Netspar

Network for Studies on Pensions, Aging and Retirement

Netspar DISCUSSION PAPERS

Xiaohong Huang and Ronald Mahieu

Performance Persistence of Dutch Pension Funds

Discussion Paper 09/2010-038

Performance Persistence of Dutch Pension Funds

Dr. Xiaohong Huang
School of Management and Governance, University of Twente,
Drienerlolaan 5, 7522 NB Enschede, the Netherlands.
E-mail: x.huang@utwente.nl

Dr. Ronald Mahieu
Department of Econometrics and Operations Research, Tilburg University,
Warandelaan 2, 5000 LE Tilburg, the Netherlands.
Tel: 0031 53 4893458, Fax: 31 53 489 2159.
E-mail: r.j.mahieu@uvt.nl *

September 15, 2010

Summary

This paper studies the investment performance of pension funds with a focus on their ability in implementing their intended investment strategy. We use a sample of Dutch industry-wide pension funds, which are obliged by law to report their investment performance according to the so-called *z-score*. The *z-score* is a risk-adjusted performance measure where the benchmarks are chosen a priori preventing manipulation in the score calculation. We find that pension funds as a group cannot beat their self-selected benchmarks and show no performance persistence. It reflects that pension funds on average do not add value in implementing their investment strategy. Cross-sectionally, it turns out that large funds are better able to beat their benchmarks persistently than small funds.

Keywords: Pension fund investment, performance evaluation, *z-score*
JEL classification: G23, C12

*The authors thank Jenke ter Horst, Carmen Pilar Marti, and other participants in the RSM PHD seminar, NETSPAR Pension Day and EFMA 2008 conference for valuable comments. The paper was written when both authors stayed at Erasmus University Rotterdam, the Netherlands. We gratefully acknowledge IQ Info (`www.iqinfo.com`) for providing us access to their pension fund database. The usual disclaimer applies.

1 Introduction

The aggregated market value of Dutch pension fund portfolios is enormous. At the end of year 2006 the total asset size of Dutch pension funds was around €691 billion, while the assets from other sources, managed by collective investment schemes such as mutual funds and hedge funds, are only about €117 billion.[1] Most of these pension assets are associated to the so-called "mandatory industry-wide pension funds" (€470 billion), which manage the pension savings for the majority of Dutch employees. The sheer size of this category of pension funds and their significant role in providing retirement income warrant a careful investigation of the performance of their investment portfolios.

In the Netherlands a mandatory industry-wide pension fund is a multi-sponsor pension plan that provides defined benefit pension services to all employees of the companies affiliated to a particular industry. Employees of these companies are obliged to participate in these schemes. The mandatory feature of these plans leads to a legal requirement that pension funds should report their investment performance in terms of a so-called *z-score*, a risk-adjusted measure of their investment returns. If a fund fails a performance test based on this *z-score*, it loses its mandatory status. Individual participating companies can then leave the fund and either join another pension fund or establish their own fund.

In this paper, we use a unique data set of *z-score* observations to provide a cross-sectional and longitudinal description of the investment performance of Dutch mandatory industry-wide pension funds. Our study adds to the current literature on pension fund performance. It provides another piece of evidence that pension funds do not add value in implementing investment strategies with respect to the indicated benchmarks. Our study also shows the variation in performance across funds of different sizes, revealing that big funds persistently outperform small funds. This finding corroborates the ongoing consolidation in the pension fund sector.

A pension fund portfolio consists of various asset classes, and the study of its investment performance can be performed at the level of both an individual asset class portfolio and the level of the overall fund portfolio. Previous empirical studies focus on the asset class level

[1] According to statistics on the website of the Dutch central bank (DNB): `www.dnb.nl`.

to investigate the ability of individual asset managers. This type of performance evaluation can aid pension fund trustees in their decisions of hiring and firing asset managers. From the perspective of pension fund participants, however, the overall investment performance of the fund is more important than the performance of an individual asset class. For example, it is the overall performance that will influence the contributions they have to pay and the benefits they receive after retirement. Our study addresses the question whether pension funds outperform their benchmarks. The lack of empirical studies pension fund investment performance is related to the fact that there is little detailed information available on the asset allocations and the returns of individual components of the investment portfolio. The Dutch sample used in our study overcomes this problem, and provides a risk-adjusted measurement that accounts for fund- and period- specific asset allocations and performance.

The total investment return of a pension fund portfolio is in general determined by the strategic asset allocation and the implementation of this allocation. The strategic allocation is typically set by the trustees with the help of consultants and investment advisers. The implementation of the strategic portfolio is delegated to internal or external asset managers with different specializations.[2] The quality of the implementation is determined by the delegation and the monitoring of the investment. Our paper focuses on the second role of trustees in their delegation and monitoring tasks. The success of the investment implementation is measured by the fund's overall portfolio performance in excess of an a priori agreed-upon benchmark portfolio.

Compared with retail investors, pension funds are more resourceful in carrying out an investment strategy. They can receive extensive help from advisors and consultants, gain valuable information before making the decisions, and can establish desired procedures to monitor the investment process. We would expect that pension fund trustees are able to select and recruit a superior group of internal and/or external asset managers and establish effective investment management procedures to encourage their asset managers to beat the pre-agreed benchmarks. A typical investment mandate awarded by a pension fund often

[2]A new trend is that an external investment firm acts as a fiduciary asset manager, who structures and monitors the total investment process from strategic asset allocation to individual asset manager selection.

has a contract life of two or more years.[3] The inconvenience of moving a large amount of pension assets across different asset managers or asset categories may also predict some type of performance persistence. However, our study finds that over our sample period pension funds do not outperform their benchmarks on average and show no performance persistence. We do find that large funds are able to persistently outperform their smaller peers. This indicates that pension funds may revalue their efforts in active asset management.

This paper is organized as follows, In Section 2 we describe some previous studies on pension fund investment performance. Section 3 provides some background on the investment processes at Dutch pension funds. Section 4 introduces the *z-score* and Section 5 the associated data. The results are discussed in Section 6. Section 7 concludes.

2 Literature review

In this part we make an overview of the major papers on pension fund performance evaluation. Both Lakonishok, Shleifer & Vishny (1992) and Coggin, Fabozzi & Rahman (1993) look at the equity portfolios of US pension funds. The first paper takes a sample, as large as it could be, to examine the double-agency structure and its relation with underperformance. The second paper makes a random selection of equity pension funds to investigate the ability of fund managers. Busse, Goyal & Wahal (2006) extend to fixed income portfolio and use a larger and recent dataset to estimate the abnormal returns. Tonks (2005) studies the UK samples and investigates the ability of investment houses rather than individual fund managers in managing pension assets. Bauer, Frehen, Lum & Otten (2007) study the aggregate equity portfolio at the fund level, and focus on the comparison with the equity portfolio in mutual funds. Brinson, Hood & Beebower (1986) and Ippolito & Turner (1987) pioneered the performance evaluation at the overall fund level, and they both use a benchmark portfolio including at most three asset classes, which is no longer proper for the current complex pension portfolio. Blake, Lehmann & Timmermann (1999) study the UK pension funds and focus on the return attribution to strategic asset allocation and managerial skills. The conclusions on

[3] Although in principal pension fund trustees can fire the asset managers within a short notice such as one day.

outperformance from pervious studies are mixed and hard to compare, because they focus on different sampling period and sampling region, use different benchmarks, and differ in fee consideration and unit of observation such as asset managers or investment houses.

Lakonishok et al. (1992) investigate the performance of 769 US equity pension fund portfolios for the period 1983Q1 through 1989Q4. Using quarterly gross return (before management fees) they find on average the equity portfolios under perform $S\&P$ 500, and conclude that active management does not add value. They propose the reason for the underperformance is the cost of double-agency structure existing in the pension fund industry where the investment is delegated from corporate management to corporate treasurer and then to portfolio managers. As persistence is concerned, the equity portfolio as a whole does not show any performance persistence for a one-year horizon. This result holds for the growth and the yield portfolio, but the value portfolio shows some persistence.

Coggin et al. (1993) shed light directly on the performance of individual asset managers for pension funds. They study the investment performance of US equity pension fund portfolios by randomly sampling 71 managers for the period January 1983 through December 1990. They distinguish monthly returns attributed to managers' stock selection and timing ability. By using both a general equity market index and a style-specific index, they find the equity asset managers for pension funds on average have positive selection ability but negative timing ability. They do not consider fees.

In addition to the equity portfolios, Busse et al. (2006) extends the analysis to other asset classes including 1,683 fixed income portfolios and 1,196 international equity portfolios held by the US pension funds over the period of 1991-2004. Using quarterly returns, they find a positive abnormal return after adjusting for the relevant risk factors for all three asset classes, even after controlling for costs. As persistence is concerned, only winner portfolios show performance persistence for a one-year horizon.

Rather than looking at the individual asset managers and their portfolios, Tonks (2005) turns his attention to fund management houses for the UK pension funds. His data are quarterly returns on the equity portfolios of 2,175 U.K. pension funds managed by 191 fund management houses from 1983Q1 to 1997Q4. Using a variety of risk adjustment, such as single factor Capital Asset Pricing Model, the Fama-French three-factor model, and a Carhart four-factor model, and averaging the abnormal returns of all portfolios under one management house, he finds a positive abnormal return and persistent performance of fund management houses over a one-year horizon. His results do not take costs into account.

In addition to the performance evaluation at the portfolio level and fund management house level, Bauer et al. (2007) study the performance of the aggregate equity portfolio for the US pension funds (including 716 Defined Benefit and 238 Defined Contribution plans) between 1992 and 2004 at the fund level. Deducting the gross annual returns by the returns to fund-specific benchmark and the costs, they find close-to-benchmark performance. This result also holds when the net returns are adjusted by the Fama-French risk factors. They argue that this is because pension funds are less exposed to agency costs than mutual funds for their monitoring capacity and negotiation power. They detect no persistence of the equity portfolios at the fund level for one year horizon.

Studies on the overall pension fund portfolio performance are even scarcer. Among the earliest is Brinson et al. (1986). They study 91 large US pension funds over a sample period of 1974-1983. Attributing returns to investment policy returns and active returns due to market timing and security selection, they document negative active returns, meaning under-performance with respect to the benchmark portfolio which includes only coarse categories of stocks, bonds and cash. Our paper in essence is similar to this paper in that our data separate active returns from policy returns, and the policy returns are represented by the benchmark return which is more accurate with a finer asset classification to reflect the current complex investment practice.

Rather than using a fund-specific benchmark portfolio, Ippolito & Turner (1987) use the same benchmark indices such as $S\&P500$ for all pension funds. Based on the net-of-fee returns of 1,526 US pension funds they find that the investment performance is very sensitive to the choice of benchmark whether it is a stock index or a bond-stock mix index. They also find that larger pension funds outperformed smaller funds substantially. Their results are not very informative for the current practice. Their choice of a broad stock or a stock-bond mix index as a benchmark is not appropriate in the current performance evaluation of pension funds, as the current pension fund investment is no longer limited to general stocks or bonds as in the paper's sample period back to 1977-1983, but is much more complicated in multiple asset classes and styles.

A more recent paper on the pension fund performance is on the UK sample by Blake et al. (1999). They investigate the monthly return of the overall pension portfolios of over 300 UK pension funds for the period of 1986-1994. Besides documenting underperformance with respect to the benchmark portfolio, their paper has a focus on decomposing the returns into a component attributed to strategic asset allocation and a component attributed to management skills. They adjust the raw returns with a benchmark portfolio where the weight to each asset class is either the time-average of realized allocation or an estimated weight by associating it to time. In addition, they apply the same external indices for all pension funds regardless of the difference of investment style within a certain asset class.

Concluding the previous papers we find there are two ways to construct a benchmark portfolio in computing abnormal returns. One way is using risk factors (such as a equity market index, a bond index or even detailed size factor, distress factor, etc.) for each asset class and estimate the loading on these factors, such as Coggin et al. (1993), Tonks (2005), and Busse et al. (2006). The benchmark portfolio is then composed by the loadings and the returns to the risk factors. This approach can be used when there is a large dataset and a proper estimation can be made. A second way is to use the holding information or asset allocation of a portfolio, such as Brinson et al. (1986), Ippolito & Turner (1987), and Blake et al. (1999). The benchmark portfolio is composed of the weighting and the index returns to respective asset classes. This approach is more accurate in accounting for the risks but requires holding information.

Our study distinguishes itself from most of the previous studies in three major aspects. Firstly, we focus on the multi-asset pension fund portfolio rather than a delegated portfolio in some asset class. This overall portfolio serves one pension plan, so its performance is highly relevant to plan sponsors and beneficiaries. Secondly, the data we use in this study provides the returns after accounting for a holding-based benchmark. Such a benchmark is more appropriate for a multi-asset portfolio than a benchmark based on some stylized risk factors. Thirdly, in the choice of a benchmark portfolio, our data improves in two aspects compared to that used in Blake et al. (1999). Firstly the precise fund-specific and period-specific weights for each asset class are used in the benchmark portfolio. Secondly the indices for the benchmark portfolio are also fund-specific. For example, some fund may use $S\&P500$ for an equity allocation, while some other fund may use a MSCI index. Therefore we have the excess return that is appropriately adjusted by the fund-specific benchmark portfolio, and it represents the value added by any deviation from the benchmark portfolio, due to either stock selection or marketing timing strategy of the selected group of asset managers.

3 Investment process and performance

Before we address the performance, we provide a brief description of the investment process of a defined benefit Dutch pension fund. Although this description is motivated by the Dutch case, it is not very different from general procedures. In brief, fund trustees typically determine the investment policy with the help of consultants. The execution of this policy is delegated to internal and/or external asset managers selected by trustees.

The investment policy is often motivated by an Asset Liability Management (ALM) study, which is an integral risk management study of the fund, taking into account the short-term and long-term objectives of the fund. The investment policy is represented by a strategic asset allocation. This is a portfolio based on the fund's (subjective) view of expected returns and risks of each asset class and an estimation of the fund's liabilities from a long-term perspective. The strategic asset allocation is often reviewed every 3 or 5 years to reflect major changes in the underlying assumptions regarding the assets and the liabilities. From the strategic asset allocation trustees define an investment plan that can be implemented by asset managers, often on an annual basis. This plan reflects a shorter term view on the risk-return profile of each asset class and sometimes indicates the ambition to exploit tactical allocation skills. The investment plan typically consists of weights to individual asset classes that may differ from the strategic weights set in the ALM study. According to the annual investment plan, trustees assign mandates for each asset class to a selected group of asset managers.[4] These managers

[4]There is a trend that trustees delegate this manager selection job to a fiduciary asset manager in order to

can be either in-house or external, one or multiple, of passive or active style. The reason for delegation is motivated by the perceived expertise of an asset manager in a particular asset class. Other reasons can be economies-of-scale in trading and record keeping (Sharpe (1981)). The performance of asset managers is monitored at a regular frequency.

The investment process reveals that investment returns are generated from three sources. One is from the strategic asset allocation, computed from the allocation and the related index return. The second source is from executing the annual investment plan, computed from the weights set in the investment plan and the index return. This part of the returns measures the added value from over- or under-weighing the strategic allocation, reflecting the timing skills of trustees when they draw up the investment plan. The last source comes from the actual execution of the investment plan delegated to a group of asset managers, computed from the actual weights and realized returns. It reflects the timing and selection skills of the asset managers as a group. Good (1984) and Brinson et al. (1986) have a similar return attribution except that they further distinguish the returns from timing and selection skills. This paper focuses on the returns from the third source, which reflect the quality of implementing the annual investment plan.

To obtain the difference between the actual returns and the returns attainable from strict adherence to an annual investment plan, a benchmark portfolio needs to be defined that represents the annual investment plan. This benchmark portfolio is a hypothetical portfolio, which is "structurally identical to the investment strategy without whatever active management takes place" as defined in Logue & Rader (1998) (p168) or a "passive mix with the same style" as in Sharpe (1992). Our performance measure defines such a benchmark portfolio. An example can be found in Table 1. The benchmark portfolio has a twofold purpose. First, the index for each component portfolio is used by trustees to evaluate the performance of individual asset managers for a particular asset class. Second, the overall return from the benchmark portfolio serves as a return target. In our study we use the benchmark portfolio in the context of the second purpose to evaluate the quality of investment implementations by asset managers.

avoid contacts with too many asset managers.

Table 1: An example of a benchmark portfolio
This is a reproduction of a benchmark portfolio. It specifies the weighting and the indices used for different investment styles. The range specifies the bound within which an active asset manager must control the allocation. Source: 2006 annual report of the Agriculture and Food Supply Pension Fund, which can be found via `www.iqinfo.com`.

Assets	Weight	Range	Index
Fixed income	**75%**	**65%-85%**	
Governments	70%	60%-80%	Citigroup Gov Bond Index
Corporates	15%	10%-20%	Citigroup non-EGBI EMU index
Private Loans	15%	10%-20%	Customized Private Loan Index
Equity	**15%**	**5%-25%**	
Europe	40%	30%-50%	MSCI Europe
USA	20%	10%-30%	MSCI North America
Pacific	15%	5%-25%	MSCI Pacific
EM Global	25%	15%-35%	MSCI EM Global
Real estate	**5,0%**	**0%-10%**	
Residential	50%	25%-75%	ROZ- IPD Woningen
Shops	50%	25%-75%	ROZ- IPD Winkels
Alternatives	**5,0%**	**0%-10%**	
Commodities	50%	0%-100%	DJ-AIG Commoditie Index
Hedge Fund	50%	0%-100%	Euro 7-day Libid

4 The performance measure: *z-score*

This section introduces the performance measure we use in this paper. Since 1998 every Dutch mandatory industry-wide pension funds must compute a so-called *z-score* to reflect their investment performance. The *z-score* is the difference between the actual return and the return on a predefined benchmark portfolio, net of expenses, and normalized by the riskiness of the portfolio, as in the following equation:

$$z_{i,t} = \frac{(R_{p,i,t} - c_{p,i,t}) - (R_{b,i,t} - c_{b,i,t})}{E_{i,t}}$$

where $R_{p,i,t}$ and $c_{p,i,t}$ are the gross investment return and internal investment cost of pension fund i at time t respectively. The internal investment cost also includes the fees paid to the external asset managers and investment related custodian and administrative cost. $R_{b,i,t}$ is the fund i's benchmark portfolio return using market indices in the respective asset categories at time t. See Table 1 for an example. $c_{b,i,t}$ is the associated investment cost of the benchmark portfolio which depends on the percentage of equity investment in the portfolio.[5] The benchmark portfolio is determined by trustees at the beginning of each year and fixed for one year. Specifically, the weights and the index for various asset classes in the benchmark portfolio are defined a priori. In addition the index should represent the asset class, be investable and objectively measurable.[6] Thus the benchmark return represents the return that an individual investor can obtain if he invests in the benchmark portfolio, and the difference between the realized return and the benchmark return reflects the excess return that a pension fund can earn by selecting the right internal or external asset managers. The pre-selected benchmark portfolio also excludes the possibility of manipulation in calculating the *z-score*.

To enable the comparison across pension funds with different investment strategies, the excess returns are scaled by the riskiness of the asset mix in the benchmark portfolio ($E_{i,t}$). The asset mix for this purpose contains two major categories: equity and fixed income (including cash). The risk percentages assigned to the equity and fixed income investment are fixed by law at 2.6% and 0.6% respectively.[7] For example, if a fund has an asset mix of 60% equity and 40% fixed income, then $E_{i,t} = 0.6 * 2.6\% + 0.4 * 0.6\% = 1.8\%$. The reported *z-score* is audited by external accountants.

The way the *z-score* is constructed reveals that it is not a measure to evaluate the effectiveness of the investment plan, but a measure of the quality of implementing the investment plan. The benchmark used in the calculation reflects the investment plan for a particular fund and a particular period. Therefore the *z-score* accurately shows the fund's ability in beating their own benchmarks. A positive (negative) *z-score* means that the fund has successfully implemented (failed to implement) its investment plan. This success (failure) is attributed to a fund's skill in selecting and monitoring its asset managers. A high (low) *z-score* reflects the relatively good (poor) ability of the fund in executing its investment plan.

The creation of the *z-score* is to have a standardized normal distribution of the returns of all funds so that the regulator can judge whether a fund's investment performance falls

[5]This cost is presented in Bpf (2000), and range from 0.10% to 0.22%. It varies to the equity proportion of the pension portfolio.

[6]See Article 5.3 in Bpf (2000).

[7]According to Bpf (2000), the riskiness of equity and fixed income investment is set at 2.6% and 0.6%, which is said to be the excess returns of the respective asset category over the past few years at the time of law enaction.

to the lower 10% of the distribution. If a fund falls to the 10% percentile, it is regarded as not delivering a satisfactory investment performance, then the participating companies are no longer obliged to join this fund, and they have the option to leave the fund to join another pension fund or establish their own. [8]

We use the *z-score* to examine the quality of investment implementation by pension funds over time and cross sectionally. There are some concerns for this measurement and we show that they do not pose a serious obstacle to its use for our purpose.

A first concern is that the benchmark portfolio is a static benchmark, in which the weighting of different asset style are fixed for one year. As a result, the intertemporal changes in the investment plan during the year cannot be captured by the benchmark portfolio used in the *z-score* calculation, but do change the return of the actual benchmark portfolio. This can hamper a fair evaluation of the implementation quality, because part of the deviations is due to the change of the benchmark portfolio and has nothing to do with the implementation ability of the selected asset managers. We believe the concern over a static benchmark portfolio is more of a conceptual problem rather than a practical one due to the following practices. Firstly, fixed weighting is a general rule, but the benchmark portfolio is allowed to be changed once when there is a considerable change in the liability structure or the old investment plan is obviously no longer appropriate for the fund.[9] Secondly, changing the investment plan during the year is more of a practice per Jan 1, 2007 when the regulation on financial assessment is implemented, which requires the investment plan to match the market value of liabilities. Thus during our sample period we do not expect material changes in the investment plan during the year. Therefore the concern of a static benchmark portfolio lacks a practical relevance in our performance evaluation.

A second concern is the risk adjustment in the denominator of the *z-score*, where the riskiness only considers the asset categories of equity and fixed income investments and the

[8] A statistical test, called *performance test*, is used to support this decision. Base on the central limit theorem, the test statistic is calculated as $P_{5\text{ year}} = (\sum_{t=1}^{5} Z_{i,t})/\sqrt{5}$, following a normal distribution. The critical value of the test is -1.28, which corresponds to a confidence level of 90% for a standardized normal distribution. If the test statistic is less than -1.28, a sponsor can choose to fire the trustees by opting out of this industry-wide pension fund.

[9] See Article 5.4 in Bpf(2000). As of November 1, 2007 funds are allowed to adjust their norm portfolio twice a year. In addition, some funds, such as Pension fund Vervoer, use a floating benchmark moving with portfolio development.

riskiness are fixed at a certain value. The ignorance of the real estate category might be a potential issue for the risk adjustment. However, the *z-score* conducts the risk adjustment in two levels. In addition to the risk adjustment in the denominator, the benchmark portfolio in the numerator adjusts the risk in real estates by using a very fine classification as shown in the example in Table 1. Moreover, the real estate investment is only a minority class (around 12%) of the total investment. Hence we believe the ignorance of the real estate class in the denominator does not distort much the distribution of the current *z-scores* among pension funds. Another concern is the moral hazard issue arising from the fixed values of 2.6 and 0.6 as the riskiness of the equity and fixed income investments, where pension funds can take advantage in calculating their scores. If a fund takes more risks than what is assumed in the benchmark, its *z-score* can be inflated. However, this only happens in the up market. The opposite happens in the down market. As our sample period covers both up and down market conditions, a higher risk taking won't give an advantage in the *z-score* calculation, unless the fund can time the market by taking a higher risk in the up market, and a lower risk in the down market. We also examine the correlation between the *z-score* and the asset mix in our sample and find it is very low and not statistically significant. This further suggests that the *z-score* has fully adjusted the risk brought by the asset mix.

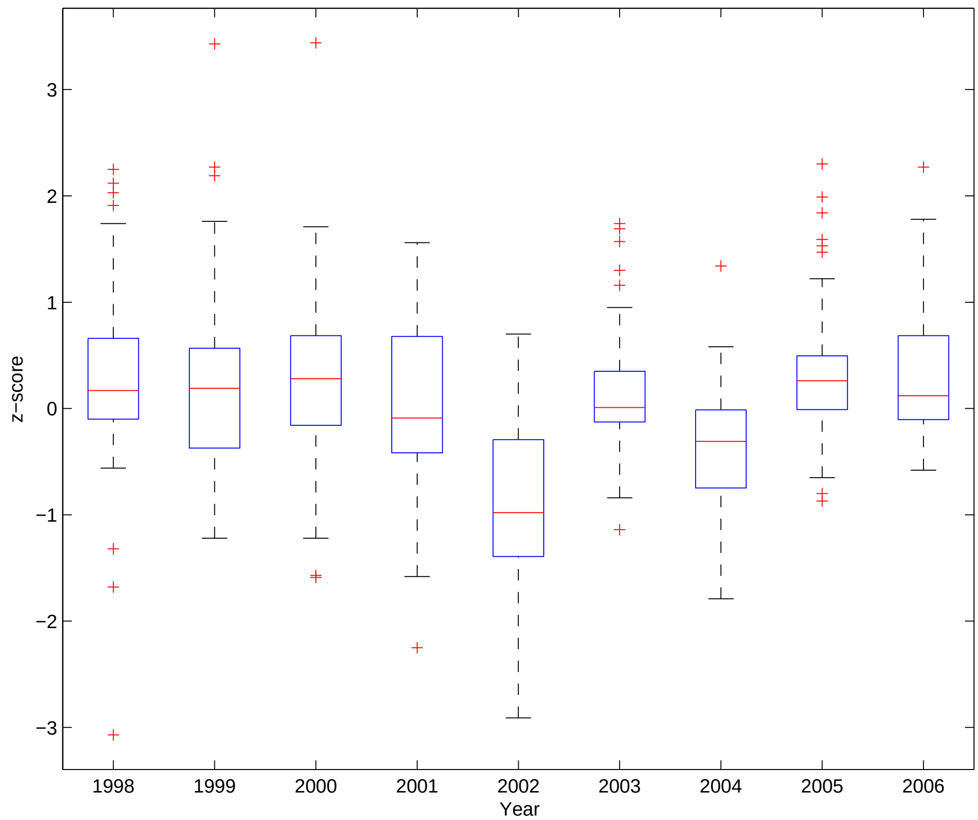
5 Data

We use the publications of the Dutch industry-wide pension fund association.[10] In addition we obtain data from *pensioninfo* which collects and composes aggregate financial information of companies and organizations including pension funds.[11] We merged and verified data from both sources. When there is a discrepancy between the *z-scores* from the two sources, we used the *z-score* reported in a fund's annual report.

Our sample runs from 1998 through 2006 and covers the entire population of mandatory industry-wide pension funds.[12] Over this sample period, the number of funds varies between

[10] In Dutch it is called the *Vereniging van Bedrijfstakpensioenfondsen* (VB). See their website at `www.vb.nl`.
[11] See their website at `http://www.iqinfo.com`.
[12] According to DNB 2007 statistics there are 71 mandatory industry-wide pension funds including 7 pre-pension funds which provide pensions for early retirement. Only mandatory funds are required to report

Figure 1: Pension fund *z-scores*: 1998 - 2006
This figure reports the box plots of the *z-scores* of all industry-wide pension funds for each of the years in our sample period of 1998-2006. The boxes around the median line represent the interquartile range. The dotted lines extend to the most extreme data values within 1.5 times the interquartile range. '+'s denote further outlying observations.



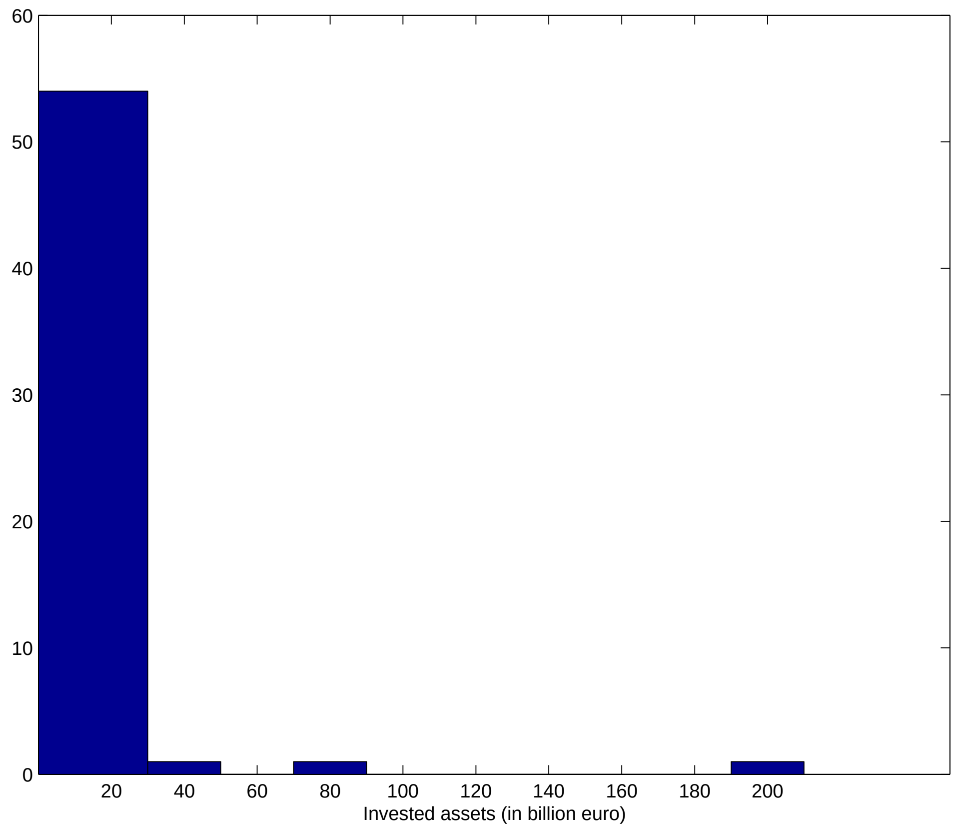
59-65 for a number of reasons. Some funds either started to exist or became mandatory after 1998, one fund merged, and two funds bought the insurance with guaranteed returns. In the end, we have a sample of 57 funds that contain a complete data set on *z-scores*, and this sample will be used for the persistence test. No funds become non-compulsory or ceases during our sample period and thus our sample does not suffer from survivorship bias. A graphic overview of the *z-scores* can be found in Figure 1.

Since there is no considerable change in the relative sizes of the pension funds in our sample, we use the value of invested assets in 2006 as a proxy for the fund size. The data is

z-scores.

Figure 2: Size histogram of 57 pension funds in 2006

This figure draws the histogram of 57 pension funds based on their invested assets in 2006.



obtained from all pension funds' 2006 annual reports and shown in Figure 2. The smallest fund in the sample is €1.47 million, the largest is €208.9 billion, and the median and mean are respectively €426 million and €7.2 billion. This reflects a large size spread among Dutch pension funds. Most funds are small- and medium-sized within €10 billion except for four multi-billion funds.

6 Empirical results

The *z-score* is based on the fund-specific benchmark portfolio and reflects a fund's ability in beating its own benchmark. Descriptive statistics in Table 2 show that through the sample period the average *z-score* varies around 0 suggesting that the *z-score* measures the out/underperformance. For a one-year horizon, except in the year 2002 and 2004, the average fund outperforms its benchmark. In total, the average fund does slightly better than its benchmark.

Table 2: Descriptive statistics of the *z-scores*

Descriptive statistics for the z-scores of Dutch industry-wide pension funds over the period of 1998-2006. (*), (**), and (***) indicate a significance at the 10%, 5% and 1% levels. T-statistics on the bottom line test whether the mean *z-score* for each year and for the pooled sample is different from 0.

	1998	1999	2000	2001	2002	2003	2004	2005	2006	Pooled
Mean	0.26**	0.27**	0.29***	0.08	-0.89***	0.14	-0.39***	0.30 ***	0.30***	0.03
Med.	0.14	0.19	0.28	-0.08	-1.00	0.04	-0.39	0.25	0.14	0.02
Maxi.	2.25	3.43	3.44	3.84	0.80	1.74	1.34	2.30	2.27	3.84
Mini.	-3.07	-1.22	-1.59	-2.25	-2.91	-1.14	-1.79	-0.87	-0.58	-3.07
Std.	0.89	0.92	0.81	0.90	0.80	0.56	0.56	0.62	0.56	0.84
Skew	-0.38	0.84	0.59	0.98	-0.26	0.67	0.16	1.15	1.03	0.22
Kurt	5.65	4.20	6.17	6.98	3.31	4.23	3.79	4.85	4.41	5.29
Obs.	59	59	60	61	62	63	65	64	62	555
t-stat	2.21	2.25	2.75	0.69	-8.83	1.98	-5.67	3.90	4.26	0.97

We perform a t-test to examine the statistic significance of the above results. For a one-year horizon, during the buoyant period of 1998 through 2000 and the recovering period of 2005 and 2006, the *z-scores* are positive at 5% significance level, while in 2002 and 2004 the *z-scores* are negative at a 5% significance level. When pooled together, the *z-score* is not significantly different from 0. Considering the possible correlation of the *z-scores* for one fund over time, we also calculate the equally-weighted *z-score* across funds and test its time average. We find the average *z-score* over time is 0.04, and t-statistic is *0.28*. In sum we can not reject the hypothesis that industry-wide pension funds as a group over time are not able to out-/under- perform their own benchmarks, namely that they deliver a close-to-benchmark performance. This result agrees with the implication of non-superior selection ability in Goyal & Wahal (2008) that fund trustees cannot time the decisions of hiring and firing asset managers successfully.

6.1 Performance persistence

The descriptive statistics show that the *average* pension fund is not able to beat its benchmark over time. In this section we focus on the performance persistence of the pension funds in our sample. Most studies suggest that there is no performance persistence within mutual funds.[13] Within the rational market framework, this is due to the free movement of competitive capital discussed in (Berk & Green (2004)). In the pension fund industry, however, mandates stay with one asset manager often for more than two years. There is no competitive supply of capital to pension asset managers, so we should expect some persistence here. To this end, we present the results from three methods.

Following the methodology of Fama & MacBeth (1973), we first run a cross-sectional regression of the future *z-score* on the past *z-score* on a yearly basis as in

$$z_{i,t} = a_t + b_t z_{i,t-1} + \epsilon_{i,t},$$

during the period 1999-2006. Using standard OLS we obtain a time series of coefficient estimates ($\hat{a}_t$ and $\hat{b}_t$). Then we perform a t-test on the average estimated coefficients, shown in Panel A of Table 3, which gives a slightly positive correlation ($\bar{\hat{b}}_t = 0.06$). It says the past *z-score* positively predicts the future *z-score*, but not statistically significant at a 5% significance level. We conclude that pension funds as a group do not show persistence in their investment performance.

We also apply a Spearman rank correlation test for persistence, which does not require a normal distribution from the underlying data. In this test we only use the funds with a complete set of *z-scores* in all 9 years, namely a sample of 57 funds. Each year we give a rank to each fund based on its *z-score*. The Spearman rank correlation coefficient for two consecutive years is then computed as

$$\rho_{t,t-1} = 1 - \frac{6\sum_{i=1}^{N} d_{i,t,t-1}^2}{N(N^2-1)},$$

[13] Among others there are Gruber (1996), Carhart (1997), Bollen & Busse (2001). Some recent studies though point out short-run persistence when using daily and monthly returns and certain performance measures such as Bollen & Busse (2005) and Huij & Verbeek (2007).

Table 3: Persistence tests based on regression and ranking

Panel A reports the average coefficients from the cross-sectional Fama-MacBeth regression $z_{i,t} = a_t + b_t z_{i,t-1} + \epsilon_{i,t}$. $\bar{\hat{a}}_t$, and $\bar{\hat{b}}_t$ are the time-average values of the estimated coefficients $\hat{a}_t$ and $\hat{b}_t$ from the Fama-MacBeth regressions. $\overline{R^2}$ is the time average of the exfundation power of the Fama-MacBeth regressions. Panel B reports the Spearman rank correlation coefficient over time and a t-test on the average coefficients. t-statistics are within brackets.

Panel A: Regression					
	$\bar{\hat{a}}_t$		$\bar{\hat{b}}_t$		$\overline{R^2}$
Average coefficient	0.01	(0.05)	0.06	(0.55)	0.09

Panel B: Ranking								
Year	'98-'99	'99-'00	'00-'01	'01-'02	'02-'03	'03-'04	'04-'05	'05-'06
$\rho_{t,t-1}$	-0.19	-0.28	-0.02	0.22	-0.29	0.15	0.15	0.44
$\overline{\rho_{t,t-1}}$				0.26	(0.26)			

where $\sum_{i=1}^{N} d_{i,t,t-1}^2$ is the sum of squared differences of ranks over two consecutive years for all funds. N is the number of funds (or ranks), i.e. N=57 in our case.

For our 9-year sample, we obtain a time series of correlation coefficients for 8 years. As in the previous regression test, we apply a t-test using the average and the standard deviation of the time series, shown in Panel B of Table 3, and find the average coefficient (0.26) is not significantly different from zero. This is consistent with our earlier result.

Table 4: Persistence test based on pension fund portfolios

This table reports the *z-score* in each year of a portfolio formed on the previous year's *z-scores*. A sample of 57 funds with a complete set of the *z-scores* are used. Panels A and B shows 3- and 5-portfolio divisions respectively.

Panel A	3 tertile portfolios							
	1999	2000	2001	2002	2003	2004	2005	2006
1(Best past performer)	-0.08	0.18	0.09	-0.78	0.11	-0.30	0.60	0.80
2	0.57	0.21	0.05	-1.02	-0.04	-0.35	0.22	-0.01
3(Worst past performer)	0.32	0.46	-0.09	-0.96	0.31	-0.40	0.14	0.09
Panel B	5 quintile portfolios							
1(best past performer)	-0.10	-0.03	-0.17	-0.63	0.16	-0.11	0.64	0.87
2	0.02	0.31	0.35	-1.01	-0.05	-0.52	0.43	0.50
3	0.59	0.48	0.05	-0.87	-0.03	-0.22	0.10	-0.07
4	0.23	-0.03	-0.03	-1.01	0.11	-0.34	0.19	-0.01
5 (worst past performer)	0.52	0.63	-0.09	-0.99	0.42	-0.53	0.25	0.22

There might be concerns that the above results may be subject to noises from a few individual funds. Therefore we construct 3 (and 5) portfolios based on their past performance

Table 5: Paired sample t-tests on pension fund portfolios
The table reports the paired sample t-test for mean differences of the *z-scores* of various portfolios from Table 4. Panels A and B shows 3- and 5-portfolio division, respectively.

Panel A:	Mean of paired difference	Std. Deviation	t-test	df	Sig. (2-tailed)
Top - Mid	0.12	0.41	0.85	7	0.42
Mid - Bottom	-0.03	0.20	-0.42	7	0.69
Top - Bottom	0.09	0.38	0.70	7	0.51
Panel B:					
Top1 - Top2	0.08	0.36	0.59	7	0.57
Top1 - Mid3	0.08	0.54	0.39	7	0.70
Top1 - Bottom4	0.19	0.38	1.41	7	0.20
Top1 - Bottom5	0.03	0.50	0.14	7	0.89
Top2 - Mid3	0.00	0.38	0.00	7	1.00
Top2 - Bottom4	0.12	0.29	1.14	7	0.29
Top2 - Bottom5	-0.05	0.35	-0.40	7	0.70
Mid3 - Bottom4	0.12	0.23	1.44	7	0.19
Mid3 - Bottom5	-0.05	0.25	-0.56	7	0.59
Bottom4 - Bottom5	-0.17	0.27	-1.76	7	0.12

and show their future performance. Specifically, every year 3 (and 5) portfolios are formed based on their *z-scores* of the previous year. Then for each individual portfolio the average *z-score* is computed. Repeating this for each year, we obtain a times series of *z-scores* for the 3 (and 5) portfolios in Table 4. If performance is persistent, the best-performing portfolios should provide the best performance in the subsequent years again. However, our results show some cases where the best performing portfolio from the past year provides the worst performance of this year. For example, the best performing portfolio in 1998 is the worst in 1999. The paired sample t-tests among the 3 (and 5) portfolios reported in Table 5 shows none of the test statistics is statistically different from zero. This again confirms no persistence in fund performance over time.

In order to understand the no-persistence better, we look further into the composition of the performance portfolios over time by applying the methodology of Fama & French (2007). In this analysis we only use the three-portfolio division. Each column in Table 6 reports the percentages of funds in the current portfolio that originated from the previous year's top, mid and bottom portfolio, respectively.

We find funds move dramatically among the top, mid and bottom portfolios. For example,

Table 6: Migration statistics

This table reports fund migrations among portfolios sorting on performance. Every year portfolio is formed into top, mid and bottom portfolio according to their *z-scores* in that year. The column shows the composition of the current portfolio that comes from the past top, mid or bottom portfolio respectively. In brackets are the t-statistics testing whether the percentage is equal to 1/3 for a sample of 57 funds. With a degrees of freedom equal to 7, critical values of 10%, 5%, 1% significance level are 1.42, 1.90, and 3 respectively. (*) indicates a significant level of 10%.

	Portfolio based on current performance		
Portfolio based on past performance	Top	mid	Bottom
Top	40%(0.95)	29%(-1.27)	31%(-0.43)
Mid	30%(-1.00)	29% (-1.04)	41%(2.01*)
Bottom	30%(-0.46)	42% (1.84*)	28% (-1.36)
Total	1	1	1

of the current top portfolio, 30% are funds that were in the previous year's bottom portfolio, and another 30% come from the mid portfolio of the previous year. Of the current bottom portfolio 31% and 41% are the funds from the top and mid portfolio in the previous year respectively. We test the hypothesis of random migration of funds among the three portfolios. The null hypothesis is that the migration probabilities should be all equal to 1/3. The test statistics show that we cannot reject the hypothesis at 5% significance level. This random movement of funds among the three performance portfolios underlines the lack of persistence that we found earlier.

In addition to the examination of how funds migrate between performance portfolios over time, we also investigate the contributions to the current *z-score* made by the migrating funds. Results are presented in Table 7.

In 1999, a large part (-0.39) of the bottom portfolio's *z-score* (-0.65) is contributed by the funds that used to be in the top portfolio in the past, while the top portfolio obtains a large chunk of its *z-score* (0.63 out of 1.27) from the funds in the previous bottom portfolio. Similar patterns can be found in year 2001 and 2005, where the current bottom portfolio's negative *z-score* is mostly contributed by the funds in the past top portfolio. In the years 2000 and 2003 the current top portfolio obtains a big portion of its *z-score* from the funds that was in the past bottom portfolio. Such dramatic changes of performance attribution between years again confirm our previous results that past performance does not tell us much

Table 7: *z-score* decomposition of portfolios over time

Portfolios are formed on the *z-scores* in the same way as in Table 6. Each column decomposes the total *z-score* of the current portfolio for each year between 1999-2006 into the *z-scores* contributed by the previous year's portfolios. The portfolio *z-score* is the average *z-score* of the funds included in the portfolio.

	Decomposition current *z-score* for year 1999 through 2006							
	1999	Top	Mid	Bottom	2000	Top	Mid	Bottom
	Total	*1.27*	*0.20*	*-0.65*	*Total*	*1.06*	*0.28*	*-0.51*
Top		0.25	0.05	-0.39		0.21	0.11	-0.15
Mid		0.39	0.13	-0.05		0.32	0.10	-0.21
Bottom		0.63	0.02	-0.20		0.53	0.08	-0.14
	2001				2002			
	Total	*0.90*	*-0.12*	*-0.74*	*Total*	*-0.05*	*-0.93*	*-1.73*
Top		0.45	-0.03	-0.32		-0.04	-0.16	-0.52
Mid		0.34	-0.04	-0.27		0.00	-0.39	-0.63
Bottom		0.12	-0.06	-0.15		-0.01	-0.38	-0.58
	2003				2004			
	Total	*0.68*	*0.04*	*-0.34*	*Total*	*0.25*	*-0.32*	*-0.94*
Top		0.22	0.02	-0.16		0.10	-0.12	-0.26
Mid		0.16	-0.01	-0.17		0.04	-0.06	-0.32
Bottom		0.31	0.02	-0.02		0.10	-0.14	-0.36
	2005				2006			
	Total	*0.97*	*0.26*	*-0.26*	*Total*	*0.94*	*0.17*	*-0.25*
Top		0.52	0.08	-0.10		0.70	0.05	0.00
Mid		0.39	0.01	-0.08		0.15	0.02	-0.11
Bottom		0.05	0.17	-0.08		0.09	0.10	-0.13

about future performance.

6.2 Performance and fund size

The previous analysis shows that as a group the Dutch pension funds in our sample do not show any out- or under-performance with respect to their benchmarks. It is interesting, however, to investigate the cross-sectional difference among funds. Ambachtsheer, Capelle & Scheibelhut (1998) investigate 80 US and Canadian pension funds for the period 1993-1996 and find that large fund size is an important driver for good pension performance, measured by risk-adjusted net value added by asset mix decision and implementation. Reasons are that large size brings economy of scale in operating cost and enables funds to support a full-time professional management team. Following this lead we test whether pension fund size is relevant for explaining the different implementation quality across Dutch pension funds. We

Table 8: Pension fund performance regressions and size
The dependent variable is the time-average *z-score* of each fund. The independent variable is the logarithm of a fund's invested assets in 2006. (***) indicates a significant level of 1%.

Variable	Coefficient	Std. Error	t-Statistic
Constant	-1.20	0.26	-4.59***
Log(assets)	0.14	0.03	4.76***
R-squared	0.54		
Adj. R-squared	0.28		

perform a regression of the time-average *z-score* on the fund's size.

The test is done on a sample of 57 funds with complete *z-scores* over the sample period. Table 8 shows that size indeed matters. Size alone explains almost 28% of the variation in a fund's average *z-score*. The larger funds have a higher average *z-score* than the smaller funds. This finding says that larger funds are more successful in implementing their investment than smaller funds. This result is supported by findings in other samples. Goyal & Wahal (2008) study the decision of hiring and firing asset managers in US pension funds. They find that fund size can explain the post-hiring excess returns, and suggest that large size allows pension fund sponsors to develop expertise in selecting asset managers. Bauer et al. (2007) study the mandate size of delegated portfolios in pension funds. They find size is not a factor driving the benchmark adjusted net return, but size does bring economy of scale in reducing costs of external managers. Both these reasonings support our findings on size, but we cannot distinguish which is exactly at work.

We also compute the *z-scores* on equally weighted portfolios sorted on size, and the results are presented in Panel A and B of Table 9. For tertile portfolios, the size effect is not obvious, but in the quintile portfolios we see a clear difference in the *z-scores* between the largest and the smallest fund. In the range of middle-sized portfolios, there seems no clear difference in their *z-scores*. This says that the top 20% of the funds ranked on size persistently outperform the bottom 20%. As Figure 2 suggests that most funds are small and medium-sized except a few huge funds, we thus form 4 size portfolios in Panel C into the categories of "smaller than €0.1 billion", "between €0.1 billion and €1 billion", "between €1 billion and €10 billion", and "larger than €10 billion".[14] Respectively they contains 8, 27, 18 and 4 funds. We apply paired

[14]Our sample of pension funds contains two very large funds, ABP and PGGM with asset sizes of €208

Table 9: Average *z-score* of size portfolios over time

Panel A and B are 3 and 5 equally weighted portfolios formed on the fund size in 2006. Panel C are 4 portfolios based on some specific size breakpoints, the number of funds are indicated in brackets. The tables reports the equally-weighted *z-score* for each portfolio over time. The sample includes 57 funds that have the complete *z-scores* over the whole sample period 1998-2006.

Panel A	3 size (tertile) portfolios								
Year	1998	1999	2000	2001	2002	2003	2004	2005	2006
1 (largest fund)	0.31	0.31	0.27	0.09	-0.46	0.08	-0.20	0.64	0.51
2	0.48	0.09	0.48	-0.17	-0.83	0.19	-0.37	0.14	0.25
3(smallest fund)	0.06	0.43	0.09	0.13	-1.41	0.11	-0.45	0.18	0.09
Panel B	5 size (quintile) portfolios								
1(largest)	0.49	0.51	0.13	0.10	-0.42	0.19	-0.11	0.68	0.65
2	0.43	-0.04	0.22	0.02	-0.48	-0.07	-0.38	0.46	0.39
3	0.22	0.22	0.52	-0.22	-0.96	0.20	-0.36	0.08	0.21
4	0.15	0.09	0.55	0.17	-1.43	0.24	-0.62	0.10	-0.04
5(smallest)	0.10	0.60	-0.01	-0.02	-1.31	0.08	-0.26	0.25	0.17
Panel C	4 size (quartile) portfolios								
>€10 bln (4)	0.31	1.45	0.57	-0.34	-0.18	0.49	0.29	0.92	1.28
€1-€10bln(18)	0.42	-0.03	0.08	0.22	-0.54	-0.02	-0.30	0.49	0.33
€0.1-€1bln (27)	0.17	0.28	0.43	-0.11	-1.10	0.16	-0.54	0.15	0.15
<€0.1bln (8)	0.37	0.34	0.07	0.15	-1.44	0.17	-0.08	0.22	0.17

sample t-tests to the largest and the smallest portfolio in Table 10. We find that the difference in the *z-scores* between the largest and the smallest size portfolios is statistically significant no matter how many portfolios are formed. To relieve the concern over the power of the t-test in this small sample, we also perform the Wilcoxon signed ranks test that is a nonparametric test suitable for small samples. The test results maintain the original conclusion. Our results are consistent with the findings on US and Canadian pension funds that size is a driver for performance. Large funds implement their investments better than small funds. Due to the data limitation, we cannot explain what causes this better performance. Possible explanation could be negotiation power in lower costs, reputation effect, better monitoring of asset managers, or more expertise in selecting superior asset managers.

billion and €81 billion respectively in 2006. Later on we also form size portfolios from a sample after removing these two biggest funds and our results remain robust.

Table 10: Paired sample t-tests of z-scores on size portfolios

The table reports the paired sample t-test for the *z-score* difference between the top portfolio and the bottom portfolio in the respective 3 (tertile), 5 (quintile) and 4 (quartile) portfolio divisions. Portfolios are formed on size, which is measured by the investment amount in 2006. With a degrees of freedom equal to 8, critical values of 10%, 5%, 1% significance level are 1.40, 1.86, and 2.90, respectively. (*)(**)(***) indicates a significant level at 10%, 5% and 1%.

	Mean of paired difference	Std. dev.	t-test	df	Sig. (2-tailed)
3-portfolio division	0.26	0.33	2.34**	8	0.05
5-portfolio division	0.29	0.29	3.00***	8	0.02
4-portfolio division	0.53	0.58	2.76**	8	0.03

7 Conclusions

One of the main tasks of pension fund trustees is to design an investment strategy that is consistent with the short and long term goals of the fund. The strategic asset allocation decision typically delivers the major part of the investment return, whereas operational and tactical allocation decisions determine excess returns from the strategic benchmarks. Our paper focuses on the effectiveness of the second decision in terms of properly selecting asset managers and monitoring the investment process. Dutch mandatory industry-wide pension funds are obliged to publish a *z-score* to show their net-of-fees investment performance relative to a priori self-selected benchmarks. These scores reflect the implementation quality of the strategic asset allocation. After a study of the *z-scores* on a comprehensive and unique data set of industry-wide pension funds in the Netherlands, we find no outperformance and no performance persistence. We conclude that pension funds on average cannot generate investment returns above their benchmarks. It means that pension funds on average do not add value in selecting and monitoring internal and/or external asset managers. However, we do find that large funds perform consistently better than small funds. This might be attributed to factors like economy of scale in costs, expertise in asset manager selection, or effective monitoring of asset managers. More detailed data is needed to distinguish among these factors to find the driving force of this outperformance. Nevertheless, our results corroborate the empirical trend that smaller pension funds either merge, or are being acquired by bigger funds in order to improve their investment performance.

References

Ambachtsheer, K., R. Capelle & T. Scheibelhut (1998), 'Improving pension fund performance', *Financial Analysts Journal* **54**(6), 15–20.

Bauer, R., R. Frehen, H. Lum & R. Otten (2007), 'The performance of US pension funds: New insights into the agency costs debate', Maastricht University, Working Paper.

Berk, J.B. & R.C. Green (2004), 'Mutual fund flows and performance in rational markets', *Journal of Political Economy* **112**(6).

Blake, D., B.N. Lehmann & A. Timmermann (1999), 'Asset allocation dynamics and pension fund performance', *Journal of Business* **72**(4), 429–461.

Bollen, N.P.B. & J.A. Busse (2001), 'On the timing ability of mutual fund managers', *Journal of Finance* **56**, 1075–1094.

Bollen, N.P.B. & J.A. Busse (2005), 'Short-term persistence in mutual fund performance', *Review of Financial Studies* **18**, 569–597.

Bpf, Vrijstellingsbesluit Wet (2000), 'Besluit van 21 december 2000'.

Brinson, G.P., L.R. Hood & G.L. Beebower (1986), 'Determinants of portfolio performance', *Financial Analysts Journal* **43**, 39–44.

Busse, J.A., A. Goyal & S. Wahal (2006), 'Performance persistence in institutional investment management', *EFA 2006 Zurich Meetings Paper.* .

Carhart, M. (1997), 'On persistence in mutual fund performance', *Journal of Finance* **52**, 57–82.

Coggin, D.T., F.J. Fabozzi & S. Rahman (1993), 'The investment performance of u.s. equity pension fund managers: An empirical investigation', *Journal of Finance* **48**(3), 1039–1055.

Fama, E.F. & J.D. MacBeth (1973), 'Risk, return, and equilibrium: Empirical tests', *Journal of Political Economy* **81**, 607–636.

Fama, E.F. & K.R. French (2007), 'Migration', *Financial Analysts Journal* **63**(3), 48–58.

Good, W.R. (1984), 'Accountability for pension fund performance', *Financial Analysts Journal* **40**(1), 39–42.

Goyal, A. & S. Wahal (2008), 'The selection and termination of investment management firms by plan sponsors', *Journal of Finance* **63**(4), 1805–1847.

Gruber, M.J. (1996), 'Another puzzle: The growth in actively managed mutual funds', *Journal of Finance* **51**, 783–810.

Huij, J.J. & M.J.C.M. Verbeek (2007), 'Cross-sectional learning and short-run persistence in mutual fund performance', *Journal of Banking and Finance* **31**(3), 973–997.

Ippolito, R.A. & J.A. Turner (1987), 'Turnover, fees and pension plan performance', *Financial Analysts Journal* **43**(6), 16–26.

Lakonishok, J., A. Shleifer & R. Vishny (1992), 'The structure and performance of the money management industry', *Brookings Papers on Economic activity. Microeconomics* **1992**, 339–391.

Logue, D.E. & J.S. Rader (1998), *Managing pension plans: A comprehensive guide to improving plan performance*, Harvard Business School Press, Boston, MA.

Sharpe, W.F. (1981), 'Decentralized investment management', *Journal of Finance* **36**, 217–234.

Sharpe, W.F. (1992), 'Asset allocation: Management style and performance measurement', *Journal of Portfolio Management* **18**(2), 7–19.

Tonks, I. (2005), 'Performance persistence of pension-fund managers', *Journal of Business* **78**, 1917–1942.