Ian Tonks
Xfi Centre for Finance and Investment, University of Exeter

Performance Persistence of Pension-Fund Managers*

Previous work on U.K. pension funds found only slight evidence of fund manager persistence but survivorship bias in the construction of these data samples 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 1-year time horizons but weaker evidence over longer horizons. Even when an allowance is made for momentum in stock returns, we find pension-fund managers exhibit performance persistence.

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% of U.K. corporate equity (Myners, 2001). A number of recent policy documents in the United Kingdom argued that pension contributions should be invested 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, using a large database on quarterly returns to U.K. pension funds, in which the fund management house managing the pension fund in each quarter is identified.[2]

* The data used in this study were provided by Russell-MellonCAPS, and we are grateful for the assistance of Alan Wilcock and Ian Ibbotson in collating the data. This paper has benefited from comments made by David Blake, Paul Draper, Alastair MacDougal, Michael Orszag, Alan Wilcock, and at seminars at the Universities of Bristol, Edinburgh, Exeter, Dundee, Lancaster, Southampton, St Andrews; Watson-Wyatt, Financial Markets Group, LSE, Said Business School, Oxford, and RTN Pensions Conference, Naples May 2003.

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

2. The term *fund manager* in this paper applies to a fund management house, rather than any individual fund manager employed by that fund management house.

(*Journal of Business,* 2005, vol. 78, no. 5)

0021-9398/2005/7805-0011$10.00

Occupational pension schemes in the United Kingdom are usually funded and require contributions throughout the employee's working life. In a funded scheme, an employee pays into a fund that 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, trustees, usually nominated by the employer, administer the fund; 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 United Kingdom. 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 outsource the management of the fund to an external fund management house. The pension funds in our sample are these segregated funded occupational pension schemes.

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 outsourced) to make the investment decisions on behalf of the fund according to some specified mandate and specific return expectation. The contract is usually on the basis of a rolling 3–5 year evaluation cycle, 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 (though of course this may be one of many managers). 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 eventual recipient of the pension, delegates pension fund decisions to the trustees, who in turn delegate the investment allocation decisions to a fund manager.

The objective of this paper is to analyze the existence of performance persistence of individual fund management houses that have been appointed as fund managers of segregated occupational pension funds. Is it possible for a pension fund trustee to identify fund management houses that 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 passive vehicle.

II. Previous evidence on performance of managed funds

The early literature of the performance of mutual funds in the United States (Jensen 1968) found that simple tests of abnormal performance did not yield significant out-performance. More recent work by Daniel et al. (1997), using normal portfolio analysis, shows that mutual fund managers, in particular those managing aggressive-growth funds, exhibit some selectivity ability but that funds exhibit no timing ability. For the United Kingdom, Blake and Timmermann (1998) examine the returns on 2300 U.K. open-ended mutual funds over 23-year period (1972–95) gross of fees. Over the period, the data include 973 dead and 1,402 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 underperformance 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 U.S. pension funds and found underperformance relative to the S&P500 Index. Lakonishok et al. (1992) provide evidence on the structure and performance of the money management industry in the United States 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 underperformed the S&P 500 by 1.3% per year throughout the 1980s. They emphasize that, although there is a long literature on the underperformance of mutual funds, pension funds also underperform relative to mutual funds on average. Coggin, Fabozzi, and Rahman (1993) investigate the investment performance of a random sample of 71 U.S. 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. 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 outperformance of 2.1% per annum but funds that adopted growth strategies underperformed by −0.96%.

Blake, Lehmann, and Timmermann, (1999) examine the asset allocations of a sample of 364 U.K. pension funds that retained the same fund manager over the period 1986–94. They find that the total return is dominated by asset allocation, with average return from stock selection negative and average return to market timing very negative. They find that U.K. fund managers are comparatively good at selecting equities, although only 16% of sample beat their 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 outperformance 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 investigated whether the outperformers 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) analyzed the short-term relative performance of no-load, growth-orientated mutual funds and found the strongest evidence for persistence in a 1-year evaluation horizon. Malkiel (1995), however, argues that survivorship bias is more critical than previous studies suggested.[3] 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 while 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 U.S. mutual funds and claim that the persistence is due mostly to funds that lag the S&P. They demonstrates that the relative performance pattern depends on the 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 the herding behavior identified in Grinblatt, Titman, and Wermers, (1995). They suggest that the market fails to discipline underperformers, and their presence in the sample contributes to the documented persistence. Carhart (1997b) demonstrates that common factors in stock returns (including a momentum factor) 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. Blake and Timmerman (1998); Allen and Tan (1999); Fletcher and Forbes (2002); and Giles, Wilsdon, and Worboys (2002) investigated

3. Malkiel (1995) 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 accrues from the tendency to run incubator funds (run 10 different products, see which are best and market those, ignoring the poor record of the rest).

performance persistence in U.K. mutual funds (unit trusts) over the last two decades and identified persistence caused by poor performers continuing to underperform.

Brown, Draper, and McKenzie (1997) examine the consistency of U.K. 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 database. They find that this limited consistency holds over different time horizons, samples, and classification schemes, although this finding seems to be influenced by the outperformance of one particular fund manager. Blake, Lehmann, and Timmermann (1999) also examine the persistence 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 U.K. equity portfolios at the 1-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).

In constructing their data samples, both the Brown et al. (1997) and Blake et al. (1999) studies of U.K. 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 database may have induced survivorship bias in these data samples, since a pension fund may continue to hire the same fund management house because its performance satisfies the pension fund trustees and does not trigger their removal. Survivorship bias can affect performance evaluation in two opposite ways. Brown et al. (1992) suggest that, if fund volatility is constant across time but varies cross-sectionally and the worst-performing funds in a period disappear, then survivorship induces spurious persistence and bias persistence upward. 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 et al. (1993) argue that, if fund survival depends on average performance over several periods, then survivorship induces spurious reversals: first-period losers must subsequently win to survive, and this biases persistence downward. Simulation results in Carpenter and Lynch (1999) suggest that persistence is weaker in samples that exhibit survivorship bias, implying that the second of the two survivorship bias effects dominates. Pension fund mandates are typically over a 3-year period, so that the survivorship criteria for U.K. pension fund managers is likely to depend on performance over several periods (Myners 2001, paragraphs 5.64–5.72). Previous studies that focused only on pension funds with the same single manager may have underestimated the true degree of persistence. In this study, we use data on all U.K. pension funds, irrespective of whether they change manager. As a consequence, we might expect that our sample will not suffer from the same extent of survivorship bias and 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 assess whether a fund-management house that has performed well in one period can repeat this feat in subsequent periods. Our database consists of the returns on pension funds managed by fund-management houses, and so to examine the consistency in performance of a specific fund-management house, we examine the performance across all the funds under management of that investment house. We do this by averaging abnormal returns across the funds under management to a particular fund manager. This averaging across pension funds is undertaken for houses both equally 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 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 single factor Capital Asset Pricing Model (CAPM), the Fama-French three-factor model, and a four-factor model that includes a momentum factor, 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:

$$AR_{Pt} = R_{Pt} - r_{ft} - \beta_P(R_{mt} - r_{ft}) - \gamma_P SMB_t - \lambda_P HML_t \qquad (1)$$

In the case of the CAPM, $\gamma_P = \lambda_P = 0$. 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 each quarter. The three-factor specification in equation (1) is the base specification in this paper, which we examine in some detail. We also reestimate abnormal returns from a four-factor model, where the fourth factor is a momentum factor, constructed as the returns on a zero investment portfolio based on a long-short position from a 1-year momentum strategy. Carhart (1997b) finds that a four-factor model can explain differences in the performance of past winners and past losers, with the momentum factor accounting for much of the explanation. He suggests that Jegadeesh and Titman's (1993) 1-year momentum in individual stock returns accounts for the persistence in mutual fund returns, since some mutual funds happen by chance to hold relatively large positions is last year's winning stocks. Wermers (1997) also finds that persistence in mutual fund performance is due to investing in past winners but suggests that this is due to an active momentum trading-rule strategy by the outperforming mutual funds. Hon and Tonks (2003) identify momentum at up to a 2-year horizon in U.K. stock returns, so that Carhart's findings could also apply to U.K. investment funds. We follow Carhart (1997b) and construct a 1-year return momentum factor as an additional factor for the U.K. stock market over the period 1984–97. To form this factor-mimicking portfolio, every month, we rank stocks listed on the London Stock Exchange from the Lspd (London Share Price Database) files on the basis of their previous 11-month returns lagged 1 month. We then compute an equally weighted portfolio return of the highest 30% of stocks minus the lowest 30% of stocks. The portfolios are reformed monthly throughout the sample period, and the quarterly return on this portfolio is compounded from the monthly returns.[4]

Performance-ranked portfolio tests sort fund managers 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 1 quarter, 4 quarters (1 year), or 12 quarters (3 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

4. We are grateful to Alan Gregory for providing the UK Momentum Factor used in Gregory and Tonks (2004).

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 next 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 periods can also be 1 quarter, 1 year, or 3 years. These procedures are followed for overlapping periods throughout the full period of the database, 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 centered on zero, which would mean that past performance is no predictor of future performance. From their simulations, Carpenter and Lynch find that the persistence test based on TDIF is best specified under the hypothesis of no persistence, and the most powerful of the alternatives considered.

In these persistence tests, we examine alternative ranking and evaluation time periods, since it may be that persistence is apparent only at particular time intervals. For example, to test for long-run persistence, 12QR12QE means we form portfolios on the basis of a 12-quarter ranking period and 12-quarter evaluation period. To test for short-run persistence, or the "hot-hands" phenomenon, we examine 1QR1QE, which means a 1-quarter ranking and 1-quarter evaluation period. We also test for asymmetric strategies, such as 12QR1QE, that allow for ranking on the basis of long-run (short-run) past returns and evaluation over short-run (long-run) future returns.

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: (1) Percentage of repeat winners, PRW = WW/(N/2) is a purely descriptive statistic that gives the percentage of the sample that are in the winner-winner box; (2) cross-product ratio CP = (WW × LL)/(WL × LW); which is also referred to as the odds-ratio and $\log(CP)/\sigma_{\log(CP)}$ has a standard normal distribution, with $\sigma_{\log(CP)} = \sqrt{(1/\mathrm{WW}) + (1/\mathrm{WL}) + (1/\mathrm{LW}) + (1/\mathrm{LL})}$, so that we may test for the statistical significance of deviations of the cross-product ratio from unity; (3) χ^2 test with 1 degree of freedon, where CHI $= \{(\mathrm{WW} - N/4)^2 + (\mathrm{WL} - N/4)^2 + (\mathrm{LW} - N/4)^2(\mathrm{LL} - N/4)^2\}/N/4$, and we may reject independence if CHI exceeds the critical value of 3.84 for a 5% test; and (4) TCS is the *t*-statistic for the slope coefficient in the cross-section ordinary least squares (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 χ^2 test performs best and the cross-sectional *t*-test (TCS) is weakest. We 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) advocate 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 use 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 to assess fund managers' private market-timing skills. Under a conditional version of the three-factor model, the abnormal returns equation (1) becomes

$$\mathrm{AR}_{Pt} = R_{Pt} - r_{ft} - \beta_P(Z_{t-1})(R_{mt} - r_{ft}) - \gamma_P \mathrm{SMB}_t - \lambda_P \mathrm{HML}_t \quad (2)$$

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

$$\beta_P(\mathrm{Z}_t) = b_0 + B'z_{t-1} \quad (3)$$

where $z_{t-1} = Z_{t-1} - E(Z)$ is a vector of deviations of Z from the 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, although 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 Combined Actuarial Performance Services Ltd (CAPS) provided the data used in this study. The data consist of quarterly returns on U.K. equity portfolios of 2,175 U.K. pension funds from March 1983 to December 1997. Typically, over this period, a U.K. pension fund invested about 57% of assets in U.K. equities, so that our database consists of returns on the major asset class in which U.K. pension funds invest. In

TABLE 1 Descriptive Statistics on Pension Funds and Fund Managers

Min	5%	25%	50%	75%	95%	Max
A. Distribution of Fund-Quarters (No. of Funds = 2,175, No. of Quarters = 59,509)						
1	4	12	24	41	56	56
B. Distribution of Manager-Quarters (No. of Managers = 191, No. of Quarters = 59,509)						
2	4	17	45	179	1,063	17,299

C. Distribution of Change in Managers across Funds (n = 2,175)

	Overall		Between Funds		Within Funds
	Frequency	%	Frequency	%	%
ΔMan code					
0	48,435	81.39	2,175	100	81.39
1	9,382	15.77	559	25.7	42.47
2	1,475	2.48	110	5.06	28.08
3	197	0.33	17	0.78	22.83
4	20	0.03	3	0.14	13.25
Total	59,509	100	2,864	131.68	71.33
D. Distribution of Managers across Funds by Category of Manager (n = 2,175)					
Fman category					
Multimanager	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-188RMan	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	141.33	63.43

NOTE.—In panel C, ΔMan = 0 denotes an observation with the manager the same as the previous manager by fund; ΔMan = 1 denotes an observation where the fund has changes its manager once up to that point in time, ΔMan = 2 denotes a fund that has changed its manager twice, and so forth.

In panel D, total within = (659 × 78.1 + 244 × 81.04 + . . .)/3,074; nRMan denotes nth ranked manager by frequency of observations and * denotes that each of these fund managers had greater than 1% of the overall frequency .

addition, for each fund quarter, the manager of the fund and the size of the fund are provided. CAPS provide a performance measurement service for about half of all segregated pension fund schemes in the United Kingdom. The full database 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 database and shows that 50% funds have 24 or fewer observations, and the average life of a fund in the data is just less than 7 years. This high attrition rate is of apparent concern, since it implies a possibility of survivorship biases, which bedevils performance evaluation studies of mutual funds. However, it is important to emphasize that the ease at which mutual funds may close down or merge poorly performing funds, documented in Elton, Gruber, and Blake (1996), is not so readily available to the firms sponsoring an occupational pension fund, since the obligation to provide a pension remains. Two major performance measurement services were used in the United Kingdom over the data sample, CAPS and WM, and UK pension funds typically subscribe to one or other of these services. According to CAPS, there are three reasons why funds enter or leave the database: (1) Structural reorganization, due to birth or death (including takeover or merger) of the sponsoring firm or mergers between pension funds within the same sponsoring firm; (2) change of fund management measurement service, due to client (pension-fund) preferences, perhaps due to dissatisfaction with the quality or price of the current service; (3) change of manager, possibly prompted by poor performance, with the new manager having a preference for one particular measurement service provider.

The second and third reasons might be related to poor performance of the pensions fund and hence could induce survivorship bias into our data set. However, it is important to emphasize that only two performance measurement services were in the United Kingdom over the data period, so that any switching between measurement services should be symmetric. A pension fund may drop out of our database because of poor performance, but there are new entrants into our database, as pension funds that have previously been assessed by the alternative measurement service switch into the CAPS measurement service. In effect, our sample may lose pension funds due to poor performance, but poor performers from the alternative measurement service replace them.

Following the suggestion of a referee, I compare the average prior 1-year performance of the funds leaving the data set with the average post 1-year performance of the funds entering the data set. There were 1,230 new funds entering the data set after the starting date, meaning that, of the 2,170 funds noted in table 2, 1,230 funds have their first observation after March 1984. There were 1,091 funds leaving the data set before the terminal date, meaning that 1,091 funds have their final observation before December 1997.[5]

5. Quarterly Returns to funds Entering and Exiting the Dataset June 1984–September 1997

	Observations	Mean	SD
Entering funds			
Post-Ret4	1,230	.0410	.0354
Post-Abret4	1,037	.0005	.0110
Exiting funds			
Pre-Ret4	1,091	.0369	.0353
Pre-Abret4	816	−.0002	.0101

Mean returns are quarterly returns based on quarterly average of the 12−month return pre and post the fund exiting or entering the data set.

TABLE 2 Descriptive Statistics

A. Returns across Quarters and Funds					
Returns, R_{Pt}	All	Weighted by Fund Size	≥12 Quarters	<12 Quarters	FT − All Share Returns
Mean	.0432	.0380	.0428	.0511	.0438
SD	.0867	.0814	.0867	.0872	.0834
Distribution of returns					
10%	−.0543	−.0537	−.0544	−.0503	
25%	.0016	.0016	.0015	.0048	
50%	.0463	.0441	.0459	.0528	
75%	.0896	.0747	.0885	.1061	
90%	.1525	.1346	.1527	.1457	
Observations	59,317	59,314	56,344	2,973	56
No. of funds	2,170	2,170	1,717	453	

B. Distribution of Fund Size across Funds (Size at Start of Quarter in £ Millions)			
	March 1983	December 1990	December 1997
Mean	25.02	50.24	102.27
SD	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
Observations	833	1131	1004

C. Returns and Fund Size across Funds and across Fund Managers

Variable	Mean	SD	Min	Max	Observations	
R_{Pt}						
Overall	.0432	.0867	−.5257	.8707	No. obs.	59,317.0
Between funds		.0165			No. funds	2,170
Within funds		.0863			Obs/fund	27.3
Between managers		.0177			No. managers	189
Within managers		.0866			Obs/manager	313.8
smv						
Overall	58.4	240.1	.0	9,108.6	No. obs	59,453.0
Between funds		174.4			No. funds	2,175
Within funds		89.6			Obs/fund	27.3
Between managers		506.0			No. managers	190
Within managers		164.8			Obs/manager	312.9
smv97						
Overall	204.2	869.1	.0	24,411.4	No. obs	59,437.0
Between funds		1,049.2			No. funds	2,170
Within funds		189.7			Obs/fund	27.4
Between managers		1,001.5			No. managers	189.0
Within managers		733.1			Obs/manager	314.5

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

It can be seen that the average quarterly return in the 12 months after a new fund has entered the database and the average quarterly return in the 12 months before a fund exits the data set are both below the quarterly returns across the entire data set reported in table 2 of 0.0432. The performance of the new funds are slightly higher than the performance of the existing funds, but the difference in the average quarterly abnormal returns (from the three-factor model, which requires a minimum of 12 time-series observations) of new and exiting funds is 0.0007. A simple *t*-test shows that this number is insignificantly different from zero. The implication is that funds leaving the database are replaced at roughly the same rate and in terms of the same returns as funds entering the database.

Carhart (1997a) distinguishes between survivorship bias and look-ahead bias: True survivorship bias is a property of the sample selection method and results from including in a sample only funds that survive until the end of the sample period. Look-ahead bias is a property of the test methodology. Tests of performance persistence require funds to exist during both the ranking and evaluation periods. Carhart defines look-ahead bias as the bias that results from eliminating funds from the sample that fail to survive a minimum period of time after the ranking period. He distinguishes between partial (PLA) and full (FLA) look-ahead bias. PLA ranks on the basis of all funds available in the ranking period and then eliminates funds that disappear before the end of the evaluation period. FLA eliminates disappearing funds before the ranking process starts.

In our study, we argue that true survivorship problems should be less of a concern, since we have the 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). By requiring at least 12 quarterly observations in the database, we impose an FLA methodology on our sample. According to the simulation results of Carpenter and Lynch (1999, tables 1 and 3), if elimination is based on a single-period criterion, then the PLA and FLA methodologies identify spurious persistence, whereas if elimination is based on a multiperiod criterion, the PLA and FLA methodologies understate persistence. In both cases, the PLA biases are more pronounced than the FLA. For these reasons, we adopted the FLA methodology in our sample construction, and by eliminating observations when there is less than 12 quarters worth of data, in many cases, we impose a multiperiod elimination criterion, which according to Carpenter and Lynch understates the true level of persistence.

Panel B of table 1 shows that the management of pension fund equity portfolios is relatively concentrated. With a total of 191 different fund manager categories (including in-house managers), 25% of fund managers manage only 17 quarters or less (across funds), and 50% manage across 45 quarters or less. Since the average life of a fund is just under 7 years (28 quarters), this implies that 50% of fund managers in the database manage only two funds.

Pension funds hire fund managers to manage the investment portfolio, and the funds may change fund managers within the database. Panel C reports the distribution of changes in managers across funds. Over 80% of the observations in the database have the same fund manager as in the previous quarter. From the middle pair of columns in panel C, we report that 559 funds, or just over 25% of the funds, change their manager at least once during the period, with three funds changing the manager four times. The final column in panel C shows the distribution of the change in fund manager within funds and reports that over 80% of the time periods on average across funds retain the same manager. These figures suggest that our database over the period 1984–97 has much less fund manager turnover than was reported in Myners (2001, paragraph 5.40), based on a survey in 2001. This may be because the Myners's *Report* noted the increased tendency in the late 1990s for large funds to appoint multiple external managers. We might conjecture that a fund that adopts a multiple-manager hiring strategy is more likely to change fund mangers.

Panel D of table 1 provides further evidence on the concentration of fund management. We rank the fund managers is terms of the number of fund quarters under management. The top-ranked fund manager[6] (1RMan) manages 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. The 1RMan manages across 244 funds, and 81.04% of these funds' observations are using 1RMan. There is also a multimanager 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 multimanager category, we have no information on the identity of the multiple managers and further the definition of multiple fund managers has changed over time. Only 85 funds use the same fund manager over the fund's life. The change of manager category sometimes applies to a fund when the fund manager has changed during the quarter, in which case if it is not possible to identify the fund manager managing the portfolio during that quarter, a change of manager category is noted.

Table 2 provides descriptive statistics on the returns to, and the size of, the U.K. equity portfolios of the pension funds in our database. From

6. This fund manager is actually identified as #28, in our database—all the fund managers are identified by a code.

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 emphasizes the variability in returns. But these pooled measures disguise an important statistic that is made clear in panel C, 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 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 behavior 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 may be difficult to identify individual fund manager performance. Our subsequent results on manager performance are all the more striking, given this feature of the data.

Table 2, panel A, also reports 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% indicates 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 estimated only when there were 12 or more quarterly returns for that fund. The cutoff value of 3 years accords with the typical fund mandate. Table 2, panel A, reports the distribution of returns of the subsample of 1,717 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 subsample is indeed representative. We also report the distribution of returns for those 453 funds with less than 12 observations that we drop from the database. In fact, the mean and median for this subsample is slightly higher than for the entire distribution and for the retained subsample. This suggests that we are not dropping poorly performing funds from the database. The standard deviations of both groups are also very similar, and it is slightly surprising that the standard deviation of the eliminated funds is no greater than that for the remaining funds, since both Carhart (1997a) and Carpenter and Lynch (1999) find the higher-volatility funds eliminated in their samples.

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 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, whereas the largest fund had an equity portfolio of over £9 billion. 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 nonconstant 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 emphasizes that pension fund size is relatively constant over time, with most variation occurring between funds.

In this study, we use data on all U.K. pension funds irrespective of whether they change manager. The Brown et al. (1997), and Blake et al. (1999) studies of U.K. pension funds specify that the pension fund have 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 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 are more likely to identify true persistence.

V. Results

The results of the persistence 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 the return to an arbitrage or zero-net-investment portfolio constructed by going long in the high-quintile portfolio (AV5) and short in the low-quintile portfolio (AV1). The 1-quarter on 1-quarter results (1QR1QE) shows only weak evidence

TABLE 3 **Persistence Tests Based on the Three-Factor Abnormal Returns of Fund-Manager Performance**

	AV5	AV1	DIF	TDIF		
A. Performance-Ranked Portfolio Tests: Unweighted Abnormal Returns						
1QR1QE	−.0002	−.0018	.0016	1.41		
4QR4QE	.0015*	−.0023*	.0039*	6.72		
12QR12QE	.0013*	−.0005	.0018*	3.10		
B. Performance-Ranked Portfolio Tests: Abnormal Returns Weighted by Fund Size						
1QR1QE	−.0004	−.0017	.0013	1.14		
4QR4QE	.0015*	−.0025*	.0041*	8.15		
12QR12QE	.0008	−.0005	.0013*	3.36		
	N	PRW	CP	Z-Statistic	CHI	TCS
C. Contingency Tables: Unweighted Abnormal Returns						
$1Q_{-t}1Q_t$	5,360	.523	1.139*	2.373	6.169*	.15
$4Q_{-t}4Q_t$	1,166	.578	1.531*	3.614	15.125*	6.17*
$12Q_{-t}12Q_t$	195	.574	1.249	.773	1.821	1.72
D. Contingency Tables: Abnormal Returns Weighted by Fund Size						
$1Q_{-t}1Q_t$	5,360	.520	1.112	1.936	4.282	−.29
$4Q_{-t}4Q_t$	1,166	.576	1.563*	3.788	15.520*	5.45*
$12Q_{-t}12Q_t$	195	.574	1.049	.167	2.149	1.23

NOTE.—For performance ranked tests, in panels A and B, 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 ranking and evaluation periods: 12QR12QE means a 3-year ranking period and 3-year evaluation period, and 1QR1QE means a 1-quarter ranking period and 1-quarter evaluation period. This procedure is followed for overlapping periods throughout the full period of the database, DIF is AV5–AV1, and TDIF is a t-statistic on DIF, allowing for the autocorrelation induced by using overlapping observations.

In panels C and D, fund managers are classified as winners or losers based on abnormal returns in each of two consecutive time periods Q_t and Q_{-t}; and the numbers of winner-winner (WW), winner-loser (WL), loser-winner (LW), and loser-loser (LL) are counted in each time period. There are three sets of consecutive periods: $12Q_{-t}12Q_t$ means the two consecutive periods are of 3-year length (12 quarters), similarly for $4Q_{-t}4Q_t$ and $1Q_{-t}1Q_t$. The following statistics are computed: (1) percentage of repeat winners, PRW = WW/(N/2); (2) cross-product ratio CP = (WW × LL)/(WL × LW), where $\log(\text{CP})/\sigma_{\log(\text{CP})}$ has a standard normal distribution and $\sigma_{\log(\text{CP})} = \sqrt{(1/\text{WW}) + (1/\text{WL}) + (1/\text{LW}) + (1/\text{LL})}$; (3) χ^2 test with 1 degree of freedom, where $\text{CHI} = [(\text{WW} - N/4)^2 + (\text{WL} - N/4)^2 + (\text{LW} - N/4)^2(\text{LL} - N/4)^2]/N/4$, and N is the number of pairs; and (4) 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.

* denotes significantly different from 0 at 95% confidence limits.

of persistence as measured by TDIF, although the longer-term abnormal returns show much stronger evidence of persistence. The results in panel A suggest that there is some persistence at all time horizons, with the strongest at 1 year. The evidence in this table is that, at the 1-year ranking and evaluation horizon, the difference in excess returns between the top and bottom quintiles averages about 1.56% per year. Further, in contrast to previous studies, which have suggested that any persistence is due to underperforming funds continuing to underperform, we find, although there is persistence in the bottom quintile, for the 4QR4QE strategy, the top-performing funds produce excess returns that are significantly above zero.

These findings are confirmed from the contingency table tests in panel C. The χ^2 test on independence is easily rejected for the 1-year horizon abnormal returns, and the odds ratio is also significantly different from unity at both the 1-quarter and 1-year horizons. Similarly, the *t*-statistic on the slope coefficient in the cross-section regression of 1-year abnormal returns on lagged 1-year abnormal returns is 6.17 for the 1-year abnormal returns, indicating significant persistence at 1 year. One slight inconsistency in these tables is that CHI implies that quarterly abnormal returns are more persistent than 3-year returns, whereas all the other measures suggest that the longer-term returns are more persistent than the shorter-term measures. Although the percentage of repeat winners is only 52% for the quarterly horizon, it rises to 57% for the longer horizons. Brown et al. (1992) suggest that some persistence 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 database. Recall that Carpenter and Lynch (1999) suggest that the TDIF measure is the most powerful from among the alternative tests for persistence. So, we should put more emphasis on the findings in panel A, which report persistence at the 1-year horizon. Blake et al. (1999) produce a statistic directly comparable with the data in panel A. They find that the return to the "zero net investment portfolio" for U.K. equities in the case of abnormal returns estimated from a multifactor model (similar to our three-factor model) for the year-on-year ranking and evaluation strategies yielded an annual return of 0.5%. This is substantially less than the annualized return of 1.56% implied by the equivalent strategy in panel A. We return to this comparison later.

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 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 pay 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, where the evidence suggests even stronger evidence of persistence at the 1-year and 3-year horizons: the

TABLE 4 **Persistence Tests Based on Alternative Measures of Abnormal Returns of Fund-Manager Performance**

	AV5	AV1	DIF	TDIF		
A. Performance-Ranked Portfolio Tests: Unweighted Four-Factor Abnormal Returns						
1QR1QE	−.0008	−.0019	.0012	1.14		
4QR4QE	.0012*	−.0025*	.0037*	6.45		
12QR12QE	.0011	−.0005	.0016*	3.18		
B. Performance-Ranked Portfolio Tests: Unweighted CAPM Abnormal Returns						
1QR1QE	.0002	−.0019	.0021	1.28		
4QR4QE	.0017*	−.0020	.0037*	3.07		
12QR12QE	−.0002	0.0008	−.0010	−1.76		
C. Performance-Ranked Portfolio Tests: Unweighted Three-Factor Conditional Abnormal Returns						
1QR1QE	.0006	−.0021*	.0027*	2.43		
4QR4QE	.0020*	−.0028*	.0048*	8.04		
12QR12QE	.0011	−.0002	.0013*	3.34		
	N	PRW	CP	*Z*-Statistic	CHI	TCS
D. Contingency Tables Based on Unweighted Four-Factor Abnormal Returns						
$1Q_{-t}1Q_t$	5,360	.515	1.079	1.331	2.123	−1.47
$4Q_{-t}4Q_t$	1,166	.568	1.507*	3.357	12.043*	4.18*
$12Q_{-t}12Q_t$	195	.590	1.320	.933	2.290	.94
E. Contingency Tables Based on Unweighted CAPM Abnormal Returns						
$1Q_{-t}1Q_t$	5,360	.526	1.164*	2.783	8.272*	4.19*
$4Q_{-t}4Q_t$	1,166	.587	1.700*	4.488	21.575*	6.37*
$12Q_{-t}12Q_t$	195	.523	.826	−.663	1.615	−3.28*
F. Contingency Tables Based on Unweighted Three-Factor Conditional Abnormal Returns						
$1Q_{-t}1Q_t$	4,980	.537	1.273*	4.245	18.593*	3.18*
$4Q_{-t}4Q_t$	1,088	.574	1.545*	3.565	13.919*	7.03*
$12Q_{-t}12Q_t$	183	.536	1.067	.218	.497	.84

NOTE.—See the note to table 3.

quarterly return on DIF is 0.41 percentage points. The contingency tests in panel D also report evidence of persistence at the 1-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 D, we investigate the effect of including a momentum factor as the fourth factor in the calculation of abnormal returns. The effect of introducing this momentum factor into the portfolio performance tests is to reduce the value of DIF slightly in all three cases in panel A. However, the value of DIF is still significantly positive for the 4-quarter and 12-quarter horizons. In the case of the contingency tables, panel D tells a similar story. The effect of introducing the momentum factor at the

1-quarter and 4-quarter horizons is to reduce the percentage of repeat winners, the cross-product ratio, the odds ratio, the χ^2 test for independence, and the coefficient of the cross-section regression. Hence, the momentum factor seems to explain a small amount of the observed persistence. However, the strong result that comes out of panels A and D is that, at the 4-quarter horizon, persistence is still a significant feature of the data set, even after allowing for momentum in individual stocks.

In panels B and E, 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 1-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 is insignificantly different from zero, and there is even some evidence of fund-manager performance reversals at the long horizon. The contingency tables in panel E for the CAPM measure also report significant persistence at the 1-year horizon and find evidence of persistence at the short horizon: the CP, CHI, and TCS statistics are all significant over successive 1-quarter periods. At the 12-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 C and F, we report the performance-ranked portfolio tests and the contingency tables derived from a three-factor model of conditional abnormal returns, where the conditioning variables include lagged macroeconomic variables. Following the Ferson and Schadt (1996) argument, by excluding lagged macroeconomic factors 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 C, this measure of fund-manager performance produces the significant consistency in performance at all horizons, with the strongest persistence at the 1-year period. The annualized value on DIF is 1.93%. The contingency tables in panel F also report persistence in the short run and medium term, though not in the longer term. The slightly stronger result on persistence in the case of the conditional benchmarks implies that the unconditional benchmarks disguise fund managers' true abilities, although to only a minor extent.

In table 5, we also examine the performance-ranked portfolio tests based on nonsymmetric ranking and evaluation periods. We undertake this test for the abnormal returns derived from the three-factor model; therefore, the results should be compared with those in table 3, panel A. Indeed, the elements in the leading diagonal of table 5 are the same as in the DIF column for table 3, panel A. It can be seen that there is significant evidence of persistence at all combinations of ranking and evaluation periods except for the very short term. The value of DIF appears to be maximized on the basis of a 12-quarter ranking and a 1-quarter

TABLE 5 Asymmetric Persistence Tests Based on Three-factor Abnormal Returns of Fund-Manager Performance (Performance-Ranked Portfolio Tests: Unweighted Abnormal Returns)

	Evaluation Period		
Ranking Period	1QE	4QE	12QE
1QR	.0016	.0024*	.0012*
	(1.41)	(4.80)	(5.07)
4QR	.0040*	.0039*	.0018*
	(4.75)	(6.72)	(3.15)
12QR	.0042*	.0041*	.0018*
	(6.52)	(5.66)	(3.10)

NOTE.—As in table 3, fund managers are sorted each year into quintile portfolios based on past performance of the pension funds under management. Each element in the table is DIF–Difference in the equally weighted average portfolio abnormal returns of the top and bottom quintile portfolios over the subsequent evaluation period. There are nine combinations of ranking and evaluation periods; for example, 4QR12QE means a 4-quarter ranking period and 12-quarter evaluation period. The term in parentheses is a *t*-statistic on DIF, allowing for the autocorrelation induced by using overlapping observations.

* denotes significance at the 95% level.

evaluation period. This yields annual returns on an arbitrage portfolio of 1.68%.

Finally, we wish to compare our results with the previous studies in the United Kingdom 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 already mentioned, both studies concentrated on pension funds that employed the same fund manager over a long time period and we suggest 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 6, we report the results of imposing two data restrictions on our database: we consider only those pension funds that have remained in the data file for all 56 quarters from March 1984 to December 1997 and, in addition, 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 database is longer in our study. In table 5, we report the results of including these restrictions into our database 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 1-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 database reduces the observed degree of persistence

TABLE 6 Persistence Tests Based on Three-Factor Abnormal Returns of Fund-Manager Performance for a Restricted Sample of Single Same-Manager Funds

A. Performance-Ranked Portfolio Tests of Fund-Manager Performance

	AV5	AV1	DIF	TDIF
1QR1QE	.0018	.0001	.0017	2.11
4QR4QE	.0028*	.0005	.0023*	3.35
12QR12QE	.0038*	.0019	.0019*	4.30

B. Contingency Tables of Fund-Manager Performance

	N	PRW	CP	*Z*-Statistic	CHI	TCS
$1Q_{-t}1Q_t$	7,060	.522	1.159*	3.093	9.738*	6.63*
$4Q_{-t}4Q_t$	1,668	.553	1.477*	3.960	15.794*	6.43*
$12Q_{-t}12Q_t$	384	.630	2.840*	4.945	25.021*	1.19

Note.—See the note to table 3.

at the 1-year horizon. Interestingly, the performance-ranked portfolio tests for the 1-quarter and 12-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 6, 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

With the advent of low cost stakeholder pensions in the United Kingdom, 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 a role for active fund management of pension funds.

We measured the abnormal returns generated by fund-management houses in managing the equity portfolios of U.K. pension funds over the period 1983–97. We found evidence of significant persistence in the performance of fund managers at the 1-year time horizon using a number of different consistency tests, as well as weaker evidence of persistence at longer 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 annualized abnormal return of 1.56%. According to Carpenter and Lynch (1999), this test is the most powerful in detecting persistence in performance. Further, we examined how robust these results are to the inclusion of a momentum factor. In contrast to Carhart (1997b) and Wermers (1997), we find that, although a momentum factor can explain a small degree of the observed persistence, after extracting the momentum factor, there is still significant persistence in the performance of pension-fund managers. The returns on a zero-investment portfolio net of momentum effects would have yielded 1.48% per annum. This is an important result and conflicts with the view presented in the Myners's *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). How can we explain this gradual erosion of persistence at longer time horizons? As in note 2, the underlying unit of analysis in this paper is the fund-management house, not individual fund managers. It may be that the fund-management skills we identified are due to individuals, but over time these individuals move between jobs, so that over longer horizons, the persistence in fund-management-house performance weakens.

We 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 database 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 made no allowance for the costs of fund management. We 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 explicitly 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 and the upper quartile of management fees are 48 basis points. In both cases, the excess return of 148 basis points that we identified clearly covers the fund management fees. Second, having identified performance persistence at the 1-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

Allen, D. E., and M. L. Tan. 1999. A test of the persistence in the performance of UK managed funds. *Journal of Business Finance and Accounting* 26, nos. 5–6 (June–July): 559–93.

Blake, D., B. Lehmann, and A. Timmermann. 1997. *Performance measurement using multi-asset portfolio data: A study of UK pension funds 1986–994.* Pensions Institute, Birkbeck College.

———. 1999. Asset allocation dynamics and pension fund performance. *Journal of Business* 72:429–61.

Blake, D., and A. Timmermann. 1998. The birth and death processes of mutual funds *European Finance Review* 2:57–77.

Brown, G., P. Draper, and E. McKenzie. 1997. Consistency of UK pension fund performance. *Journal of Business Finance and Accounting* 24 (March): 155–78.

Brown, S., and W. Goetzmann. 1995. Performance persistence. *Journal of Finance* 50: 679–98.

Brown, S., W. Goetzmann, R. Ibbotson, and S. Ross. 1992. Survivorship bias in performance studies. *Review of Financial Studies* 5:553–80.

Carpenter, J. N., and A. W. Lynch. 1999. Survivorship bias and attrition effects in measures of performance persistence. *Journal of Financial Economics* 54:337–74.

Carhart, M. 1997a. Survivorship bias and mutual fund performance. Unpublished paper, University of Southern California.

———. 1997b. On persistence in mutual fund performance. *Journal of Finance* 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* 48:1039–55.

Consumers' Association. 1997. *A blueprint for better pensions.* London: Consumers' Association.

Daniel, K., M. Grinblatt, S. Titmanm and R. Wermers. 1997. Measuring mutual fund performance with characteristic based benchmarks. *Journal of Finance* 52:1035–58.

Department of Social Security. 1998. *A new contract for welfare: Partnership in pensions*. London: HM Stationery Office.

Elton, E. J., M. J. Gruber, and C. R. Blake. 1996. Survivorship bias and mutual fund performance. *Review of Financial Studies* 9, no. 4:1097–1120

Ferson, W. E. and R. Schadt. 1996. Measuring fund strategy and performance in changing economic decisions. *Journal of Finance*, 51:425–62.

Financial Services Agency. 1999. *Comparative information for financial services.* London: Financial Services Agency.

Fletcher, J., and D. Forbes. 2002. An exploration of the persistence of UK unit trust performance. *Journal of Empirical Finance* 9, no. 5, 475–96.

Giles, T., T. Wilsdon, and T. Worboys. 2002. Performance persistence in UK equity funds: An empirical analysis. Report No. D03374-00, London: Charles River Associates.

Gregory, A., and I. Tonks. 2004. Performance of personal pension schemes in the UK. Discussion paper no. 22 FMG/UBS Series. London School of Economics.

Grinblatt, M., and S. Titman. 1992. Persistence in mutual fund performance. *Journal of Finance* 47:1977–84.

Grinblatt, M., S. Titman, and R. Wermers. 1995. Momentum investment strategies, portfolio performance and herding: A study of mutual fund behaviour. *American Economic Review* 85:1088–1105.

Hendricks, D., J. Patel, and R. Zeckhauser. 1993. Hot hands in mutual funds: Short run persistence of relative performance, 1974–1988. *Journal of Finance* 48:93–130.

Hon, M., and I. Tonks. Momentum in the UK stock market. *Journal of Multinational Financial Management* 13, no. 1 (February 2003): 43–70.

Ippolito, R. A., and J. A. Turner. 1987. Turnover, fees and pension plan performance. *Financial Analysts Journal* 43:16–26.

Jegadeesh, N., and S. Titman. 1993. Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance* 48:65–91.

Jensen, M. C. 1968. The performance of mutual funds in the period 1945–1964. *Journal of Finance* 23:389–416.

Lakonishok, J. A., A. Shleifer, and R. W. Vishny. 1992. The structure and performance of the money management industry. *Brookings Papers on Economic Activity*: 339–91.

Malkiel, B. G. 1995. Returns from investing in equity mutual funds 1971 to 1991. *Journal of Finance* 50:549–72.

Myners, P. 2001. *Report on institutional investment.* London: H.M. Treasury.

Occupational Pensions Regulatory Authority. 2001. *Pensions registry.* London: Occupational Pensions Regulatory Authority.

Office of Fair Trading. 1997. *Report of the director general's inquiry into pensions.* London: Office of Fair Trading.

Thomas, A., and I. Tonks. 2001. Equity performance of segregated pension funds in the UK. *Journal of Asset Management* 1, no. 4:321–43

Wermers, R. 1997. Momentum investment strategies of mutual funds, performance persistence and survivorship bias. Working paper, University of Colorado at Boulder.