researchers who thought that CAPM's inadequacy could be developed. With this in mind, **Fama and French (1992)** added two more factors to the standard CAPM, namely three-factor model. They assumed that size (SMB - small minus big) and book to market value (HML – high minus low) affects the price of a security. When they tested this three-factor model, they saw that they were right in their suspicions and this new model was much more descriptive than the standard CAPM.

$$R_p = \alpha_p + R_{f+} \beta_{p_1} (R_m - R_f) + \beta_{p_2} (SMB) + \beta_{p_3} (HML) + \varepsilon_p \ ,$$

where SMB is the size premium, β_{p_2} is the size sensitivity of security P, HML is the value premium and β_{p_2} is the value sensitivity of security P.

Unlikely many other studies, **Grinblatt and Titman (1993)** created a model (which is called Portfolio Change Measure - PCM) in which they could measure performance of a fund without necessity for a benchmark, since they predicted that a performance of an asset is sensitive to the selected benchmark.

$$PCM = \sum_t \sum_i (R_{i,t} (w_{i,t} - w_{i,t-j}) / T) \ ,$$

where $R_{i,t}$ is the rate of return on asset i at time t, $w_{i,t}$ and $w_{i,t-j}$ are the portfolio weights at the beginning of period t and t-j respectively.

They used the same dataset, which they used in their previous study in 1989. The results of the research were consistent with the previous study. For instance, the strongest evidence of superior performance was again found for aggressive growth funds. But this superior performances of some fund groups do not mean that investors earn abnormal net returns from these funds because high-yield funds have also high costs.

Not only mutual funds' performance, but also studies of pension funds' performance were examined. **Coggin, Fabozzi and Rahman (1993)** studied the performance of 71 pension funds between 1983 and 1990. They used two different models (Treynor & Mazuy and Bhattacharya & Pfleiderer) and 4 different benchmarks to test the sensitivity of performances according to these variables.

They thought that models are inadequate because of ignoring the securities' negative market timing. They gave reference to Coggin and Hunter's study to support their idea. They modified and used the models in order to prevent this inadequacy and to measure the negative timing ability. One of the major findings of this study was that the selectivity measure was higher and timing measure was lower than the average. Secondly, they found that the ability of timing and selectivity were seemed to be sensitive to the chosen benchmark. Lastly, they found negative correlation between selectivity and timing.

Malkiel (1995) investigated the performance, performance persistency and cost of all kinds of mutual funds in US market which were offered to the public, during 1971-1991 period on the base of Capital Asset Pricing Model. Study showed that the funds could not beat the market according to their pre-cost and after-cost returns.

Additionally, he noticed that there are some cracks. First of all, funds' returns had positive persistence in short time period and negative persistence of long time period. This attitude is contradicted to EMH, which claims that historical prices and returns would not provide any information to predict future price and returns. But the performance persistency applied only in the 1970s, but it was not valid for the 1980s.

Other cracks were some seasonal and day-of-the-week patterns and foreseeability of funds' returns based on their firm size, price to earnings (P/E) ratios, and price to book value (P/B) ratios.

Four years after Fama & French's study, **Carhart (1997)** extended Fama and French's three-factor model with an additional risk factor, the one-year momentum (MOM) effect.

$$R_p = \alpha_p + R_f + \beta_{p_1}(R_m - R_f) + \beta_{p_2}(SMB) + \beta_{p_3}(HML) + \beta_{p_4}(MOM) + \varepsilon_p$$

He examined 1892 funds in 1962-1993 and classified funds into 3 fund categories; aggressive growth, long term growth and growth-income. He measured the performance with two different methods; CAPM and his four-factor model. The

results showed that the four-factor model better fit the data more than standard CAPM. He also indicated that there is persistence in short term funds' returns. He sorted the funds according to their returns. If a fund manager has a long position on first decile funds and a short position on a last decile funds, the fund manager would earn an annual return of 8%. On the other hand, he could not find any persistence in long term. If a fund manager practiced the same strategy in the long term, he/she would have a much smaller spread. He also reiterated that there is a significant and negative correlation between total cost of managing fund and its return.

Modigliani & Modigliani (1997) thought that performance of two securities (for example; a fund and its benchmark) can be compared if and only if they have the same level of risk. So they set up a risk-adjusted performance measurement ratio, which was derived from Sharpe's study in 1966. It adjusts the risk of a given portfolio by matching it with the risk of a benchmark and then calculates the return for that portfolio.

$$M^2 = \frac{\sigma_m}{\sigma_p}\left(R_p - R_f\right) + R_f$$

The idea behind this method was that if a fund manager received higher returns than its same-risk-benchmark, it means that the manager has the selectivity ability. Thereby, if M^2 is greater than R_m, it certifies that the fund/portfolio manager reach the abnormal returns according to market.

In the traditional methods which I have reviewed above, unconditional methods were used to measure fund performance. In contrast to these studies, **Ferson and Schadt (1996)** thought that using unconditional methods for evaluating the performance of a security might be dysfunctional. Just as return and risk of a security can change over time, the alpha or other factor sensitivities should also be measured with a condition.

Their method was based on the same assumptions with the semi-strong form of Efficient Market Hypothesis. The reason for choosing these assumptions was to separate fund managers' real market timing ability from standard public

information. They formulized the risk sensitivities of a security with a multiplication of time-varying vectorial variables. They adapted this approach to CAPM, APT and multi-factor models. For example, the adaption of multiple-factors model can be formulated as follows;

$$E(R_{i,t+1}|Z_t) = \alpha_0(Z_t) + \sum_{j=1}^{K} \beta_{i,j}(Z_t) + \alpha_j(Z_t) \qquad \text{i=0,1,2,...,K and t=0,1,2,...,T-1}$$

To test their approach, they used 67 open-end mutual funds' monthly returns between 1968 and 1990. The findings supported the main theme of the study. The use of conditioning information in performance measurement is both statistically and economically significant.

Most of the studies took side with the Efficient Market Hypothesis and certified that actively managed funds could not outperform the index funds. **Gruber (1996)** tried to find some tangible reasons to why people tend to invest actively managed funds. He offered four reasons for this question. First three of these reasons were, lower managerial cost, customer services and diversification opportunity. Even if a fund manager has any of the management abilities, the managerial cost would be lower than the investors' own transaction cost. Customer service also offers record keeping and transferring the money into or out of funds conveniently. These there reason could not discriminate between actively managed funds and passive index funds. But the fourth and last reason was the professional portfolio management, which makes the main difference between types of portfolio managements. He used three different measures for the performance of funds; abnormal returns relative to market return, excess return of fund from a single-factor model and excess return of fund from four-factor model. His sample included 270 funds in 1985-1994. The main result of the study was that, funds could not outperform the market. Relative to average market return, funds underperformed the market by 1.94%. According to single-factor model and four- factor model, the underperformance of funds were 1.56% and 0.65% respectively. So, he reinforced that actively managed funds cannot outperform the market.

Some researchers thought that the issue of outperforming depends on luck, while others claimed that it is about the managers' skill. **Fama and French (2010)** conducted a study to distinguish skill and luck. They moot the idea about the weakness of persistence tests. They argued that the previous studies about performance persistency did not extend to long period, so these studies were unsatisfactory for measuring short-term persistency. To overcome this circumstance, they worked with a sample of 3.156 funds in 1984-2006, covering more than 20 years. They also used two different models, their three-factor and Carhart's four-factor model. The result of study contained that, the average net α of the funds was -0.81% while the gross α of the funds was 0.13% according to their three-factor model. They concluded that management fees and expenses reduce the returns of funds and managers were not skillful enough to cover their fees/expenses.

3. DATA AND METHODOLOGY

3.1 Data

The purpose of this study is not to understand which pension company is performing better but to see the performance of pension system as a whole. For this reason, the funds have not been examined on a company basis, but examined in an integrated manner, taking into account their weighted average. To calculate the weighted average returns, we used the proportion of each fund's asset value to the total asset value of their own fund group. The performance of 196 pension mutual funds' examined on the basis of fund group.

Each fund is assessed within its own fund group which are listed below and these groups are considered as a single fund.

- Government Bonds and Bills Funds
- Stock Funds
- Flexible Funds
- Government Bonds and Bills (FX) Funds
- Standard Funds
- Liquid Funds-Government
- Liquid Funds-Flexible
- Balanced Funds
- Index Funds
- Composite Bonds And Bills Funds

In the table below, we can see fund groups and the number of companies operating for each fund groups. The number of companies operating in some fund groups is few. Along with reminding, the system was considered within its integrity, even so, that kind of funds groups were also taken into consideration, since the funds were

continuously active during the period 2007-2016. These funds were still taken into account as the system was considered in its integrity and the related funds were operating continuously.

Table 1 : The number of operating companies and funds according to fund title

Fund Title	Number of Companies	Number of Funds
Government Bonds and Bills Funds	14	24
Stock Funds	17	26
Flexible Funds	17	79
Government Bonds and Bills (FX) Funds	12	20
Standard Funds	16	17
Liquid Funds-Government*	6	7
Liquid Funds-Flexible	11	15
Balanced Funds*	2	2
Index Funds*	2	3
Composite Bonds And Bills Funds*	2	3

*Although there are few companies and a small number of funds in some fund groups, they were also included in this study, because they were operating continuously during the period of 2007-2016.

The study has considered the monthly returns of the fund groups which were operating in Turkish Individual Pension System, during the period 2007-2016.

There are two main reasons for data selection. First one is, the system was started at the end of 2003 and funds showed too much instability for entering and leaving the system in the first years. The other reason is, the stipulation criteria for entitlement to retirement benefit in Individual Pension System, which is to be a participant for at least 10 years. In other words, it is necessary that savings would be accumulate along with their returns for at least 10 years. For this reason, in order to take the funds in our sample into the analysis, it would be appropriate to examine the data of at least 120 months.

The second thing that is necessary to shape the model is to decide about the benchmarks. The benchmarks for fund titles are based on the portfolio allocations for each fund in the data taken from the Capital Markets Board of Turkey (CMB). Each fund's portfolio allocations have a specific percentage of the following benchmarking criteria.

- Stock (HS)
- Government Debt Securities on Foreign Currencies (KD)
- Government Debt Securities on Turkish Lira (KB)
- Reverse Repo Transactions (TR)
- Debt Securities Of Corporations (OZ)

The benchmarks selected for each comparison criterion are as follows; BIST 100 for Stock, 2-year benchmark bond on TL for Government Debt Securities on TL, 2-year foreign exchange benchmark bond for Government Debt Securities on Foreign Currencies, O/N Gross Repo Index for Reverse Repo Transactions and Private sector benchmark bond for Debt Securities of Corporations.

If the distribution of a fund portfolio is sensitive to multiple comparison measures, it would not be sufficient to compare it with a single benchmarking criterion. For instance, if it is known that the stock fund is sensitive to more than one benchmark criteria aside from stocks (for our example, the other criteria are Government Debt Securities on TL and Reverse Repo Transactions), it would not be completely accurate to compare stock fund with only BIST100.

The table below explains the benchmark criterion content of each fund group, reminding that portfolio allocations are different on the basis of fund's title.

Table 2 : Benchmarking criteria in each fund groups' portfolio

		Benchmarking Criteria				
		Stock (HS)	Government Debt Securities on Foreign Currencies (KD)	Government Debt Securities on Turkish Lira (KB)	Reverse Repo Transactions (TR)	Debt Securities Of Corporations (OZ)
Fund Group	Government Bonds and Bills Funds	√		√	√	√
	Stock Funds	√		√	√	
	Flexible Funds	√	√	√	√	√
	Government Bonds and Bills (FX) Funds		√		√	
	Standard Funds	√		√	√	√
	Liquid Funds-Government	√		√	√	√
	Liquid Funds-Flexible	√		√	√	√
	Balanced Funds	√		√	√	√
	Index Funds	√			√	
	Composite Bonds And Bills Funds	√	√		√	√

The portfolio allocations of the each fund is different from day to day and from company to company. In order to handle this situation, two methods have been used to determine the benchmarks, in the light the studies of Lehmann&Modest (1987) and Coggin&Fabozzi&Rahman (1993).

1. Value-weighted benchmark determination
2. Equally-weighted benchmark determination

The first method is value-weighted benchmark determination, which has been examined to see what has already been actualized in the whole system. In this method, each fund's asset value in a specific fund group were proportioned to the

total asset value of their own fund group. Then, the value-weighted benchmarks were calculated as a vectorial multiplication of asset value ratios and portfolio allocations of each funds. The second method is equally-weighted benchmark determination, which has also been examined in order to eliminate any neglected situation because of the dominant leader pension companies in the sector. In this method, Portfolio benchmarks were calculated with equal weights, regardless of the fund size of the pension companies. Since the funds are considered on their monthly returns, the benchmarks' returns are also calculated on a monthly basis.

3.2 Methodology

Our data samples for both funds and their benchmarks consist of 120 months. We used the sample of 10-year data out of 15-year-old system (population). So population and sample are not so different from each other. When the variances of the sample of fund and sample of benchmark are calculated, we saw that they are unequal. For this reason, it is assumed that the variances of both populations are also unequal. Therefore, in the test phase of the study, two sample upper-tailed t-tests were used, assuming different populations' variances between each fund group and its benchmark.

Null and alternative hypotheses are determined as follows;

H_0 (Null Hypothesis) : $\mu_{PMF} \leq \mu_{Benchmark}$

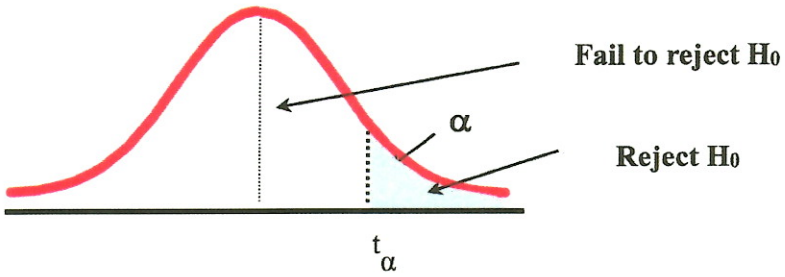
H_1 (Alternative Hypothesis) : $\mu_{PMF} > \mu_{Benchmark}$,

where μ_{PMF} is the mean return of the Pension Mutual Fund (PMF) and $\mu_{Benchmark}$ is the mean return of Pension Mutual Fund's benchmark. To find the t-statistic of the model, we should determine the degrees of freedom with the below formula;

$$v = \frac{\left[\left(\frac{S^2_{PMF}}{N_{PMF}}\right) + \left(\frac{S^2_{Benchmark}}{N_{Benchmark}}\right)\right]^2}{\left(\frac{S^2_{PMF}}{N_{PMF}}\right)^2 /(N_{PMF} - 1) + \left(\frac{S^2_{Benchmark}}{N_{Benchmark}}\right)^2 /(N_{Benchmark} - 1)},$$

where S_{PMF} is the standard deviation of Pension Mutual Fund's sample, $S_{Benchmark}$ is the standard deviation of benchmark's sample, N_{PMF} is the sample size of Pension Mutual Fund and $N_{Benchmark}$ is the sample size of benchmark.

Two α-values were used, so the tests were performed for 90% and 95% confidence level. T-statistic values for each fund were found. Tests ended with a comparison of t-statistic values with t-critical values.



As can be seen from the above graph, if the t-statistic value is greater than the t-critical value, we reject the null hypothesis. If the t-statistic value is less than the t-critical value, we fail to reject the null hypothesis.

4. RESULTS

Before the tests, it was seen that each of the fund groups outperformed their benchmarks. However, with the assistance of tests, we can see if these superior returns were statistically significant.

Tests were conducted to compare the returns of fund groups, with both value-weighted benchmarks and equally-weighted benchmarks. If we look at the value-weighted benchmark comparison, only 2 (Composite Bonds and Bills Funds and Government Bonds and Bills on Foreign Currencies Funds) out of 10 fund groups' superior performance are statistically significant at 95% confidence level. But at 90% confidence level, the number rises to 4 groups (Composite Bonds And Bills Funds, Government Bonds and Bills on Foreign Currencies Funds, Standard Funds and Government Bonds and Bills Funds) that generated superior returns compared to their benchmarks.

Table 3 : T-tests result for value-weighted benchmark comparison*

Fund Group	**Degrees of freedom**	**t-Stat**	**At %90 Confidence Level ($t_{0,1}$=1,282)**	**At %95 Confidence Level ($t_{0,05}$=1,645)**
Government Bonds and Bills Funds	128	1.602	S	NS
Stock Funds	238	0.119	NS	NS
Flexible Funds	199	1.275	NS	NS
Government Bonds and Bills (FX) Funds	205	2.430	S	S
Standard Funds	125	1.526	S	NS
Liquid Funds-Government	127	1.040	NS	NS
Liquid Funds-Flexible	126	0.407	NS	NS
Balanced Funds	185	1.117	NS	NS
Index Funds	194	0.214	NS	NS
Composite Bonds And Bills Funds	210	2.239	S	S

*S for significantly outperform the benchmark, NS for not significantly outperform the benchmark

On the other part of tests, the same results were obtained in the equally-weighted benchmark comparison at each level of confidence. Only 2 (Composite Bonds And Bills Funds and Government Bonds and Bills on Forein Currencies Funds) out of 10 fund groups' significantly generated higher returns than their benchmarks at 95% confidence level. But, at 90% confidence level, the number of funds with statistically superior performance rises to 4 (Composite Bonds And Bills Funds, Government Bonds and Bills on Foreign Currencies Funds, Standard Funds and Government Bonds and Bills Funds) out of 10.

Table 4 : T-tests result for equally-weighted benchmark comparison*

Fund Group	**Degrees of freedom**	**t-Stat**	**At %90 Confidence Level ($t_{0,1}$=1,282)**	**At %95 Confidence Level ($t_{0,05}$=1,645)**
Government Bonds and Bills Funds	127	1.604	S	NS
Stock Funds	238	0.129	NS	NS
Flexible Funds	204	1.222	NS	NS
Government Bonds and Bills (FX) Funds	212	2.525	S	S
Standard Funds	126	1.516	S	NS
Liquid Funds-Government	129	1.238	NS	NS
Liquid Funds-Flexible	137	1.118	NS	NS
Balanced Funds	184	1.111	NS	NS
Index Funds	195	0.216	NS	NS
Composite Bonds And Bills Funds	215	2.244	S	S

*S for significantly outperform the benchmark, NS for not significantly outperform the benchmark

As it can be seen from the two tables above, the test results for both benchmarks are very close to each other. The sub-finding of the study consists on that there is no remarkable difference in the management strategy (determining portfolio allocations) between sector leader companies and other companies.

The other sub-finding of the study, is that the risk levels of pension mutual funds are less than the benchmarks. It is briefly stated that pension mutual funds have relatively higher returns with a relatively lower risk level without any exception.

5. CONCLUSIONS

In this study, we analyzed a sample of 196 pension funds that are continuously operating in the Turkish Individual Pension System, over the period of 2007-2016. We separated funds into 10 fund groups and examined the performances of the fund groups with reference to their benchmark returns.

When we look at the actualized 10-year returns, regardless of the test results, it is seen that the performances of funds' were better than their benchmarks. However, only 2 out of 10 fund groups are statistically significant at 95% confidence level, when 4 out of 10 funds groups are statistically significant at 90% confidence level.

Nevertheless, it can be described as a noteworthy opportunity to maintain the same return persistence for 10 years. According to that, it can be a good reason to join the Individual Pension System.

Participants of system are also divided in terms of their point of view on the Individual Pension System. The first group regards the system as a saving instrument and the second group regards it as an investment instrument. It is seen that pension mutual funds provides higher returns with a lower risk as compared to their benchmark which are a portfolio of passively managed investment instruments.

6. REFERENCES

Treynor Jack L., 1965, "How to Rate Management of Investment Funds", Harvard Business Review, 43(1), p. 63-75

Sharpe, W., 1966, "Mutual Fund Performance," The Journal of Business, 39, p. 119-138

Jensen, M., 1968, "The Performance of Mutual Funds in the Period 1945-1964," The Journal of Finance, 23, p. 389-416

Carlson, R., "Aggregate Performance of Mutual Funds, 1948-1967," Journal of Financial and Quantitative Analysis, p. 1-32

Treynor, J. and K. Mazuy, 1966, "Can Mutual Funds Outguess the Market?" Harvard Business Review, July, p. 131-136

McDonald, J., 1974, "Objectives and Performance of Mutual Funds, 1960-1969," Journal of Financial and Quantitative Analysis, p. 311-333

Kon, S. and F. Jen, 1979, "The Investment Performance of Mutual Funds: An Empirical Investigation of Timing, Selectivity and Market Efficiency," The Journal of Business, 52, p. 263-289.

Kon, S., 1983, "The Market-Timing Performance of Mutual Fund Managers," The Journal of Business, 56, p. 323-347.

Henriksson, R.D. and R.C.Merton.(1981). "On Market Timing and Investment Performance.II.Statistical Procedures For Evaluating Skills". Journal of Business, 54.No.4, p. 513-553

Lehmann, B. and D. Modest, 1987, "Mutual Fund Performance Evaluation: A Comparison of Benchmarks and Benchmark Comparisons," The Journal of Finance, 42, p. 233-265.

Grinblatt, M. and S. Titman, 1989, "Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings," The Journal of Business, 62, p. 393-416.

Ross, S., 1976. The Arbitrage Theory of Capital Asset Pricing, Journal of Economic Theory 13, p. 341–60

Fama, E. F.; French, K. R., 1993. Common Risk Factors in the Returns on Stocks and Bonds, Journal of Financial Economics 33, p. 3-56

Fama, E. F., French, K. R., 1992. The Cross-Section of Expected Stock Returns, Journal of Finance, 47, p. 427-465.

Carhart, M., 1997, "On Persistence in Mutual Fund Performance," The Journal of Finance, 52, p. 57-82.

Malkiel, B., 1995, "Returns from Investing in Equity Mutual Funds: 1971-1991," The Journal of Finance, 50, p. 549-572

Modigliani, F., & Modigliani, L. (1997). Risk-adjusted performance. The Journal of Portfolio Management, 23(2), p. 45-54.

Chang, E. and W. Lewellen, 1984, "Market Timing and Mutual Fund Investment Performance," The Journal of Business, 57, p. 57-72

Grinblatt, M. and S. Titman, 1993, "Performance Measurement without Benchmarks: An Examination of Mutual Fund Returns," The Journal of Business, 66, p. 47-68.

Ferson, W. and R. Schadt, 1996, "Measuring Fund Strategy and Performance in Changing Economic Conditions," The Journal of Finance, 51, p. 425-461

Coggin, T., Fabozzi, F., and Rahman, S., 1993, The Investment Performance of U.S. Equity Pension Fund Managers: An Empirical Investigation, Journal of Finance, 48, p.1039-1055

Bhattacharya S. and P. Pfleiderer (1983) A Note on Performance Evaluation, Stanford University, Technical Report 714, p. 1-23

Gruber, M., 1996, "Another Puzzle: The Growth in Actively Managed Mutual Funds," The Journal of Finance, 51, p. 783-810.

Fama, E., and French, K. (2010), Luck versus Skill in the Cross Section of Mutual Fund Returns, Journal of Finance, 65, p. 1915-1947

7. APPENDIX

7.1 Value-weighted benchmark comparison

Table 5: t-Test Results for Standard Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Standard Funds	*Standard Funds' Benchmark*
Mean	0,736%	-0,213%*
Variance	0,012%	0,451%
Observations	120	120
Hypothesized Mean Difference	0	
df	125	
t Stat	1,526	
P(T<=t) one-tail	0,06470111	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Standard Funds	*Standard Funds' Benchmark*
Mean	0,736%	-0,213%*
Variance	0,012%	0,451%
Observations	120	120
Hypothesized Mean Difference	0	
df	125	
t Stat	1,526	
P(T<=t) one-tail	0,06470111	
t Critical one-tail	1,645	

*Under normal conditions, as it is a long-term and passively managed portfolio, the average return of benchmark should not be negative. But, because of the economic instability in 2007, 2009 and 2014, Government Debt Securities on TL had negative return in these years. Majority of the benchmark's portfolio allocation consists of Government Debt Securities on TL. As an effect of these reasons, the average of benchmark' return was also calculated as negative.

Table 6: t-Test Results for Liquid Funds (Government)

t-Test Result for α=0,10 (at %90 Confidence level)

	Liquid Funds (Government)	*Liquid Funds' (Government) Benchmark*
Mean	0,725%	0,551%
Variance	0,001%	0,032%
Observations	120	120
Hypothesized Mean Difference	0	
df	127	
t Stat	1,040	
P(T<=t) one-tail	0,15019601	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Liquid Funds (Government)	*Liquid Funds' (Government) Benchmark*
Mean	0,725%	0,551%
Variance	0,001%	0,032%
Observations	120	120
Hypothesized Mean Difference	0	
df	127	
t Stat	1,040	
P(T<=t) one-tail	0,15019601	
t Critical one-tail	1,645	

Table 7: t-Test Results for Liquid Funds (Flexible)

t-Test Result for α=0,10 (at %90 Confidence level)

	Liquid Funds (Flexible)	*Liquid Funds' (Flexible) Benchmark*
Mean	0,747%	0,678%
Variance	0,001%	0,033%
Observations	120	120
Hypothesized Mean Difference	0	
df	126	
t Stat	0,407	
P(T<=t) one-tail	0,34241799	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Liquid Funds (Flexible)	*Liquid Funds' (Flexible) Benchmark*
Mean	0,747%	0,678%
Variance	0,001%	0,033%
Observations	120	120
Hypothesized Mean Difference	0	
df	126	
t Stat	0,407	
P(T<=t) one-tail	0,34241799	
t Critical one-tail	1,645	

Table 8: t-Test Results for Composite Bonds And Bills Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Composite Bonds And Bills Funds	*Composite Bonds And Bills Funds' Benchmark*
Mean	0,973%	0,120%
Variance	0,055%	0,119%
Observations	120	120
Hypothesized Mean Difference	0	
df	210	
t Stat	2,239	
P(T<=t) one-tail	0,01309208	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Composite Bonds And Bills Funds	*Composite Bonds And Bills Funds' Benchmark*
Mean	0,973%	0,120%
Variance	0,055%	0,119%
Observations	120	120
Hypothesized Mean Difference	0	
df	210	
t Stat	2,239	
P(T<=t) one-tail	0,01309208	
t Critical one-tail	1,645	

Table 9: t-Test Results for Government Bonds and Bills (FX) Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Government Bonds and Bills (FX) Funds	*Government Bonds and Bills (FX) Funds' Benchmark*
Mean	1,011%	0,122%
Variance	0,048%	0,112%
Observations	120	120
Hypothesized Mean Difference	0	
df	205	
t Stat	2,430	
P(T<=t) one-tail	0,00797268	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Government Bonds and Bills (FX) Funds	*Government Bonds and Bills (FX) Funds' Benchmark*
Mean	1,011%	0,122%
Variance	0,048%	0,112%
Observations	120	120
Hypothesized Mean Difference	0	
df	205	
t Stat	2,430	
P(T<=t) one-tail	0,00797268	
t Critical one-tail	1,645	

Table 10: t-Test Results for Government Bonds and Bills Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Government Bonds and Bills Funds	*Government Bonds and Bills Funds' Benchmark*
Mean	0,785%	-0,186%*
Variance	0,016%	0,425%
Observations	120	120
Hypothesized Mean Difference	0	
df	128	
t Stat	1,602	
P(T<=t) one-tail	0,05585738	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Government Bonds and Bills Funds	*Government Bonds and Bills Funds' Benchmark*
Mean	0,785%	-0,186%*
Variance	0,016%	0,425%
Observations	120	120
Hypothesized Mean Difference	0	
df	128	
t Stat	1,602	
P(T<=t) one-tail	0,05585738	
t Critical one-tail	1,645	

*Under normal conditions, as it is a long-term and passively managed portfolio, the average return of benchmark should not be negative. But, because of the economic instability in 2007, 2009 and 2014, Government Debt Securities on TL had negative return in these years. Majority of the benchmark's portfolio allocation consists of Government Debt Securities on TL. As an effect of these reasons, the average of benchmark' return was also calculated as negative.

Table 11: t-Test Results for Stock Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Stock Funds	*Stock Funds' Benchmark*
Mean	0,962%	0,852%
Variance	0,517%	0,525%
Observations	120	120
Hypothesized Mean Difference	0	
df	238	
t Stat	0,119	
P(T<=t) one-tail	0,45283742	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Stock Funds	*Stock Funds' Benchmark*
Mean	0,962%	0,852%
Variance	0,517%	0,525%
Observations	120	120
Hypothesized Mean Difference	0	
df	238	
t Stat	0,119	
P(T<=t) one-tail	0,45283742	
t Critical one-tail	1,645	

Table 12: t-Test Results for Flexible Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Flexible Funds	*Flexible Funds' Benchmark*
Mean	0,783%	0,295%
Variance	0,049%	0,127%
Observations	120	120
Hypothesized Mean Difference	0	
df	199	
t Stat	1,275	
P(T<=t) one-tail	0,10195081	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Flexible Funds	*Flexible Funds' Benchmark*
Mean	0,783%	0,295%
Variance	0,049%	0,127%
Observations	120	120
Hypothesized Mean Difference	0	
df	199	
t Stat	1,275	
P(T<=t) one-tail	0,10195081	
t Critical one-tail	1,645	

Table 13: t-Test Results for Index Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Index Funds	*Index Funds' Benchmark*
Mean	1,046%	0,876%
Variance	0,200%	0,559%
Observations	120	120
Hypothesized Mean Difference	0	
df	194	
t Stat	0,214	
P(T<=t) one-tail	0,41553728	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Index Funds	*Index Funds' Benchmark*
Mean	1,046%	0,876%
Variance	0,200%	0,559%
Observations	120	120
Hypothesized Mean Difference	0	
df	194	
t Stat	0,214	
P(T<=t) one-tail	0,41553728	
t Critical one-tail	1,645	

Table 14: t-Test Results for Balanced Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Balanced Funds	*Balanced Funds' Benchmark*
Mean	0,792%	0,225%
Variance	0,072%	0,238%
Observations	120	120
Hypothesized Mean Difference	0	
df	185	
t Stat	1,117	
P(T<=t) one-tail	0,13280053	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Balanced Funds	*Balanced Funds' Benchmark*
Mean	0,792%	0,225%
Variance	0,072%	0,238%
Observations	120	120
Hypothesized Mean Difference	0	
df	185	
t Stat	1,117	
P(T<=t) one-tail	0,13280053	
t Critical one-tail	1,645	

7.2 Equally-weighted benchmark comparison

Table 15: t-Test Results for Standard Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Standard Funds	*Standard Funds' Benchmark*
Mean	0,736%	-0,167%*
Variance	0,012%	0,413%
Observations	120	120
Hypothesized Mean Difference	0	
df	126	
t Stat	1,516	
P(T<=t) one-tail	0,06603691	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Standard Funds	*Standard Funds' Benchmark*
Mean	0,736%	-0,167%*
Variance	0,012%	0,413%
Observations	120	120
Hypothesized Mean Difference	0	
df	126	
t Stat	1,516	
P(T<=t) one-tail	0,06603691	
t Critical one-tail	1,645	

*Under normal conditions, as it is a long-term and passively managed portfolio, the average return of benchmark should not be negative. But, because of the economic instability in 2007, 2009 and 2014, Government Debt Securities on TL had negative return in these years. Majority of the benchmark's portfolio allocation consists of Government Debt Securities on TL. As an effect of these reasons, the average of benchmark' return was also calculated as negative.

Table 16: t-Test Results for Liquid Funds (Government)

t-Test Result for α=0,10 (at %90 Confidence level)

	Liquid Funds (Government)	*Liquid Funds' (Government) Benchmark*
Mean	0,725%	0,538%
Variance	0,001%	0,026%
Observations	120	120
Hypothesized Mean Difference	0	
df	129	
t Stat	1,238	
P(T<=t) one-tail	0,10899528	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Liquid Funds (Government)	*Liquid Funds' (Government) Benchmark*
Mean	0,725%	0,538%
Variance	0,001%	0,026%
Observations	120	120
Hypothesized Mean Difference	0	
df	129	
t Stat	1,238	
P(T<=t) one-tail	0,10899528	
t Critical one-tail	1,645	

Table 17: t-Test Results for Liquid Funds (Flexible)

t-Test Result for α=0,10 (at %90 Confidence level)

	Liquid Funds (Flexible)	*Liquid Funds' (Flexible) Benchmark*
Mean	0,747%	0,627%
Variance	0,001%	0,013%
Observations	120	120
Hypothesized Mean Difference	0	
df	137	
t Stat	1,118	
P(T<=t) one-tail	0,13276367	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Liquid Funds (Flexible)	*Liquid Funds' (Flexible) Benchmark*
Mean	0,747%	0,627%
Variance	0,001%	0,013%
Observations	120	120
Hypothesized Mean Difference	0	
df	137	
t Stat	1,118	
P(T<=t) one-tail	0,13276367	
t Critical one-tail	1,645	

Table 18: t-Test Results for Composite Bonds And Bills Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Composite Bonds And Bills Funds	*Composite Bonds And Bills Funds' Benchmark*
Mean	0,973%	0,140%
Variance	0,055%	0,110%
Observations	120	120
Hypothesized Mean Difference	0	
df	215	
t Stat	2,244	
P(T<=t) one-tail	0,01291176	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Composite Bonds And Bills Funds	*Composite Bonds And Bills Funds' Benchmark*
Mean	0,973%	0,140%
Variance	0,055%	0,110%
Observations	120	120
Hypothesized Mean Difference	0	
df	215	
t Stat	2,244	
P(T<=t) one-tail	0,01291176	
t Critical one-tail	1,645	

Table 19: t-Test Results for Government Bonds and Bills (FX) Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Government Bonds and Bills (FX) Funds	*Government Bonds and Bills (FX) Funds' Benchmark*
Mean	1,011%	0,123%
Variance	0,048%	0,100%
Observations	120	120
Hypothesized Mean Difference	0	
df	212	
t Stat	2,525	
P(T<=t) one-tail	0,00614569	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Government Bonds and Bills (FX) Funds	*Government Bonds and Bills (FX) Funds' Benchmark*
Mean	1,011%	0,123%
Variance	0,048%	0,100%
Observations	120	120
Hypothesized Mean Difference	0	
df	212	
t Stat	2,525	
P(T<=t) one-tail	0,00614569	
t Critical one-tail	1,645	

Table 20: t-Test Results for Government Bonds and Bills Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Government Bonds and Bills Funds	*Government Bonds and Bills Funds' Benchmark*
Mean	0,785%	-0,207%*
Variance	0,016%	0,443%
Observations	120	120
Hypothesized Mean Difference	0	
df	127	
t Stat	1,604	
P(T<=t) one-tail	0,05555278	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Government Bonds and Bills Funds	*Government Bonds and Bills Funds' Benchmark*
Mean	0,785%	-0,207%*
Variance	0,016%	0,443%
Observations	120	120
Hypothesized Mean Difference	0	
df	127	
t Stat	1,604	
P(T<=t) one-tail	0,05555278	
t Critical one-tail	1,645	

*Under normal conditions, as it is a long-term and passively managed portfolio, the average return of benchmark should not be negative. But, because of the economic instability in 2007, 2009 and 2014, Government Debt Securities on TL had negative return in these years. Majority of the benchmark's portfolio allocation consists of Government Debt Securities on TL. As an effect of these reasons, the average of benchmark' return was also calculated as negative.

Table 21: t-Test Results for Stock Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Stock Funds	*Stock Funds' Benchmark*
Mean	0,962%	0,841%
Variance	0,517%	0,541%
Observations	120	120
Hypothesized Mean Difference	0	
df	238	
t Stat	0,129	
P(T<=t) one-tail	0,44885084	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Stock Funds	*Stock Funds' Benchmark*
Mean	0,962%	0,841%
Variance	0,517%	0,541%
Observations	120	120
Hypothesized Mean Difference	0	
df	238	
t Stat	0,129	
P(T<=t) one-tail	0,44885084	
t Critical one-tail	1,645	

Table 22: t-Test Results for Flexible Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Flexible Funds	*Flexible Funds' Benchmark*
Mean	0,783%	0,329%
Variance	0,049%	0,117%
Observations	120	120
Hypothesized Mean Difference	0	
df	204	
t Stat	1,222	
P(T<=t) one-tail	0,11149109	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Flexible Funds	*Flexible Funds' Benchmark*
Mean	0,783%	0,329%
Variance	0,049%	0,117%
Observations	120	120
Hypothesized Mean Difference	0	
df	204	
t Stat	1,222	
P(T<=t) one-tail	0,11149109	
t Critical one-tail	1,645	

Table 23: t-Test Results for Index Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Index Funds	*Index Funds' Benchmark*
Mean	1,046%	0,875%
Variance	0,200%	0,550%
Observations	120	120
Hypothesized Mean Difference	0	
df	195	
t Stat	0,216	
P(T<=t) one-tail	0,41468863	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Index Funds	*Index Funds' Benchmark*
Mean	1,046%	0,875%
Variance	0,200%	0,550%
Observations	120	120
Hypothesized Mean Difference	0	
df	195	
t Stat	0,216	
P(T<=t) one-tail	0,41468863	
t Critical one-tail	1,645	

Table 24: t-Test Results for Balanced Funds

t-Test Result for α=0,10 (at %90 Confidence level)

	Balanced Funds	*Balanced Funds' Benchmark*
Mean	0,792%	0,226%
Variance	0,072%	0,240%
Observations	120	120
Hypothesized Mean Difference	0	
df	184	
t Stat	1,111	
P(T<=t) one-tail	0,13402994	
t Critical one-tail	1,282	

t-Test Result for α=0,05 (at %95 Confidence level)

	Balanced Funds	*Balanced Funds' Benchmark*
Mean	0,792%	0,226%
Variance	0,072%	0,240%
Observations	120	120
Hypothesized Mean Difference	0	
df	184	
t Stat	1,111	
P(T<=t) one-tail	0,13402994	
t Critical one-tail	1,645	