Performance Evaluation of Polish Pension Funds

Dariusz Stanko*

August 2002

JEL classification: G-23
Keywords: performance evaluation, pension funds, pension reform

*Graduate School of Economics, Osaka University, stanko@srv.econ.osaka-u.ac.jp

Performance Evaluation of Polish Pension Funds

Contents

ABSTRACT:

The paper is the first performance evaluation study of the Polish pension funds operating from 1999. The unconditional performance evaluation models are used. It is shown that pension fund managers did produce additional value due to active management. Therefore, unsatisfactory overall results for the public pension system cannot be attributed to the inefficiency of the investment process. The research presents some facts on pension fund's investment behaviour (successful diversification, returns clustering around median manager, positive timing skills).

I. INTRODUCTION

Funded pension schemes are becoming a key point for modern economics and economic policy. Increasing demographic pressure combined with the need of reforming the existing ineffective and politically vulnerable systems bring about the current trends to privatise the modern welfare state and to use capital market-based solutions in the old-age provision.

The urge for reform is particularly strong within the industrialised countries. Poland is one of the first countries in Europe to have adopted a public, fully funded second old-age security pillar. The heart of the system consists of privately managed

yet publicly owned pension funds that are the investment vehicles for retirement savings.

There is growing literature on performance evaluation of pension funds. However, not much research has been done into performance evaluation of mandatory public funds. This paper (up to the best of author's knowledge) is the first performance evaluation study of abnormal returns (alphas) of the Polish pension funds. It presents results for different from UK or US institutional and legal framework. The study represents part of a wider research (Stanko, 2002) into the system. Three issues are of particular concern:

- How did the results of the funds compare to other investment opportunities during 1999 – 2002?
- Did the active management of the pension funds produce any value added?
- What can be said about their investment activity (diversification, asset allocation, asset selection and timing)?

The structure of the paper is as follows. The second section sketches the Polish pension reform and describes current pension market. Next, some previous studies on pension funds are discussed. Section IV presents information on the data used. Section V compares funds' recent results with other investment opportunities.

Composite performance measures are calculated in section VI. Section VII introduces performance evaluation models used in the paper. Section VIII discusses the empirical findings. The last section concludes.

II. Brief Introduction To The Polish Pension System

The Polish new retirement system started in April 1999. There are currently 17 pension funds[i] run by private managing companies. They run individual accounts for each insured and invest their savings in the capital market. The retirement age is set at 60 for women and 65 for men. Before reaching that age, the insured cannot withdraw their retirement assets. However, in case of death of the insured, the assets are subject to inheritance law.

Current assets of the pension funds system (July 2002) comprise of approx. 6.15 b PLN with 10.84 m individual accounts. The market is of an oligopoly type – the first four funds have almost 75% of assets. The law roughly predefines the investment style as investment limits permit to hold maximum of 55% of portfolio in stock.

There exists a wide range of regulations concerning safety measures (reporting, guarantee and reserve funds, fund as a legal entity separated from

managing company), minimal guaranteed rate of return, evaluation framework and so on.

This paper deals only with performance evaluation of pension funds. The overall economic efficiency of the second pillar (costs and performance monitoring issues) is discussed in Stanko (2002).

III. Previous Studies

Up to the author's knowledge, there is no research into Polish pension funds that uses finance theory to evaluate abnormal returns (alphas). The supervisory body, Committee of Insurance and Pension Fund Supervision) issues monthly and quarterly bulletins on the results of the pension funds. However, these include only calculation of internal rate of return, the weighted average rate of return and mandatory minimum rate of return, specified by law.

The pension funds representative body, Chamber of Pensions Fund Administrators released a short commentary (IGFE, 2002) that presented the results of industry in comparison to other investment vehicles.

When it comes to international publications, there are some recent papers that discuss the problem of pension funds performance and measurement. However, the

nature and legislation environment of researched funds are different from each other. The closest to the Polish system are the Latin American schemes. No estimates of alphas are available, though.

Blake and Board (2000, p. 552) study the UK private pension funds and find that the average fund underperformed the market average by 0.45 percentage points per annum, before deducting any fund management fee. Another UK study into over 2000 segregated pension funds by Thomas and Tonks (2001, p. 17) during the period 1983-1997 found that most of the funds are "close-trackers" to the FT-All Share Index and that their average outperformance was significantly different from zero, around one half of a percentage point per year. The average selectivity alpha and the average timing parameters were both negative (Thomas and Tonks, 2001, p. 14). Also, Blake et al. (2001, p. 15) present evidences that the funds' results are very close to the benchmark and on average slightly underperform it. Blake et al. (1999) found a stock selection negative and the average market timing very negative.

There are relatively many studies into the American pension funds. Ippolito and Turner (1987) studied over 1500 US ERISA-based pension funds during the period of 1977-1983, and Lakonishok et al. (1992) examined 769 defined benefit funds in 1983-1989. Both studies conclude that, on average, the pension managers

significantly underperformed the passive management style (represented by S&P index). Lakonishok et al. (1992) relates the average underperformance of 1.3% annually to the agency problems ("window-dressing"). A study of Coggin et al. (1993) on a random sample of 71 US equity pension funds during 1983-1990 found a significant positive selectivity and negative timing. Christopherson et al. (1998), using conditional performance evaluation framework evaluate 261 manager portfolios over 1980-1996 to the Russell 3000 benchmark and find that the average manager outperforms the Russel by 0.72% per annum.

IV. Data Description

A. Source

The sample covered in the analysis ranges from the beginning of June 1999 till the end of June 2002 and is the most comprehensive one. The system started in April 1999, but it was not before June 1999 when the majority of the funds launched their operations in the market.

The closing daily data on pension funds unit prices, stock and bond market indices were obtained from the Internet site of the economic journal "Parkiet" (www.parkiet.com/dane). The Bloomberg GOPL (Polish Government Bonds) index

was given courtesy of Mr. Phil Galdi From Merrill Lynch Bank. Mr. Janusz Zielinski provided the data on Treasury Bills yields for the secondary market from the National Bank of Poland. The data concerning Polish mutual funds investing in bonds were retrieved from financial page www.money.com.pl.

B. Data description

The continuously compounded rates of return were calculated. The paper employs weekly and monthly returns. The assumption was that an investment starts in the morning of the first day of each period (week, month) and ends in the morning of the first day of the following period. Therefore, the closing values of the preceding days were used (i.e. the closing value of the last day was used).

The returns from an investment in the stock market were calculated with the use of WIG and WIG20 indices (Warszawski Indeks Gieldowy, Warsaw Stock Index). The former is a total return index that includes dividends and pre-emptive rights. The latter consists of 20 blue chips and does not account for dividends[ii]. The pension funds are obliged by law to follow investment limits. Therefore, the pension managing houses invest heavily in blue chips equity, which are described by WIG20 index.

For bond returns, the Bloomberg GOPL and MFUND indices were used. The former is a proxy of profitability from investing in Polish government bonds and reflects accrued interest income. The latter index is the returns arithmetic average of the biggest three mutual funds investing in bonds and money instruments. It serves as another benchmark and reflects in particular the pension fund's strategy of preserving the portfolio's accumulated value.

In the case of missing data, an artificial time-weighted data was computed.

C. Distribution and statistical properties of returns and indexes

There are 774 daily observations are used to form weekly and monthly returns. Almost all of monthly returns for the pension funds are normal. In the case of weekly returns, 14 out of 21 funds reveal normality (at 5% significance level). The weekly and monthly returns for stock and bond markets are normally distributed, while the mutual bond market and treasury bills are not.

As the data revealed heteroskedasticity, White (1980) heteroskedasticity —consistent covariance matrix estimators are employed.

The investment results for pension funds are highly correlated with one another, suggesting existence of common return generator factor or factors.

The augmented Dickey-Fuller and Phillips-Peron unit root tests for all fund, stock and bond excess returns, as well as for the averages, rejected the hypothesis of non-stationarity.

V. Performance vs Alternative Investment Vehicles

Table 1 presents the discretely compounded rates of return for the pension funds, stock, bond and treasury bills market for the period 1 June 1999 - 29 March 2002.

Nominal rates of return for the survived pension funds (Panel A) range between 35% and 54% for the whole period, or 10% and 15% annually. The results of the three bank-based funds that appeared on the market later (Panel B) are somehow lower than the outcomes of those that had started earlier. Half of the discontinued funds were outperforming the market average at that time.

Interestingly, the best performing fund in the discontinued funds group was the fund that also started earlier than the others (Pioneer). The second fund (Rodzina), although performing very well, discontinued its operations. The probably explanation lies in fact that this fund started as the last one amongst all 21 pension providers.

It is logical to assume then that performance was not categorically the motivation for the mergers or acquisitions in the market. Rather, some funds did badly during their marketing campaigns. Having failed to achieve an economically sound fragment of the market they were forced to quit. One can hardly claim that the market competition mechanism, exercised by the members voting for the outperforming fund, contributed to these changes in the market.

Thus, as long as the inferior investment performance is not the main cause for funds' disappearance from the market, the problem of survivorship bias[iii] should not have any effect on the performance evaluation.

Table 1 Pension funds. The investment results for the period 1 June 1999 – 28 June 2002.

The geometric real rates of return for the survived funds are quite diverse (Panel C). The best funds earned in real terms around 5-6% while the worst revealed negative or mere 1% growth.

Table 2 Pension funds vs. other investments. The investment results for the period 1 June 1999 – 28 June 2002.

An inspection of table 2 reveals that in the investigated period it was a much wiser strategy to follow the bond markets than to invest in the stock market. The

results of the best pension funds (first quartile) are comparable with the Merrill Lynch Polish Government Bonds index and the three biggest mutual funds investing in bonds and money instruments. However, 12 out of 17 the funds did considerably worse. The pension funds present themselves much better in comparison to mutual funds, especially those investing in shares (equity, balanced and growth funds) and retirement funds. Three quartiles of the funds win with both of these investment vehicles.

Due to constant bear market, the rates of return earned by administrators of the pension funds are not much higher than the one earned by following the most naïve strategy that is saving monies in the bank account. Three of the funds earned even less than the bank account interests.

Although these outcomes definitely cannot be labelled as fabulous, it is important to remember that the funds are supposed to follow the diversification rule and also to operate within the investment limits specified by law. Therefore, they could not invest 100% of their assets into one asset bearing the highest return at the time (for instance in the bank accounts). The asset allocation benchmark that reflects the typical stock-bond combination employed by pension funds would have earned around 30-33% at the same period.

VI. Composite Performance Measures

Table 3 presents the Sharpe and Treynor ratios[iv] for pension funds and other investment vehicles during the period of 1 June 1999 – 28 June 2002. The calculation period is appropriately reduced in cases when a fund or instrument existed for shorter time. The ratios were calculated on the basis of daily continuously compounded returns.

The ratios are mostly negative. The absolute values of the Sharpe indicator are small. Even the best results achieved by mutual funds that invested in bond and money instruments are still much lower (0.07-0.16) than the historical Sharpe ratio for the American stock market that was around 0.5. Again, such effect can be attributed to the bear market dominating during last three years. The stock market investment brought a 9 percent loss. An investment in bonds or Treasury Bills would produce a 51-52 percent gain.

Table 3 Sharpe and Treynor ratios.

Both indicators do not quantify the value added of active portfolio management. However, they help in ranking the efficiency of the funds' investments. The Sharpe ratio uses the standard deviation for calculating its total risk of a portfolio. Thus,

the ratio does not say anything about the diversification issues. Since the pensioners cannot split their retirement assets within several funds at the same time, this ratio is a useful measure. Since they put all their premium in a single well-diversified fund, the total risk is important. The Treynor grade is more suitable for the portfolios that are already diversified since only the non-diversifiable (systematic) part is taken into consideration.

An important function of the institutional investors do is risk diversification. The total risk for a well-diversified portfolio should be identical (in theory) or very similar (in practice) to the systematic risk, then. In the case of pension funds, as far as they do their diversification job well, both of the rankings should be identical. The rankings achieved by Sharpe and Treynor ratios (Table 8) are essentially the same and it implies that their portfolios were well diversified. Only three funds have different ranks.

Table 4 Comparison of ranking persistency.

VII. METHODOLOGY

This section shows the methodology for testing whether the pension funds' active management create some additional (compared to passive investing) value

for the system members. Since the market is very new, the data series are short. Moreover, the author does not possess full information concerning the pension funds' portfolio structure (only the general public information was accessible). Some of the basic indices still do not exist (for example the dividend index). These facts limit severely the usage of possible models of performance evaluation, particularly the conditional evaluation framework.[v]

A. Unconditional market model

A basic Jensen's regression was used. As Cesari and Panetta (2002) remark, the Jensen's alpha can be treated as an unbiased performance measure if the manager of portfolio (fund) possesses security-specific information but no timing information. When the manager achieves successful timing, the measure is usually downward biased (Cesari and Panetta, 2002). The unconditional version of the market model is as follows:

$$\tilde{r}_{it} - r_{ft} = \alpha_i + \beta_{im}(\tilde{r}_{mt} - r_{ft}) + \tilde{\varepsilon}_{it}$$

where: r_{it} is the return of the i^{th} fund at the period t and r_{ft} is the risk-free return at the period t, r_{mt} represents the return of the benchmark market portfolio and β_{im} is the fund's beta, that is its systematic risk. The tildes denote random variables. The

returns in this paper include brokerage, service, depositary and asset management fees. They do not include the up-load payment[vi].

In the context of available data, the stock market indices (WIG and WIG20) were used as proxies for the market portfolio. It is hoped that as the pension market grows a new pension-oriented index will be created. This question is beyond the scope of the paper.

B. Unconditional two-index model

The portfolio structure of pension funds is an important issue in their performance evaluation. As opposed to the heavily researched mutual funds, the pension funds also hold considerable amounts of bonds and other interest-bearing instruments. Recently, they invested around 60-70% of assets in bonds and only 20-30% in stock. According to the modern portfolio theory, bonds should be treated as a part of all risky portfolios. Usually (mainly due to data availability and frequency), it is the stock market index that represents the risky assets. The value of the beta shows the sensitivity of the fund's return to the return of the stock market benchmark. However, in the case of the bond instruments, their correlation with shares is merely 0.19 for daily, and 0.3 for weekly and monthly returns.

The risk measure for bond instruments is probably better associated with the duration term. Since such data is not directly available, the solutions might be as follows. One may try to regress the two-index model and to estimate the parameters. Alternatively, as in Elton et al. (1993), one first regresses the bond returns against the stock exchange index and then uses such orthogonalized index to measure marginal return contribution to the stock index (i.e. the part of the returns that are uncorrelated with the main stock index). Also, Blake et al. (2001, p. 15) use the multiple-index Jensen regression arguing that such approach "is likely to be more appropriate for the aggregate portfolio" (Blake et al., 1998, p. 5).

The following two-index model is considered:

$$\tilde{r}_{it} - r_{ft} = \alpha_i + \beta_{im}(\tilde{r}_{mt} - r_{ft}) + \theta_{im}(\tilde{r}_{bt} - r_{ft}) + \tilde{\varepsilon}_{it}$$

where: r_{it} is the return of the i^{th} fund at the period t and r_{ft} is the risk-free return at the period t, r_{mt} and r_{bt} represents the returns of the benchmark (stock and bond) market portfolios and β_{im}, θ_{im} are the fund's betas for stock and bond investments, respectively. The tildes denote random variables.

In the case when funds investments are mainly concentrated on specific subgroups of securities the market model and the two-index model might not describe properly a fund's investment strategy. Cesari and Panetta (2002) propose

the five-factor model estimated by maximum likelihood method. However, in the context of Poland most of the investment in stocks is concentrated in blue chips and national bonds, therefore two-index model should suffice.

C. Asset allocation benchmark

A synthetic index A_t represents the investment returns from strategic asset allocation portfolio at time t, employed by a pension fund. The asset allocation structure in last three years was approximately 30% of assets invested in stock and 70% in bond.

$$\tilde{r}_{it} - r_{ft} = \alpha_i + \beta_i(\tilde{A}_t - r_{ft}) + \tilde{\varepsilon}_{it}$$

Blake and Timmermann (2002, p. 110) argue that the strategic asset allocation is a risk decision, not the investment decision and is usually determined by maturity structure of the anticipated liability cash flows. However, one may use the strategic asset allocation benchmark to judge the stock selection and market timing (i.e. tactical asset allocation) decisions. A comparison of empirical alphas and alphas derived from asset allocation portfolio gives some insights for portfolio attribution. The paper employed four benchmarks being a combination of stock (WIG, WIG20) and bond (GOPL, MFUNDS) indices.

D. Unconditional timing models

If a fund's performance is based not only on the security-specific information possessed by an investment manager but also on his or her timing strategy, then the alpha estimates from time series undervalue this timing ability (Cesari and Panetta, 2002).

This is so because the composition, and therefore the risk of the portfolio, changes as the managers adjust their exposure to risk in reaction to the market trends. The performance measurement must therefore recognize the manager's micro-forecasting ability (security selection) and macro-forecasting ability (market timing).

The paper considered two models to estimate market timing. The first, based on Treynor-Mazuy test (1966) assumes a non-linear relationship between the risk and return:

$$\tilde{r}_{it} - r_{ft} = \alpha_i + \tilde{\beta}_i(\tilde{r}_{mt} - r_{ft}) + TM_i(\tilde{r}_{mt} - r_{ft})^2 + \tilde{\varepsilon}_{it}$$

The intercept alpha estimate measures the security selection ability, while the squared term represents the additional amount of return as a product of the timing ability. When the TM parameter is positive, this ability is superior, while negative

value shows that the fund is losing the shareholders' money by engaging in speculative activity.

To derive the equation above, one needs to start with the standard Jensen model and then assume that changes of beta are only due to the market timing activity:

$$\tilde{r}_{it} - r_{ft} = \alpha_i + \beta_i(\tilde{r}_{mt} - r_{ft}) + \tilde{\varepsilon}_{it} \quad and \quad \beta_i = \tilde{\beta}_i + TM\,(\tilde{r}_{mt} - r_{ft})$$

Merton (1981), and Henriksson and Merton (1981) define the market-timing ability as the skill of predicting whether the excess return on risky assets will be positive or negative, i.e. whether the return on risky assets will be higher than the risk-free rate at a particular period (Jagannathan and Korajczyk, 1986, p.220). If such a situation is predicted, then the manager weights his or her portfolio more heavily towards the risky instruments and the new beta of portfolio increases. Therefore, the Merton-Henriksson model assumes:

$$\tilde{r}_{it} - r_{ft} = \alpha_i + \beta_i(\tilde{r}_{mt} - r_{ft}) + TM_i \max[0, -(\tilde{r}_{mt} - r_{ft})] + \tilde{\varepsilon}_{it}$$

VIII. EMPIRICAL RESULTS

A. Alphas

Table 5 presents the results of the OLS regression for the unconditional market and two-index models. An arithmetic average of all fund's returns was used to represent the industry's results. Panel A shows the Jensen's alphas for all 14 funds that operated on the market within last three years. The estimates for various models are very similar and range between 3.7 and 4.0 per cent per annum. All alphas are significant (most at 5%, the other at 10% level). The funds' active management has created an additional value comparing to the results that would have been obtained by a passive investment in stock and bond indices.

Table 5 Pension industry performance: Jensen's alpha (monthly returns).

Using the WIG20 (blue chips) and MFUNDS (mutual funds investing bonds and money instruments) gives higher R^2 ratios and better statistical significance of estimates. Both indices are closer to the real investment style applied by pension funds. Additionally, error terms for models employing WIG20 are formally accepted as normally distributed.

Panel B shows the outcomes for a wider sample of survived funds (17). The group incorporates three funds that started their activity later. The number of monthly observations is shorter (34 instead of 37) and that is probably the reason

(along with lower investment efficiency itself) why the estimates are somehow less significant. The estimated alphas are around 2.8 – 3.3 per cents and are lower than that in Panel A. Therefore, it suggests that the investment efficiency of the latecomers was lower.

Panel C demonstrates results for a variable number of funds present at the market during the whole period of 37 months. Their alphas are significant and are a little lower than the first group (Panel A) but higher than the second group (Panel B). The values vary between 3.4 and 3.7% per annum. It implies that the latecomers and the discontinued funds had altogether somehow lower results. Due to short series of observations, the estimates for the survived latecomers and for the discontinued funds alone are not significant and are not presented here.

Table 6 Pension industry performance: Jensen's alpha (weekly returns).

Table 6 contains results for the same models estimated on the basis of weekly returns. Most of the models revealed strong negative autocorrelation; therefore autoregression AR error models were used. For instance, a first-order serial correlation of error term would result in the following model:

$$\tilde{r}_{it} - r_{ft} = \alpha_i + \beta_{im}(\tilde{r}_{mt} - r_{ft}) + \theta_{im}(\tilde{r}_{bt} - r_{ft}) + \tilde{\varepsilon}_{it}$$

where

$$\tilde{\varepsilon}_{it} = \rho\varepsilon_{it-1} + \tilde{v}_{it} \text{ and } \tilde{v}_{it} \sim N(0, \sigma_{it}^2)$$

The weekly estimates of alpha for all the groups are slightly lower for the first (by 0.1-0.3%) and the third group (by 0.1%) than the monthly results. The group of the survived funds that started early (Panel A) reveals alpha between 3.4-3.9% per annum. The third group (Panel C) shows very similar results to the monthly ones: the estimates are around 3.3-3.6%.

The situation is much different for the survived funds that started their operations later. All estimates are considerably lower (0.7%) than the monthly equivalents. However, only one alpha is significant (2.36%).

The results for three groups follow the same pattern as in the case of the monthly models. The first group of the funds has the best alphas, while all the funds that existed are by some 0.3% lower. The worst performance skill results distinguish the late-coming survived funds (Panel B). It might be the case that weekly returns are more prone to distortions by random effects. Moreover, one had to deal with the autocorrelation problem. A comparison of explanatory power between models (R^2) and t-values for the estimates supports the usage of the monthly approach.

Once the general boundaries for the alpha values have been drawn, it is

interesting to analyse how the funds were performing on the individual basis. This information is presented in Table 7 (monthly returns) and Table 8 (weekly returns). Depending on the monthly model, there are around 7 – 9 significant alphas at the 10% significance level. Apparently, the significant estimates dominate in the group with the longest history (i.e. the early started survived funds). Around half funds in this group reveal significant alphas.

The two-index models using the government bonds yields (GOPL) produce 6-7 significant alphas, while the models with the MFUNDS index give 7-8 significant results, depending on the stock market index. The returns from investing in blue chips (WIG20) should be underestimated and the performance overestimated[vii]. However, the empirical results do not vary much. For all the models, WIG20 used in place of WIG produces very similar estimates of alphas. The probable explanation is that high-capitalized WIG20 drives the whole market index and is more volatile then the whole market. The funds invest mainly in blue chips. During the last three years, the WIG20 experienced more negative return than WIG (see Table 2). That is perhaps why that effect offset the dividend upshot.

The weekly approach brings about a similar number of significant individual estimates of alpha. However, weekly observations are more prone to random errors.

The estimates are much more volatile, the models' explanatory power lower and autocorrelation present. Therefore, the author assumes that the monthly results are more consistent and reliable.

Table 7 Individual pension funds performance: Jensen's alpha. Monthly gross returns (up to 34 observations).

Table 8 Individual pension funds performance: Jensen's alpha. Weekly gross returns (up to 134 observations).

Table 9 presents the individual funds alphas and the industry average. Top managers achieved much higher results than the whole group. A direct comparison between funds is not possible because the portfolios have various levels of systematic risk (betas). A higher alpha usually means that the manager took more unsystematic risk by using private selection information to invest more in particular shares. Interestingly, there is a strong relationship between estimated alpha and the position of a fund in the total real return (see also Table 1) and Sharpe ratio rankings (Table 4).

Table 9 Pension funds with significant alphas.

The length of available data does not give us a comfort of precise estimation.

However, in summary, one may assess, that the "added-value" product of the active pension management in Poland was on average 3.4-3.7 % per annum for the industry as a whole and 3.7-4.0 % for the early-starters. The best funds achieved alphas at the level of 5.0-6.0%.

Following the methodology of Blake et al. (2001, p. 10-14), the annualised interquartile range in each of the models was calculated for raw annual returns (taken from Table 1) and alphas (taken from Table 6). The interquartile range shows the difference between top 75% and 25% results. In the case of all survived funds, this range is 215 basis points for raw returns and 177-204 basis points for alphas arriving from six risk-adjustment regression models. However, if just the early starters are taken into account[viii], the variability between return and abnormal return distributions increases considerably. The interquartile range for raw returns is 160 basis points, with range for alpha between 177-209 annualised basis points. The models 4 and 6 (Table 6) have the highest number of significant estimates and produce the widest range of 200-210 basis points. These results suggest, that: 1) the fund industry returns had tendency to cluster around the median value, and 2) the Jensen unconditional models managed to detect a considerable portion of performance variability.

B. Asset allocation

Table 10 presents information about alphas from the asset allocation models for monthly returns. Those alphas represent so-called tracking error, which is the difference between the managed portfolio return and the benchmark portfolio return. The weekly framework produced fewer significant results, and is not presented here.

Table 10 Excess returns from asset allocation and two-index models – comparison.

The asset allocation and two-index managed portfolio results are very similar. For GOPL-based models, the asset allocation alphas are lower than the alphas from two-index models. The situation is opposite when MFUNDS is used. Besides stock and bonds, a pension fund maintains also cash and money-market instruments. The MFUNDS index does account for this part of holdings while the GOPL does not. That is probably why, under the bear market, the alphas from asset allocations for MFUNDS are higher.

The excess returns from asset allocation portfolios represent over 90% of total excess returns.

C. Timing

Timing models with monthly returns bring many significant estimates (Tables 11 and 12). In some cases, to handle the autocorrelation, the autoregressive error term models were used. Apart from one, all the listed funds have positive significant estimates of the timing factor. It supports the observed facts: pension funds switched from stock assets towards bonds in order to defend against the bear market. A comparison of timing estimates sheds some light on the relative importance of timing strategy for each fund in its investment policy. Unfortunately, nothing can be directly inferred about their security selection skills, for with one exception, all the alphas are insignificant. Only the Pioneer fund reveals a negative timing skill and a positive selection capacity (alpha positive at 5% level and equal to 0.1101).

Nevertheless, bringing together the information on timing and asset selection gives us some indirect idea about stock selection. Excess returns for GOPL-based models are lower than total excess return, while the MFUNDS alphas follow the opposite. Knowing that timing was positive for all of the above models induces that security selection capability was negative.

The weekly models produced merely two positive significant estimates with no

significant alphas; the results are not presented here.

Table 11 Treynor-Mazuy timing model. Monthly returns.

Table 12 Henriksson-Merton timing model. Monthly returns.

D. Size effect and performance

Regression of funds' returns on their final magnitudes did not produce any significant relationship. The average real rate of return for all the funds (13.09% p.a.) is higher than the average weighted by market share at the end of the period[ix] (12.81%), which would imply that small funds had better returns in general than the large funds. One of the possible conjectures why the size effect is not caught in more evident form is that the time horizon is too short and the big funds have not yet: (a) achieved the critical magnitudes (organizational problems); (b) dominated the stock markets (impact of trading size on market).

E. Polish results in comparison to other studies

Comparing to the international results, the Polish pension industry recorded high economically and statistically significant abnormal results. There might be several reasons for being so. Firstly, the market during last three years was

definitely bearish; therefore the low-risk strategy (loading off the idiosyncratic stock risk and investing more in government instruments with high real yield) relatively easy bitten up the indexes. Funds also had longer investment horizon than the individuals who do not possess much information and who are more prone to panic during market corrections. This issue is somehow supported by positive timing abilities revealed. The second likely explanation is that the market indices used were not efficient. It is particularly probable since the Polish capital market is definitely still far from being semi-efficient and its size and depth are limited. Finally, there might be other types of information that were not accounted for by the model used here.

IX. CONCLUSIONS

The gross investment results for the pension fund market are satisfactory. The market as a whole and half of the existing sample produced significantly positive results. Funds that started earlier experienced higher excess returns. The industry asset management results could be perhaps even better if some of the system-built problems were limited. The unsatisfactory net return for pension funds' members must be ascribed to the overall regulatory and organizational flaws affecting the

pension framework (Stanko, 2002).

Positive alphas were achieved due to asset allocation and market timing. With respect to the performance attribution, asset allocation played a dominant role. Security selection during that time produced negative results (bear market), however one could question its relative importance. It is because most of the investments were located in blue chips partly in response to the investment law, partly due to herding around mode manager (effect of mandatory minimum rate of return).

The funds applied active trading techniques (positive timing) and accomplished diversification jobs. No relation between the size and performance was observed, probably due to the market's immaturity.

However, the annualised interquartile range in raw returns and in estimated alphas shows the industry's tendency to gather around the median fund manager. These findings are in line with a UK study by Davis et al. (2001) who attribute the effect in to: the predominance of a single investment style, the fee structures and performance evaluation incentives. In the Polish context, these issues are discussed in Stanko (2002).

A study's limitation is that it did not employ the conditional performance

evaluation models due to lack of proper equivalents for the Fama-French (1993) multi-factor models and proxies for information variables. One of the interesting questions is whether the funds follow the contrarian strategy. These issues are intended to be addressed in the future.

REFERENCES

Baks, K., A. Metrick, and J. Wachter (2001), 'Should Investors Avoid All Actively Managed Mutual Funds? A Study in Bayesian Performance Evaluation', *Journal of Finance*, 56(1), 45-85

Blake, D. and A. Timmermann (2001), 'Performance Benchmarks for Institutional Investors; Measuring, Monitoring and Modifying Investment Behaviour', Discussion Paper PI-0106, The Pensions Institute

Blake, D. and J. Board (2000), 'Measuring Value Added in the Pensions Industry', *The Geneva Papers on Risk and Insurance*, 25(4), 539-567

Blake, D., B. Lehmann and A. Timmermann (2001), 'Performance Clustering and Incentives in the UK Pension Fund Industry', Discussion Paper PI-9901, The Pensions Institute

Blake D., B. Lehmann and A. Timmermann (1999), 'Asset Allocation Dynamics and

Pension Fund Performance', *Journal of Business*, 72(4), 429-462

Blake D., B. Lehmann and A. Timmermann (1998), Performance Clustering and Incentives in the UK Pension Fund Industry, Financial Markets Group, Discussion Paper Series, No. 294

Cesari R. and F. Panetta (2000), 'The performance of Italian Equity Funds', *Journal of Banking & Finance*, 26, 99-126

Christopherson, J., W. Ferson and D. Glassman (1998), 'Conditioning Manager Alphas on Economic Information: Another Look at the Persistence of Performance', *Review of Financial Studies*, 11(1), 111-142

Coggin, T. D., F. J. Fabozzi, S. Rahman (1993), 'The Investment Performance of US Equity Pension Fund Managers: An Empirical Investigation', *Journal of Finance*, 48(3), 1039-1055

Daniel, K., M. Grinblatt, S. Titman and R. Wermers (1997), 'Measuring Mutual Fund Performance with Characteristic Based Benchmarks', *Journal of Finance*, 52(3), 1035-1058

Del Guercio, D. and P. Tkac (2000), The Determinants of the Flow of Funds of Managed Portfolios: Mutual Funds versus Pension Funds, Federal Reserve Bank of Atlanta, Working Paper 2000-21

Elton, E.J., Gruber, M. J., Das, S., and M. Hlavka (1993), 'Efficiency with Costly Information: A Reinterpretation of Evidence from Managed Portfolios', *Review of Financial Studies*, 6(1), 1-22

Fama, E., and K. French (1993), 'Common Risk Factors in the Returns on Stocks and Bonds', *Journal of Financial Economics*, 33(1), 3-56

Farnsworth, H., W. Ferson, D. Jackson and S. Todd (2002), 'Performance Evaluation with Stochastic Discount Factors', NBER Working Paper 8791

Ferson, W. and K. Khang (2002), 'Conditional Performance Measurement Using Portfolio Weights: Evidence for Pension Funds', *Journal of Financial Economics*, 65(2), (in press)

Henriksson, R. and R. Merton (1981), 'On Market Timing and Investment Performance. Part II. Statistical Procedures for Evaluating Forecasting Skills', *Journal of Business*, 54(4), 513-33

IGFE (2002), Chamber of Pensions Fund Administrators, Bulletin No. 3

Ippolito, R. A. and J.A. Turner (1987), 'Turnover, Fees and Pension Plan Performance', *Financial Analyst Journal*, 43, 16-26

Jagannathan, R. and R. Korajczyk (1986), 'Assessing the Market Timing Performance of Managed Portfolios', *Journal of Business*, 59(2), 217-235

Lakonishok, J.A., A. Shleifer and R.W. Vishny (1992), 'The Structure and Performance of The Money Management Industry', *Brooking Papers on Economic Activity*, Macroeconomics, 339-391

Merton, R. (1981), 'On Market Timing and Investment Performance. Part I: An Equilibrium Theory of Value for Market Forecasts', *Journal of Business*, 54(3), 363-406

Pastor, L. and R. F. Stambaugh (2002), 'Mutual Fund Performance and Seemingly Unrelated Assets', *Journal of Financial Economics*, 63(3), 315-349

Sharpe, W. (1992), 'Asset Allocation: Management Style and Performance Measurement', *Journal of Portfolio Management*, 18, 7-19

Stanko, D. (2002), 'Polish Pension Funds – Does The System Work? Cost, Efficiency and Performance Measurement Issues', mimeo, Osaka University, Graduate School of Economics

Thomas, A. and I. Tonks (2000), Equity Performance of Segregated Pension Funds in the UK, Centre for Market and Public Organisation, University of Bristol,

Treynor, J. and K. Mazuy (1966), 'Can Mutual Funds Outguess the Market?', *Harvard Business Review*, 43, 63-75

White, H. (1980), 'A Heteroskedasticity-Consistent Covariance Matrix Estimator

and a Direct Test for Heteroskedasticity', *Econometrica*, 48, 817-838

FOOTNOTES

[i] There will be 16 funds from 2003 as an effect of recent merger announcement.

[ii] Income from dividends for the industry in 2001 amounted to nearly 64 m PLN that is to 5.4% of whole revenues from investments (KNUiFE, Quarterly Bulletin 4/2001).

[iii] Survivor effect is a problem when the evaluated performance is (usually) biased upside "due to exclusive focus on those institutions that survived throughout the evaluation period." (Davis and Steil, 2001). For this topic see for example Goetzmann et al. (1992), Elton et al. (1996), Garcia and Gould (1993).

[iv] The Sharpe ratio is calculated as:

$$\frac{r_i - r_f}{\sigma_i^2}$$

where r_i is the return earned by a fund during a period, r_f is the return earned on a risk-free investment (usually Treasury Bills) and σ_i^2 is the volatility of the fund's returns during the period. The Treynor Ratio is calculated as:

$$\frac{r_i - r_f}{\beta_i}$$

$$\text{where } \beta_i = \frac{\text{cov}(r_i, r_m)}{\sigma_m^2}$$

is the measure of systematic risk, or in other words, the relation between the covariance of fund and market portfolio returns and the volatility of the market portfolio.

[v] For conditional models based on returns see Ferson and Schadt (1996); models employing portfolio holdings are presented in Ferson and Khang, (2002). Daniel et al. (1997) introduce characteristic-based benchmarks. Also, there exist numerous other methodologies employing: style analysis (Sharpe, 1992) ordered mean difference (Bowden, 2000), seemingly

unrelated assets (Pastor and Stambaugh, 2002), stochastic discount factor (Farnsworth et. al, 2002), Bayesian performance (Baks et al., 2001) and flows (Guercio and Tkac, 2000).

[vi] According to the Polish law, the up-load fees are deducted before contributions are calculated into investment units. Investment costs are calculated every working day. The fee for management (equal to monthly 0.05% of accumulated assets) is calculated alike; however, the deduction takes place on the month's last working day. Such an operation affects directly the value of the investment unit in case of weekly returns (the last month's week return is lower by 0.05% in comparison to other weeks). However, monthly returns are not affected, as they are calculated on the basis of last days of the succeeding months (the operation consequences cancel out).

[vii] Alpha is a result not explained by the model. Investing in market portfolio should produce capital gains and dividend income. The later part is not accounted for by the index. Therefore the dividends artificially increase the value of alpha.

[viii] None of the models gave significant alpha estimates for the survived latecomers. Thus, the estimation of interquartile range is higher if one skips those three funds.

[ix] One should weight the returns by the value of the fund at the end of each period (ex. quarter). However, since there was no considerable change in the market structure within this period for the survived pensions, even such a rough calculation should work well.