Performance Persistence of Pension Fund Managers

by

Ian Tonks

Centre for Market and Public Organisation
University of Bristol

January 2002

CMPO is a Leverhulme funded research centre. Information about **CMPO** can be found at http://www.bris.ac.uk/Depts/CMPO/

The data used in this study was provided by CAPS Ltd, and we are grateful for the assistance of Alan Wilcock and Ian Ibbotson. This paper has benefited from comments made at seminars at the Universities of Bristol, Edinburgh, Exeter, Dundee, Lancaster, Southampton, St Andrews; Watson-Wyatt, Financial Markets Group, LSE, and Said Business School, Oxford.

Abstract

This paper examines persistence over time in the performance of fund managers responsible for making the investment decisions of UK pension funds. Previous work on UK pension funds found little evidence of fund manager persistence, but we argue that this may have been due to survivorship bias in the construction of these data samples, which may have disguised true persistence. Using a large sample of pension funds over the period 1983-97 in which there is less survivorship bias, we find strong evidence of persistence in abnormal returns generated by fund managers over one year time horizons.

Keywords: Pension funds, fund management, performance measurement

JEL classification: G23

I Introduction

In this paper we examine whether fund managers consistently add value to the performance of the funds under their management. This is a general question in the context of delegated portfolio management, and we focus on the specific area of the investment decisions of pension funds. Pension funds are major investors in financial markets, owning 20 per cent of UK corporate equity (Myners Report, 2001). A number of recent policy documents in the UK have argued that pension contributions should be investing in tracker funds, on the basis that "there is little evidence that active fund management can deliver superior investment returns for the consumer"[1]. The purpose of this paper is to assess this claim making use of a large dataset on quarterly returns to UK pension funds, in which the fund manager managing the pension fund in each quarter is identified.

Occupational pension schemes in the UK are usually funded and require contributions throughout the employees working life. In a funded scheme an employee pays into a fund which accumulates over time, and then is allowed to draw on this fund in retirement. These schemes are provided by an employer and may pay on a defined benefit or a defined contribution basis. Defined benefit (or final salary) schemes offer a pension, guaranteed by the employer, usually defined in terms of some proportion of final year earnings, and are related to the number of years of employment. Defined contribution (or money purchase) schemes are always funded and convert the value of the pension fund at retirement into an annuity. Under both types of scheme the fund is administered by trustees, usually nominated by the employer, and the trustees, following advice from actuaries, decide whether to invest the assets of the fund in a pooled or segregated investment vehicle.

According to the Occupational Pensions Regulatory Authority (2001), there are nearly 110,000 occupational pension schemes in the UK. The vast majority of these schemes have less than 100 members and are run by insured fund management or as pooled investment schemes. The trustees of the remaining relatively large pension funds typically delegate the management of the pension fund portfolio to fund managers. These fund managers may be in-house, employed directly by the pension fund, or the trustees may out-source the management of the fund to an external fund management house. The pension funds in our sample are these segregated funded occupational pension schemes.

---

[1] para. 420, p. 71 Office of Fair Trading (1997). See also Consumers' Association (1997); Department of Social Security (1998); Financial Services Agency (1999)

In a pooled vehicle, the fund simple purchases units of a diversified investment from a financial institution such as an insurance company. In a segregated vehicle the trustees hire a fund manager (in-house or out-sourced) to make the investment decisions on behalf of the fund according to some specified mandate. The length of this contract is usually three years with the fund manager reporting back to the trustees on a quarterly basis, [Myners (2001), paragraph. 5.64]. According to the Myners' Report (2001) "one-third of schemes had changed manager in the past 12 months . . . . 64 per cent of trustees from smaller funds said they had not changed their manager for more than three years" (paragraph 5.40). Lakonishok, Shleifer and Vishney (1992) refer to the fund management of pension funds as a double agency situation, since the employee as principal, and who will eventually become the recipient of the pension, delegates pension fund decisions to the trustees who in turn delegates the investment allocation decisions to a fund manager.

The objective of this paper is to analyse the existence of performance persistence of individual fund management houses who have been appointed as fund managers of segregated occupational pension funds. Is it possible for a pension fund trustee to identify fund management houses who consistently outperform the benchmark? The significance of this work for trustees and plan advisors is compelling. At the most fundamental asset allocation level, the conclusions of the analysis of the distribution of returns will aid trustees in their decision as to whether to invest their pension fund monies in an active or in a passive vehicle.

II Previous Evidence on Performance of Managed Funds

The early literature of the performance of mutual funds in the US [Jensen (1968)] found that simple tests of abnormal performance did not yield significant returns. More recent work by Daniel, Grinblatt, Titman and Wermers (1997) using normal portfolio analysis shows that mutual fund managers – in particular aggressive-growth funds, exhibit some selectivity ability but that funds exhibit no timing ability. For the UK Blake and Timmermann (1997) examine the returns on 2300 UK open ended mutuals over 23 year period (1972- 1995) gross of fees. Over the period the data includes 973 dead and 1402 surviving funds, and by studying the termination of funds, they are able to shed light on the extent of survivorship bias. They find

economically and statistically very significant under-performance that intensifies as the termination date approaches, and they conclude that survivorship does not alter the results significantly.

The evidence on the average performance of pension funds relative to external benchmarks has also been disappointing. Ippolito and Turner (1987) examined returns on 1,526 US pension funds and find under-performance relative to the S&P500 Index. Lakonishok, Shleifer and Vishney (1992) provide evidence on the structure and performance of the Money Management Industry in the US in general, but focus on the role of pension funds, examining 769 pension funds, with total assets of $129 billion at the end of 1989. They find the equity performance of funds under-performed the S&P 500 by 1.3% per year throughout the eighties. They emphasise that although there is a long literature on the under-performance of mutual funds, pension funds also under-perform relative to mutual funds on average. Coggin, Fabozzi and Rahman (1993) investigate the investment performance of a random sample of 71 US equity pension fund managers for the period January 1983 through December 1990, and find that the average selectivity measure is positive and average timing ability is negative. Though both selectivity and timing are sensitive to the choice of benchmark when management style is taken into consideration. For example they find that funds that target value strategies yielded out-performance of 2.1 per cent per annum, but funds that adopted growth strategies under-performed by -0.96 per cent.

In the UK Blake, Lehmann, & Timmermann, (1999) examine the asset allocations of a sample of 364 UK pension funds who retained the same fund manager over the period 1986-1994. They find that the total return is dominated by asset allocation. Average return from stock selection is negative, and average return to market timing very negative. Although UK equity managers comparatively good at selecting equities – although only 16% of sample beat peer group average. Thomas and Tonks (2001) in a large sample of pension funds find little evidence of any abnormal performance, but find that pension funds seem to follow very similar investment strategies, so that identifying out-performance is difficult.

Although on average fund managers do not outperform, in any sample there is a distribution to the performance, and more recently research on performance measurement has investigated whether the out-performers in the sample continue to outperform in the future. Grinblatt and Titman (1992) find that differences in mutual fund performance between funds persist over 5-year time horizons and this persistence is consistent with the ability of fund managers to earn abnormal returns. Hendricks, Patel and

Zeckhauser (1993) analysed the short-term relative performance of no-load, growth orientated mutual funds, and found the strongest evidence for persistence in a one year evaluation horizon. Malkiel (1995) however argues that survivorship bias is more critical than previous studies have suggested.[2] When an allowance is made for survivorship bias in aggregate, funds have underperformed benchmark portfolios both after management expenses and even gross of expenses. Further he finds that whilst considerable performance persistence existed in the 1970s, there was no consistency in fund returns in the 1980s. Brown and Goetzmann (1995) examine the performance persistence of US mutual funds and claim that the persistence is mostly due to funds that lag the S&P. They demonstrates that relative performance pattern depends on period observed and is correlated across managers, suggesting that that persistence is probably not due to individual managers – it is a group phenomenon, due to a common strategy that is not captured by standard stylistic categories or risk adjustment procedures. This is consistent with herding theories of behaviour (Grinblatt, Titman and Wermers, 1994). They suggest that the market fails to discipline underperformers, and their presence in the sample contributes to the documented persistence. Carhart (1997) demonstrates that common factors in stock returns and investment expenses explain persistence in equity mutual funds' mean and risk-adjusted returns. The only significant persistence not explained, is concentrated in strong underperformance by the worst return mutual funds. His results do not support the existence of skilled or informed mutual fund portfolio managers.

Brown, Draper and McKenzie (1997) examine the consistency of UK pension fund performance, and find "limited evidence of persistence in performance" (p. 155) for a small number of fund managers. Their sample consists of 232 funds 1981-90 and 409 funds 1986-92, and they construct their sample of funds from those that retained the same single fund manager over the time-span of their dataset. They find that this limited consistency holds over different time horizons, samples and classification schemes, though this finding seems to be influenced by the out-performance of one particular fund manager. Blake, Lehmann, & Timmermann, (1999) also examine persistency of long-lived pension fund with a sample of funds that retain the same fund manager. Although they find evidence of persistence in fund returns for UK equity portfolios at the one-year horizon, they argue that the persistence results are entangled with an inverse relationship between fund size and fund performance. They conclude that when an allowance is made for fund size "these regularities [of persistence] are second order" (page 37).

[2] Malkiel points out that only the more successful mutual funds survive. Higher risk funds that fail tend to be merged into other products to hide their poor performance. Also bias from tendency to run incubator funds – run ten different

In constructing their data samples, both the Brown *et al* (1997) and Blake *et al* (1999) studies of UK pension funds specify that the pension fund have the same single fund manager over the length of their respective samples. However this specification of the dataset may have induced survivorship bias in these data samples, since pension fund may have continued to hire the same fund management house, because their performance has satisfied the pension fund trustees, and not triggered their removal. Survivorship bias can affect performance evaluation in two opposite ways. Brown, Goetzmann, Ibbotson and Ross (1992) suggest that if fund volatility is constant across time, but varies cross-sectionally, and if the worst performing funds in a period disappear, then survivorship will induce spurious persistence and bias persistence upwards. Conditional on surviving the best funds tend to have high volatility: in a sample of survivors, first-period winners tend to have high volatility and subsequently win in the second period. On the other hand Grinblatt and Titman (1992) and Hendricks, Patel and Zeckhauser (1993) argue that if fund survival depends on average performance over several periods, then survivorship induces spurious reversals: first-period losers must subsequently win in order to survive, and this biases persistence downwards. Simulation results in Carpenter and Lynch (1999) suggest that persistence is weaker in samples that exhibit survivorship bias, implying that it is the second of the two survivorship bias effects that dominates. Since pension fund mandates are typically over a three year period, the survivorship criteria for UK pension fund managers is likely to depend on performance over several periods [Myners (2001), paragraphs 5.64 - 5.72], and previous studies that have only focused on the same single-manager pension funds may have underestimated the true degree of persistence. In this study we use data on all UK pension funds irrespective of whether they change manager. We might expect that since our sample does not suffer from the same extent of survivorship bias, we will be more likely to identify true persistence.

III Measuring Fund Performance

We examine the consistency or persistence of fund manager performance. That is, we wish to assess whether a fund manager who has performed well in one period can repeat this feat in subsequent periods. Our dataset consists of the returns on pension funds managed by fund managers, and so to examine the consistency in performance of a specific fund manager need to obtain a measure of the his performance across all the funds under his management. We do this by averaging abnormal returns across the funds under management to a particular fund manager. This averaging across pension funds is both equally

---

products – see which are best and market those, ignoring the poor record of the rest

weighted and weighted by fund size, since it could be argued that fund managers put greater effort into managing larger funds. Blake *et al* (1997) has difficulty distinguishing between fund performance and fund size, but this difficulty of interpretation is less important with the average performance of the funds under management. There are a number of tests for persistence, and recently Carpenter and Lynch (1999) have assessed the power of these difference tests particularly in the presence of different types of survivorship bias. Carpenter and Lynch classify persistence tests into two types: performance ranked portfolio strategies, and contingency tables.

Fund manager performance is measured as the average abnormal returns on the funds under management, where the abnormal returns AR_{Pt} for each pension fund P are computed from an asset pricing model. Alternative asset pricing models are the Fama-French three-factor model and the single factor CAPM, where the factor loadings are estimated over the whole sample period. In the three-factor model the standard three factors are the excess return on the market $R_{mt} - r_{ft}$, the returns on a size factor SMB_t which is the difference between the returns on a portfolio of small companies and a portfolio of large companies, and a book-to-market factor HML_t which is the difference in returns on a portfolio of high book-to-market companies and low book-to-market companies.

$$AR_{Pt} = R_{Pt} - r_{ft} - b_P (R_{mt} - r_{ft}) - g_P SMB_t - l_P HML_t \qquad (1)$$

In the case of the CAPM $g_P = l_P = 0$. To calculate abnormal returns we follow a two step procedure,: we first regress returns on the factors to produce the factor loading requiring a minimum of 12 time series observations to estimate the regression parameters. In the second stage we calculate the abnormal return on each fund. The abnormal returns in equation (1) relate to the performance of a pension fund P, managed by fund manager F. To obtain a measure for the performance of fund manager F, the abnormal returns from equation (1) relating to pension fund P of the pension funds managed by fund manager F are averaged for each fund manager and for each quarter.

Performance ranked portfolio tests sort fund manager each period into portfolios based on past performance. Over an initial period, called the ranking period, the performance of fund managers are compared and ranked. The ranking period can be either one quarter, four quarters (one year) or twelve quarters (three years). This averaged abnormal return across pension funds is attributed to the skills of fund

manager *F* in the ranking period. Fund managers are ranked on the basis of the average return on the funds under management in the ranking period, and five portfolios are formed on the basis of this ranking, with equal numbers of fund managers in each portfolio. The top portfolio consists of those fund managers with the highest average abnormal returns in the ranking period, down to the bottom portfolio with those fund managers with the lowest average abnormal returns.

We then compute the equally weighted average portfolio abnormal return of the top and bottom portfolios over a subsequent evaluation period, which we denote AV5(τ) and AV1(τ) respectively, where τ denotes the particular evaluation period. We then advance the ranking period by one period, and repeat the ranking process and subsequent evaluation. We report the average abnormal returns AV5 and AV1 of the top and bottom portfolios, in the evaluation periods, averaged over all evaluation periods. The evaluation period can also be either one quarter, one year or three years. These procedures are followed for overlapping periods throughout the full period of the dataset, and we compute DIF as AV5-AV1, and then report TDIF, which is a t-statistic on DIF, which is calculated after allowing for the autocorrelation induced by the overlapping observations. Under the null hypothesis of no persistence the value of DIF should be centred on zero, which would mean that past performance is no predictor of future performance. From their simulations Carpenter and Lynch find that the that the persistence test based on TDIF is the best specified under the hypothesis of no persistence, and the most powerful against the alternatives considered.

In these persistency tests we examine alternative ranking and evaluation time periods, since it may be the case that persistency is only apparent at particular time intervals. For example to test for long run persistency 12QR12QE means we form portfolios on the basis of twelve-quarter ranking period and twelve quarter evaluation period. To test for short-run persistency, or the "hot-hands" phenomenon, we examine 1QR1QE, which means one quarter ranking and one quarter evaluation period.

Contingency tables classify funds as winners or losers in each of two consecutive time periods, and the numbers of winner-winner (WW), winner-loser (WL), loser-winner (LW), and loser-loser (LL) combinations are counted. We compute the following related statistics: a) Percentage of repeat winners, PRW = WW/(N/2) is a purely descriptive statistic which gives the percentage of the sample that are in the winner-winner box; b) Cross-product ratio CP = (WW x LL)/(WL x LW); which is also referred to as the odds-ratio and $\log(CP)/\sigma_{\log(CP)}$ has a standard normal distribution, with $\sigma_{\log(CP)} = \sqrt{[(1/WW) +$

(1/WL) + (1/LW) + (1/LL)], so that we may test for the statistical significance of deviations of the cross-product ratio from unity; c) Chi-Squared test with 1 d.o.f. where CHI = $\{(WW - N/4)^2 + (WL - N/4)^2 + (LW - N/4)^2 (LL - N/4)^2\}/N/4$; and we may reject independence if CHI exceeds the critical value of 3.84 for a 5% test; and d) TCS is the t-statistic for the slope coefficient in the cross-section OLS regression of evaluation period abnormal returns on ranking period abnormal returns. Carpenter and Lynch (1999) find that in the presence of survivorship bias the Chi-Squared test performs best, and that the cross-sectional t-test (TCS) is the weakest. We will discuss the results of these persistence tests in the light of the simulation results of Carpenter and Lynch (1999) on the relative power of these alternative tests.

Ferson and Schadt (1996) have advocated allowing for the benchmark parameters to be conditioned on economic conditions: called conditional performance evaluation, on the basis that some market timing skills may be incorrectly credited to fund managers, when in fact they are using publicly available information to determine future market movements. In which case Ferson and Schadt argue that the predictable component of market movements should be removed in order to assess fund managers private market timing skills. Under a conditional version of the three-factor model, the abnormal returns equation (1) becomes

$$AR_{Pt} = R_{Pt} - r_{ft} - b_P(Z_{t-1})\,(R_{mt} - r_{ft}) - g_P\,SMB_t - l_P\,HML_t \qquad (2)$$

where Z_{t-1} is a vector of instruments for the information available at time t (and is therefore specified as $t-1$) and $b_P(Z_t)$ are time conditional betas, and their functional form is specified as linear

$$b_P(Z_t) = b_0 + B'z_{t-1} \qquad (3)$$

where $z_{t-1} = Z_{t-1} - E(Z)$ is a vector of deviations of the Zs from their unconditional means. Implementing this approach involves creating interaction terms between the market returns and the instruments. The instruments used in this study are: lagged treasury bill rate, dividend yield, a default premium (the difference between low and high quality corporate bonds), and the slope of the term structure (the difference between long and short run government bond yields). To implement the conditional performance evaluation tests, we follow the same two step procedure outlined earlier, though we now require 20 time-series observations to produce meaningful coefficients. First for each fund we run a time series regression of excess returns

against the three factors, with the interaction terms included, which enables us to calculate the abnormal return on each fund. In the second stage we compute the abnormal returns.

IV Data

The data used in this study was provided by the Combined Actuarial Performance Services Ltd (CAPS). It consists of quarterly returns on UK equity portfolios of 2,175 UK pension funds from March 1983 to December 1997. Typically over this period a UK pension fund invested about 57% of assets in UK equities, so that our dataset consists of returns on the major asset class in which UK pension funds invest. In addition for each fund-quarter the manager of the fund and the size of the fund is provided. CAPS provide a performance measurement service for about half of all segregated pension fund schemes in the UK. There is one other major provider of pension fund performance: WM Ltd. The full dataset consists of a total of 59,509 observations on quarterly returns and fund size, and the maximum number of Quarters is 56.

Table 1, Panel A illustrates the Distribution of fund quarters over the dataset, and shows that 50 per cent funds have 24 or less observations, and the average life of a fund in the data is just less than seven years. This high attrition rate is partly explained by the closure of funds due to the sponsoring companies merging, or becoming insolvent, but the predominant reason is due to the fund switching to an alternative performance measurement service. As we have already mentioned there are two major performance measurement services in the UK: CAPS and WM, and pension funds will typically subscribe to one or other of these two services. When a pension fund changes fund manager, it may be that the fund manager has a preference to be assessed by one of these two measurement services. If the new fund manager has been appointed following a run of poor performance by the previous fund manager, and the new fund manager switches performance measurement services, there is a possibility of survivorship biases, which bedevils performance evaluation studies. However this switching between measurement services should be symmetric: so that although a pension fund may drop out of our dataset because of poor performance, there will be new entrants into our dataset as pension funds that have previously been assessed by the alternative measurement service, and changed fund manager because of poor performance, switch into the CAPS measurement service. In effect our sample loses pension funds due to poor performance, but they are replaced by poor performers from the alternative measurement service. Carhart (1997) draws a distinction between survivorship bias and look-ahead bias. True survivorship bias is a property of the

sample selection method, and results from only including funds in a sample that survive until the end of the sample period. Look-ahead bias is a property of the test methodology, and results from imposing conditions on the funds in the sample to produce meaningful econometric results. In our study true survivorship problems should not be a concern, since we have replacement of poorly performing funds with other poorly performing funds. However look-ahead bias may affect our results, since we require at least 12 time series observations to estimate the parameters in equation (1).

Panel B shows that the management of pension fund equity portfolios is relatively concentrated: There are a total of 189 different fund managers (including in-house managers), 2 per cent of fund managers manage only 17 quarters or less (across funds), and 50 per cent manage across 45 quarters or less. Since the average life of a fund is just under seven years (28 quarters), this implies that fifty per cent of fund managers in the dataset are managing only two funds. Panel C provides further evidence on the concentration of fund management. We have ranked the fund managers is terms of the number of fund-quarters under management. The top ranked fund manager[3] (1RMan) managers 10.8% of observations, the second ranked 2RMan managers 5.6% and 3RMan managers 4.8%, and another 14 fund managers (4RMan-18RMan) manage a total of 23.14% of observations. 1RMan manages across 244 funds, and 81.04% of these funds' observations are using 1RMan. There is also a multi-manager category and a change of manager category (Δman). Most funds use a single fund manager in any quarter, but 659 funds have multiple fund managers at some time, and 29.07% of all observations have multiple fund managers. In the case of the multi-manager category we do not have information on the identity of the multiple managers, and further the definition of multiple fund manager has changed over time. Only 85 funds use the same fund manager over the fund's life.

The Myners' Report (2001) draws a distinction between two types of investment decision making "Peer Group Total Fund benchmark" and "Customised total fund benchmark". In the former, pension fund trustees delegate both asset allocation and security selection to one or more fund management companies. Under the latter, the trustees first decide upon an asset allocation strategy, and then employ one or more fund managers to manage the assets of a particular asset class. Under both alternatives the trustees may employ multiple managers, but the fund is more likely to employ multiple managers under the latter category. According to Figure 3.2 of the Myners' Report (2001), over time there has been a decline in the

use of the peer group benchmarks with a single fund manager managing a balanced portfolio, with a trend towards funds employing multiple managers with specific expertise.

Tables 2 provides descriptive statistics on the returns to, and the size of, the UK equity portfolios of the pension funds in our dataset. From Panel A, the average discrete quarterly return over all funds over all quarters is 4.32%, compared with an average discrete return of 4.38% for the FT-All Share Index. The overall standard deviation of these returns is 8.67%, and the distribution of returns also emphasises the variability in returns. But these pooled measures disguises an important statistic that is made clear in Panel C, which is that the between funds standard deviation is much less than the within fund distribution. This implies that for a particular quarter the distribution of fund returns is tightly packed around the mean, but that over time the variability of returns is much higher. In fact the correlation between the time series values of the FT-All Share index and the average return each quarter across the pension funds is 0.995. The contrast in the within and between standard deviations might be indicative of the herding behaviour of pension funds suggested by Lakonishok *et al*. (1992). The between variation of fund returns by manager is much smaller than the within manager standard deviation, which implies that is may be difficult to identify individual fund manager performance. Our subsequent results of manager performance are all the more striking, given this feature of the data.

Table 2 Panel A also report on the distribution of returns weighted by the value of the fund at the beginning of each quarter. The value weighted average return of 3.80% implies that small funds have a slightly higher return than large funds. In the subsequent regression analysis, we require a minimum number of observations to undertake a meaningful statistical analysis, and we imposed the requirement that time series fund parameters are only estimated when there were 12 or more quarterly returns for that fund. This cut-off value of three years accords with the typical fund mandate. Table 2 Panel A reports the distribution of returns of the sub-sample of 1724 funds with at least 12 time series observations, and this may be compared with the distribution of returns across the whole sample, to check that the sub-sample is indeed representative.

In Panel B of Table 2 we report statistics of the size of the equity portion of the pension funds in our sample, at three different dates at the start, in the middle and at the end of our sample. The size distribution

[3] This fund manager is actually identified as #28, in our dataset - all the fund managers are identified by a code.

is highly skewed with a large number of very small funds. For example in 1997 the median size fund had an equity portfolio of 28 million pounds, whereas the largest fund had an equity portfolio of over 9 billion pounds. In Panel C we also report the distribution of fund size across funds and across fund managers. We report two measures of fund size: smv is the starting market value of the equity portfolio of the fund at the start of each quarter; smv97 is the starting market value of the funds, with the fund value inflated to December 1997 values. This measure of fund size at constant prices is obtained by compounding to December 1997 fund size (smv) in each quarter by the average rate of return over the life of the fund. Panel C shows that the distribution of firm size when measured at non-constant prices is bigger between funds than within funds. This difference in the between and within distribution of fund size is much sharper when size is measured at constant prices, where the within variation is only a fifth of the between variation. This statistic emphasises that pension fund size is relatively constant over time, with most variation occurring between funds.

In this study we use data on all UK pension funds irrespective of whether they change manager. The Brown *et al* (1997), and Blake *et al* (1999) studies of UK pension funds specify that the pension fund has the same fund manager over the length of their respective samples. As we have argued, it is likely that survivorship bias is more of an issue in same manager funds, since pension fund trustees who have retained the same fund manager, are likely to have been satisfied with that fund manager's performance. We expect that our sample does not suffer from this survivorship bias, and consequently we will be more likely to identify true persistence.

V Results

The results of the persistency tests of fund manager performance for the base case of the three factor model of abnormal returns are reported in Table 3. Panel A reports the performance ranked portfolio tests, and Panel C the contingency table tests for the three factor model. Each panel has three rows representing the number of time periods over which the ranking and evaluation periods have been evaluated. The first two columns in Panel A report the average evaluation period returns of top and bottom quintile portfolios, formed on the basis of ranking period fund manager abnormal returns. In panels A and C the measure of fund manager abnormal returns is the equally weighted average abnormal return of the funds under management in a particular quarter.

It can be seen that for each row the mean return on the high quintile portfolio is always greater than that on the low quintile portfolio (DIF is always positive). We might think of DIF as being the return to an arbitrage portfolio which has been constructed by going long in the high quintile portfolio (AV5) and short in the low quintile portfolio (AV1). The one quarter on one quarter results (1QR1QE) shows only weak evidence of persistency as measured by TDIF, though the longer term abnormal returns show much stronger evidence of persistency. The results in Panel A suggest that there is some persistency at all time horizons, with the strongest at one year.

These findings are confirmed from the contingency table tests in Panel C. The chi-squared test on independence is easily rejected for the one year horizon abnormal returns, and the odds ratio is also significantly different from unity at both the one quarter and one year horizons. Similarly the t-statistic on the slope coefficient in the cross-section regression of one-year abnormal returns on lagged one year abnormal returns is 6.17 for the one year abnormal returns, indicating significant persistence at one year. One slight inconsistency in these tables is that CHI implies that quarterly abnormal returns are more persistent that three year returns, whereas all the other measures suggest that it is the longer term returns are more persistent than the short-term measures. Though the percentage of repeat winners is only 52 per cent for the quarterly horizon, but rises to 57 per cent for the longer horizons. Brown *et al* (1992) suggest that some persistency in performance may be due to consistently poor performance of some funds, which for institutional reasons are allowed to continue. The percentage of repeat winners in the PRW columns does not support that finding in our dataset. Recall that Carpenter and Lynch (1999) suggest that the TDIF measure is the most powerful from among the alternative tests for persistency. So that we should put more emphasis on the findings in Panel A which report persistence at the one year horizon.

Blake *et al* (1999) produce a statistic which is directly comparable with the data in Panel A. They find that the return to the "zero net investment portfolio" for UK equities in the case of abnormal returns estimated from a multi-factor model (similar to out three-factor model) for the year on year ranking and evaluation strategies yielded a annual return of 0.5 of a per cent. This is substantially less than the annualised return of 1.56 per cent implied by the equivalent strategy in panel A. We Will return to this comparison below.

Our measure of fund manager performance was computed by taking the equally weighted average abnormal return of the pension funds under management in a particular quarter, as a measure of the fund

manager's performance in that quarter. A potential criticism of this approach is that fund managers will give a better service to larger pension funds, since if fees are *ad valorem* and based on the value of assets under management (Myners, 2001), the larger pension funds are paying a higher fee to the fund manager. We therefore computed an alternative measure of fund manager performance as the average performance of funds under management weighted by the fund size at the beginning of the quarter. The results of this alternative measure of fund manager performance, still based on a three-factor model of pension fund abnormal returns, is given in panels B and D of Table 3. The performance ranked portfolio test results are given in panel B, and it can be seen that the evidence suggests even stronger evidence of persistence at the one year and three year horizons: the quarterly return on DIF is 0.41 percentage points. The contingency tests in panel D also report evidence of persistence at the one year horizon, with less evidence in the shorter or longer terms.

In Table 4 we report the results of recalculating the measures of pension fund abnormal returns using alternative asset pricing models. In panels A and C we report the performance ranked portfolio tests and the contingency tables derived from a CAPM measure of abnormal returns. Again we find that DIF has its highest vale and is significant at the one year horizon, with the value of DIF at 0,37 percentage points being very similar to the equivalent value of DIF in the three-factor model. At other time horizons, the value of DIF being insignificantly different from zero, and there is even some evidence of fund manager performance reversals at the long-horizon. The contingency tables in panel C for the CAPM measure also report significant persistence at the one year horizon, and also find evidence of persistence at the short horizon: the CP, CHI, and TCS statistics are all significant over successive one quarter periods. At the twelve quarter horizon there is evidence of fund manager performance reversals both in a negative TCS and the fact that the odds ratio (CP) is less than unity.

In panels B and D we report the performance ranked portfolio tests and the contingency tables derived from a three factor model measure of conditional abnormal returns, where the conditioning variables include lagged macroeconomic variables. Following the Ferson and Schadt (1996) argument, by excluding lagged macroeconomic factors which were publicly available at the time that fund managers make their investment decisions, the resulting abnormal returns are more likely to reflect the fund managers' true abilities. According to panel B this measure of fund manager performance produces the significant consistency in performance at all horizons, with the strongest persistence at the one year period. The annualised value on

DIF is 1.93 per cent. The contingency tables in panel D also report persistence in the short-run and medium-term, though not in the longer-term. The slightly stronger results on persistence in the case of the conditional benchmarks implies that the unconditional benchmarks are disguising fund managers true abilities, though to only a minor extent.

Finally we wish to compare our results with the previous studies in the UK on performance persistence of pension funds. Earlier work by Brown *et al* (1997) and Blake *et al* (1999) found little evidence of performance persistence in the returns of pension funds, but as we have already mentioned, both studies concentrated on pension funds that employed the same fund-manager over a long time period, and we have suggested that this may have led to a survivorship bias that disguised the true level of persistence. To examine the effects of imposing these restrictions on our data we redefined our sample using similar criteria. In table 5 we report the results of imposing two data restrictions on our dataset: we only consider those pension funds that have remained in the datafile for all 56 quarters from March 1984 to December 1997, and in addition have retained the same fund manager over those 56 quarters. This results in a sample of 129 pension funds, which is smaller than the number of pension funds in the Brown *et al* (1997) and the Blake *et al* (1999) samples, because the length of the dataset is longer in our study. In Table 5 we report the results of including these restrictions into our dataset on our performance persistence statistics, for the unconditional three factor model of fund performance. From Panel A, it can be seen that the imposition of the same fund manager, and the requirement of long-lived funds, reduces the value of DIF for the one year horizon from 0.0039 to 0.0023. A simple t-statistic on these two values shows that they are significantly different, with the implication that imposing the restrictions on our dataset reduces the observed degree of persistence at the one year horizon. Interestingly the performance ranked portfolio tests for the one quarter and twelve quarter horizons are not affected by the data restrictions. The contingency tables of fund manager performance for the restricted sample in panel B of table 5 on the other hand, seems to find stronger evidence of persistence than for the full sample from panel C in Table 3: the CHI and CP values are significant at all time horizons. This may seem surprising, but is exactly the result predicted by the simulation findings of Carpenter and Lynch (1999) who find that in the presence of survivorship bias the contingency tests are more robust, "Overall, in the absence of survivor bias, the DIF t-test using one year evaluation periods appears to be the best specified under the null hypothesis of no persistence and one of the most powerful against the alternatives that we consider. Also well-specified and powerful in large samples the chi-squared test is the most robust to the presence of survivor bias". (page 367).

VI Conclusions

Brown *et al* (1997) report that the persistence is caused by repeat losers - not in our case

With the advent of low cost stakeholder pensions in the UK, there has been a continuing trend into index funds and a movement away from active fund management, as a portfolio strategy for pension funds. However the results in this paper, from a large sample of occupation pension funds, suggest that there appears to be a role for active fund management of pension funds.

We have measured the abnormal returned generated by fund management houses in managing the equity portfolios of UK pension funds over the period 1983-97. We have found evidence of significant persistence in the performance of fund managers at the one-year time horizon using a number of different consistency tests, as well as some evidence of persistence at other time intervals. We found that the returns on a zero investment portfolio of a long position in a portfolio of fund managers that performed well over the previous 12 months and a short position in a portfolio of fund managers that performed poorly, would have yielded an annualised abnormal return of 1.56. According to Carpenter and Lynch (1999) this test is the most powerful in detecting persistence in performance. This is a significant result and conflicts with the evidence presented in the Myners' Report (2001) which states "selecting managers according to past performance figures first and brand second is widely acknowledged to be a poor way to select a manager" (paragraph 5.44). We have argued that earlier studies of performance persistence in pension fund returns may have induced a selection bias by restricting the data sample to the same fund manager over a long time period, and this survivor bias may have reduced the level of persistence in the sample. Using our dataset with a restriction that only long-lived funds with the same fund manager be included, did indeed reduce the return on a zero investment portfolio.

Two caveats are in order. First we have made no allowance for the costs of fund management. We have found that some fund managers generate consistent abnormal returns above the benchmark portfolios, but whether these abnormal returns outweigh the costs of active fund management is not an issue that we have addressed. Myners (2001) suggests that annual fund management costs are typically around 40 basis points of the funds under management for a £100 million mandate, which sound attractive relative to the 156 basis points we have identified. However we do not know whether the fund managers with the persistent performance are charging fees that reflect their abilities. Second, having identified performance

persistence at the one year horizon, it is less clear how pension fund trustees could take advantage of this fact. One implication of these results might be that a pension fund mandate should be set up on a yearly basis, however this would ignore the substantial transactions costs involved in shifting a pension funds assets from one fund manager to another, on such regular intervals.

References

Blake, D., Lehmann, B. and Timmermann, A. "Performance Measurement Using Multi-Asset Portfolio Data; A Study of UK Pension Funds 1986 - 1994", Pensions Institute.

Blake, D., B. Lehmann and A. Timmermann (1999) "Asset Allocation Dynamics and Pension Fund Performance", *Journal of Business,* vol. 72, 429-461.

Blake, D. and Timmermann A. (1997) "The Birth and Death Processes of Mutual Funds" *European Finance Review*,

Brown, G. P. Draper and E. McKenzie (1997) "Consistency of UK Pension Fund Performance", *Journal of Business Finance and Accounting* vol. 24, March, 155-178.

Brown, S., W. Goetzmann , R. Ibbotson and S. Ross (1992) "Survivorship Bias in Performance Studies", *Review of Financial Studies*, vol. 5, 553-580.

Brown, S. and W. Goetzmann (1995) "Performance Persistence", *Journal of Finance,* vol. 50, 679-98.

Carpenter, J.N., and A.W. Lynch (1999) "Survivorship Bias and Attrition Effects in Measures of Performance Persistence", *Journal of Financial Economics*, vol. 54, 337-374.

Carhart, M. (1997) "On Persistence in Mutual Fund Performance", *Journal of Finance*, vol. 52, 57-82.

Coggin, T.D., F.J. Fabozzi, and S. Rahman (1993), "The Investment Performance of US Equity Pension Fund Managers: An Empirical Investigation", *Journal of Finance*, vol. 48, 1039-55.

Consumers' Association, *A Blueprint for Better Pensions* (1997)

Daniel, K., M. Grinblatt, S. Titman and R. Wermers (1997) "Measuring Mutual Fund Performance with Characteristic Based Benchmarks", *Journal of Finance*, vol. 52, 1035-58.

Department of Social Security (1998), *A New Contract for Welfare: Partnership in Pensions* (HMSO)

Financial Services Agency (1999), *Comparative Information for Financial Services.*

Grinblatt, M. and S. Titman (1992) "Persistence in Mutual Fund Performance", *Journal of Finance*, vol. 47, 1977-84.

Hendricks, D., J. Patel and R. Zeckhauser (1993) "Hot Hands in Mutual Funds: Short Run Persistence of Relative Performance, 1974 – 1988", *Journal of Finance*, vol. 48, 93-130.

Ippolito, R.A. and J.A. Turner (1987) "Turnover, fees and pension plan performance", *Financial Analysts Journal*, vol. 43, 16-26.

Lakonishok, J.A., A. Shleifer and R.W. Vishny (1992) "The structure and performance of the money management industry", *Brookings Papers on Economic Activity*, 339-391.

Malkiel, B.G. (1995) “Returns From Investing in Equity Mutual Funds 1971 to 1991”, *Journal of Finance*, vol. 50, 549-72.

Myners, P. (2001) *Report on Institutional Investment*, (H.M. Treasury)

Occupational Pensions Regulatory Authority (2001), *Pensions Registry*.

Office of Fair Trading (1997), *Report of the Director General’s Inquiry into Pensions.*

Thomas, A. and I. Tonks (2001) "Equity Performance of Segregated Pension Funds in the UK", *Journal of Asset Management*, vol. 1 no. 4, 321-343

Table 1: Descriptive Statistics on Pension Funds and Fund Managers

Panel A: *Fund-Quarters*

No. of Funds	2,175	No. of Quarters	59,509

Distribution of Fund-Quarters

min	5%	25%	50%	75%	95%	max
1	4	12	24	41	56	56

Panel B: *Manager-Quarters*

No. of Managers	191	No. of Quarters	59,509

Distribution of Manager-Quarters

min	5%	25%	50%	75%	95%	max
2	4	17	45	179	1,063	17,299

Panel C: Distribution of Managers Across Funds

Fman code	Overall		Between		Within
	Freq	%	Freq.	%	%
Multi-manager	17,299	29.07	659	30.3	78.10
1RMan	6,410	10.77	244	11.22	81.04
2RMan	3,318	5.58	184	8.46	59.55
3RMan	2,881	4.84	116	5.33	73.40
4RMan-17RMan[#]	13,758	23.14	681	31.31	68.16
18RMan-189RMan	15,595	26.22	965	44.65	58.84
ΔMan	248	0.42	225	10.34	2.64
Total	59,509	100.0	3,074 (n=2,175)	141.33	63.43

where total within = (659*78.1+244*81.04+ .....)/3,074; 'N'RMan denotes first ranked fund manager by frequency of observations. [#] denotes that each of these fund managers had greater than 1% of the overall frequency.

Table 2: Descriptive Statistics

Panel A: Returns Across Quarters and Funds

Returns, R_{Pt}	All	Weighted by fund size	>12 Quarters	FT-All ShareRets
Mean	0.0432	0.0380	0.0428	0.0438
Std. Dev.	0.0867	0.0814	0.0867	0.0834
Distribution of returns:				
10%	-0.0543	-0.0537	-0.0543	
25%	0.0016	0.0016	0.0015	
50%	0.0463	0.0441	0.0459	
75%	0.0896	0.0747	0.0885	
90%	0.1525	0.1346	0.1527	
Obs.	59,317	59,314	56,403	56
# funds	2170	2170	1724	

Panel B: Distribution of Fund Size Across Funds

	Fund Size at start of Quarter (£m)		
	March 1983	Dec 1990	Dec 1997
Mean	25.02	50.24	102.27
Std. Dev.	85.01	194.45	387.30
Distribution of Fund size:			
10%	0.441	1.36	6.02
25%	1.06	3.31	12.39
50%	3.20	8.35	28.12
75%	14.25	27.36	70.14
90%	51.64	102.88	221.90
Obs.	833	1131	1004

Panel C: Returns and Fund Size Across Funds and Across Fund Managers

Variable		Mean	Std. Dev	Min	Max	Observations	
R_{Pt}	overall	0.04323	0.08672	-0.5257	0.8707	No. obs	59317
	between funds		0.01652	-0.1285	0.2366	# funds	2170
	within funds		0.08628	-0.5144	0.71385	Obs per fund	27.335
	between managers		0.0177	-0.1147	0.117045	# managers	189
	within managers		0.0866	-0.5241	0.872249	Obs per man	313.847
smv	overall	58.4044	240.130	0	9,108.619	No. obs	59453
	between funds		174.391	0.013	5,096.643	# funds	2175
	within funds		89.5995	-3,352.67	4070.38	Obs per fund	27.3347
	between managers		506.0059	0.2137	6,747.353	# managers	190
	within managers		164.8412	-2,865.32	4,258.418	Obs per man	312.911
smv97	overall	204.17	869.1149	0	24,411.38	No. obs	59437
	between funds		1,049.211	0.0044241	21,804.67	# funds	2170
	within funds		189.6773	-5,106.381	7,187.195	Obs per fund	27.3903
	between managers		1,001.512	0.0849746	11,851.75	# managers	189
	within managers		733.1108	4,268.246	24,260.76	Obs per man	314.481

where R_{Pt} is the quarterly return on fund P in quarter t; and smv is the fund market value at the beginning of the quarter

Table 3: Persistence Tests based on 3-factor Abnormal Returns of Fund Manager Performance

Panel A. Performance ranked portfolio tests of fund manager performance (unweighted abnormal returns)

	AV5	AV1	DIF	TDIF
1QR1QE	-0.0002	-0.0018	0.0016	1.41
4QR4QE	0.0015	-0.0023	0.0039	6.72*
12QR12QE	0.0013	-0.0005	0.0018	3.10*

Panel B. Performance ranked portfolio tests of fund manager performance (weighted by fund size)

	AV5	AV1	DIF	TDIF
1QR1QE	-0.0004	-0.0017	0.0013	1.14
4QR4QE	0.0015	-0.0025	0.0041	8.15*
12QR12QE	0.0008	-0.0005	0.0013	3.36*

For performance ranked tests, fund managers are sorted each year into quintile portfolios based on past performance of the pension funds under management - average (weighted and unweighted) abnormal returns of each fund over the ranking period. The equally weighted average portfolio abnormal returns of the top and bottom portfolios over the subsequent evaluation period is computed; AV5 and AV1 are the abnormal returns of the top and bottom portfolios in the evaluation period, averaged over all time periods in the sample. There are three different ranking and evaluation periods: 12QR12QE means three-year ranking period and three year evaluation period, and 1QR1QE means a one quarter ranking period and one quarter evaluation period. This procedure is followed for overlapping periods throughout the full period of the dataset, and DIF is AV5-AV1, and TDIF is a t-statistic on DIF, allowing for the autocorrelation induced by using overlapping observations.

Panel C: Contingency tables of fund manager performance (unweighted abnormal returns)

	N	PRW	CP	Z-stat	CHI	TCS
1QR1QE	5,360	0.523	1.139*	2.373	6.169*	0.15
4QR4QE	1,166	0.578	1.531*	3.614	15.125*	6.17*
12QR12QE	195	0.574	1.249	0.773	1.821	1.72

Panel D: Contingency tables of fund manager performance (weighted by fund size)

	N	PRW	CP	Z-stat	CHI	TCS
1QR1QE	5,360	0.520	1.112	1.936	4.282	-0.29
4QR4QE	1,166	0.576	1.563*	3.788	15.520*	5.45*
12QR12QE	195	0.574	1.049	0.167	2.149	1.23

Fund managers are classified as winners or losers based on abnormal returns in each of two consecutive time periods, and the numbers of winner-winner (WW), winner-loser (WL), loser winner (LW) and loser-loser (LL) are counted. The following statistics are computed: a) Percentage of repeat winners, PRW = WW/(N/2); b) Cross-product ratio CP = (WW x LL)/(WL x LW); where $\log(CP)/\sigma_{\log(CP)}$ has a standard normal distribution, and $\sigma_{\log(CP)} = \sqrt{[(1/WW) + (1/WL) + (1/LW) + (1/LL)]}$; c) Chi-Squared test with 1 d.o.f. where CHI = $\{(WW - N/4)^2 + (WL - N/4)^2 + (LW - N/4)^2 (LL - N/4)^2\}/N/4$, and N is the number of pairs; and d) TCS is the t-statistic for the slope coefficient in the pooled cross-section OLS regression of evaluation period abnormal returns on ranking period abnormal returns.

Table 4: Persistence Tests based on Alternative Measures of Abnormal Returns of Fund Manager Performance

Panel A. Performance ranked portfolio tests of fund manager performance: CAPM abnormal returns

	AV5	AV1	DIF	TDIF
1QR1QE	0.0002	-0.0019	0.0021	1.28
4QR4QE	0.0017	-0.0020	0.0037	3.07*
12QR12QE	-0.0002	0.0008	-0.0010	-1.76

Panel B. Performance ranked portfolio tests of fund manager performance: conditional abnormal returns

	AV5	AV1	DIF	TDIF
1QR1QE	0.0006	-0.0021	0.0027	2.43*
4QR4QE	0.0020	-0.0028	0.0048	8.04*
12QR12QE	0.0011	-0.0002	0.0013	3.34*

See notes to Table 3

Panel C: Contingency tables of fund manager performance based on CAPM abnormal returns

	N	PRW	CP	Z-stat	CHI	TCS
1QR1QE	5,360	0.526	1.164*	2.783	8.272*	4.19*
4QR4QE	1,166	0.587	1.700*	4.488	21.575*	6.37*
12QR12QE	195	0.523	0.826	-0.663	1.615	-3.28*

Panel D: Contingency tables of fund manager performance based on conditional abnormal returns

	N	PRW	CP	Z-stat	CHI	TCS
1QR1QE	4,980	0.537	1.273*	4.245	18.593*	3.18*
4QR4QE	1,088	0.574	1.545*	3.565	13.919*	7.03*
12QR12QE	183	0.536	1.067	0.218	0.497	0.84

See notes to Table 3

Table 5: Persistency Tests based on 3-factor Abnormal Returns of Fund Manager Performance for Restricted Sample of Single Same-Manager Funds

Panel A. Performance ranked portfolio tests of fund manager performance

	AV5	AV1	DIF	TDIF
1QR1QE	0.0018	0.0001	0.0017	2.11
4QR4QE	0.0028	0.0005	0.0023	3.35*
12QR12QE	0.0038	0.0019	0.0019	4.30*

see notes to Table 3

Panel B: Contingency tables of fund manager performance

	N	PRW	CP	Z-stat	CHI	TCS
1QR1QE	7,060	0.522	1.159*	3.093	9.738*	6.63*
4QR4QE	1,668	0.553	1.477*	3.960	15.794*	6.43*
12QR12QE	384	0.630	2.840*	4.945	25.021*	1.19

see notes to Table 3