Pensions Institute

DISCUSSION PAPER PI-9904

The Hazards of Mutual Fund Underperformance: A Cox Regression Analysis

Asger Lunde, Allan Timmermann and David Blake

1999

ISSN 1367-580X

The Pensions Institute
Cass Business School
City University
106 Bunhill Row London
EC1Y 8TZ
UNITED KINGDOM

http://www.pensions-institute.org/

ELSEVIER

Journal of Empirical Finance 6 (1999) 121–152

Journal of EMPIRICAL FINANCE

The hazards of mutual fund underperformance: A Cox regression analysis

Asger Lunde [a], Allan Timmermann [b,c,*], David Blake [d]

[a] Institute of Economics, University of Aarhus, Aarhus, Denmark
[b] London School of Economics, Financial Markets Group, Department of Economics, Houghton Street, London, WC2A 2AE, UK
[c] University of California, San Diego, CA 92093, USA
[d] Economics Department, Birkbeck College, University of London, London, W1P 2LL, UK

Abstract

This paper investigates the process determining mutual funds' conditional probability of closure, i.e., their hazard function. Using a nonparametric approach to estimate the effects of a fund's age on its hazard rate, we find a distinctly non-linear, inverse U-shaped pattern in the relationship. Hence, young and very old funds are least likely to be closed down. A fund's relative performance and (less significantly) the level of return in the sector in which the fund operates are also identified as important factors in the closure decision. Results from semiparametric Cox regressions are compared with those from the discrete choice probit model used by Brown and Goetzmann [Brown, S.J., Goetzmann, W., 1995. Performance persistence, Journal of Finance. Vol. 50, pp. 679–698]. Finally, we provide a complete summary of the fund attrition process by estimating the survivor function, indicating the proportion of funds that survive up to a given age, and we identify the effect of fund attrition on standard measures of persistence of fund performance.

JEL classification: C14; G14; G23

Keywords: Mutual funds; Duration analysis; Cox regression; Discrete choice model

* Corresponding author. Tel.: +44-171-955-6394; Fax: +44-171-955-7420; E-mail: a.timmermann@lse.ac.uk

PII: S0927-5398(98)00013-9

1. Introduction

Using a large panel of mutual funds, this paper identifies and quantifies the significance of a range of factors influencing the process governing the rate at which funds are terminated. We have access to a new data set containing monthly return records on an almost complete sample of the unit trusts (open-ended mutual funds) that were operating in the UK during the period 1972–1995. The sample is unique both in terms of its length (spanning 23 years or 281 months) and in terms of the number of dead and surviving funds included (973 and 1402, respectively). The sample is long enough for us to measure the complete history of a large number of funds and so enables us to address a range of issues that have not previously been analyzed, such as the duration dependence of fund closures, as indicated by the shape of the mutual fund hazard function, hazard rate dependence on common and fund specific factors, and the fund survivor function.

Investigating the process governing mutual fund attrition rates is important for a number of reasons. Recent studies of mutual fund performance have found a sizeable survivorship bias associated with the underperformance of funds that have been closed. [1] A better understanding of mutual fund performance thus requires a thorough analysis of the factors determining the size and nature of survivorship bias. The average lifetime of a fund and the relationship between a fund's abnormal performance and its probability of being closed affects the size of the survivorship bias, and we shed new light on both these factors. Fund attrition is also likely to affect the estimated persistence of mutual fund performance to the extent that the funds that are closed down are also the ones that display the poorest historical track record. Second, measuring the duration profile of mutual funds is important for understanding the incentives under which fund managers operate. Fund management groups appear to have an incentive to close down poorly performing funds which otherwise reduce the average performance of the funds in their stable. If these funds are generally closed down after only a very short period, then fund managers can be expected to be under significant pressure to perform in the short run. This might give the fund managers a strong incentive to follow 'short-termist' investment strategies. [2] Third, assuming that funds are most likely to close down as a result of loss of investor interest arising from poor performance, the termination process might also be informative of the strategies pursued by investors. For example, if persistent relative underperformance does not strongly affect the likelihood of closure, a case could be made that investors do not pursue rational strategies. Finally, the time it takes to close down funds also

[1] See, e.g., Grinblatt and Titman (1989), Brown and Goetzmann (1995), Malkiel (1995) and Gruber (1996).

[2] Brown et al. (1996) argue that the practice in the US portfolio management industry of rewarding fund managers based on their annual relative performance generates a short-term perspective in the managers' objective function.

sheds light on the information extraction process that investors are confronted with when attempting to determine a fund's performance. If poorly performing funds tend to close down after only a short period, this suggests that investors possess good information about the fund performance, whereas a longer average time before closure might indicate that performance signals are weak, so that investors need more time before they can identify a fund's genuine performance.

The investigation of mutual fund attrition processes is a fairly recent phenomenon. Brown and Goetzmann (1995) examine a sample of US equity mutual funds by estimating a probit function for the event that a fund dies. We model nonparametrically the impact of fund age on the hazard rate and discover an interesting inverse U-shaped pattern to this relationship: young and very old funds seem to have a lower risk of closure. This result differs from the parametric approach used by Brown and Goetzmann (1995) which models the probability of a fund's closure as a linear function of the fund's age. There is nevertheless a close relationship between the two approaches: when we estimate a discrete choice model with a piecewise linear effect of fund age on the hazard rate, we can recover the inverse U-shaped pattern from the Cox regression model and the estimated coefficients on the covariates converge to their values from the Cox regression model as the number of age dummies increases.

A more complete picture of the fund attrition process is provided by our analysis of the survivor function which measures the proportion of funds dying as a function of age. Mutual funds have a relatively long average lifetime and precise estimation of the hazard function requires a data sample as long as ours.

The paper proceeds as follows. Section 2 describes the data set. Sections 3 and 4 analyze the proportional hazard and discrete choice models. Section 5 summarizes the estimated survivor function, while Section 6 analyses the relationship between fund attrition and persistence of performance. Section 7 concludes.

2. Description of the data

Our data set consists of monthly return records on a sample of unit trusts that were in existence in the UK at some time during the period February 1972–June 1995, a total of 281 months. In the UK, unit trusts allow individuals and companies to buy an easily realizable stake in a diversified portfolio of marketable securities that is managed by a professional fund management group. Monthly returns on these funds are calculated using bid prices and net income and hence, do not include transaction costs or management fees. The data set, which was provided by Micropal, contains the return records of 973 funds that died [3] before

[3] A fund's death refers to the event that the fund ceases to exist in its previous form. Hence, a fund can die either as a result of a name change, a merger with another fund, or a wind up.

Table 1
Births and deaths of funds (1972–1995)

Year	Funds born during year	Birth rate (%)	Funds dying during year	Death rate (%)	Funds alive at end of year
< 1972	46				46
1972[a]	35	76.1	NA	NA	81
1973	169	208.6	NA	NA	250
1974	40	16.0	NA	NA	290
1975	41	14.1	NA	NA	331
1976	47	14.2	NA	NA	378
1977	35	9.3	2	0.5	411
1978	33	8.0	10	2.4	434
1979	52	12.0	9	2.1	477
1980	55	11.5	17	3.6	515
1981	75	14.6	24	4.7	566
1982	62	11.0	29	5.1	599
1983	103	17.2	18	3.0	684
1984	138	20.2	24	3.5	798
1985	174	21.8	21	2.6	951
1986	176	18.5	21	2.2	1106
1987	205	18.5	48	4.3	1263
1988	163	12.9	57	4.5	1369
1989	165	12.1	75	5.5	1459
1990	139	9.5	91	6.2	1507
1991	71	4.7	130	8.6	1448
1992	98	6.8	90	6.2	1456
1993	118	8.1	118	8.1	1456
1994	87	6.0	109	7.5	1434
1995[b]	48	3.3	80	5.6	1402

The birth and death rates were computed as the number of funds that were born or died during a particular calendar year divided by the number of funds in existence at the end of the previous calendar year.
[a] Data for 1972 excludes January.
[b] Data for 1995 includes January to June.

the end of the sample and 1402 funds that survived until the end of the sample. [4] According to Micropal, the data set is an almost complete record of all authorized unit trusts that were in existence during the sample period under investigation.

Table 1 provides some descriptive statistics for the total number of fund births and deaths during the calendar years spanned by the sample period. A large number of funds started up in 1973 and also during the mid-to-late-eighties on the back of stock market booms (column 2). Since the number of funds in existence has increased over time, a more suitable measure of industry activity is provided by the fund birth rate, calculated as the number of funds born during a given

[4] We do not, however, have data on other fund attributes such as size or management fee.

Table 2
Sectorial composition of the data set by survivorship status

Sectoral	Surviving funds	Non-surviving funds
UK equity growth	144	95
UK equity general	111	61
UK equity income	111	112
UK smaller companies	74	30
UK gilt and fixed interest	53	52
UK balanced	58	18
Financial and property	11	15
Investment trust units	13	8
Commodity and energy	13	48
International equity growth	74	80
International equity income	9	19
International fixed interest	37	4
International balanced	36	21
Fund of funds	68	18
North America	127	92
Europe	121	74
Japan	88	22
Far East including Japan	39	28
Far East excluding Japan	70	12
Australasia	4	21

For each sector, columns two and three give the number of surviving and non-surviving funds contained in the sample.

calendar year divided by the number of funds in existence at the end of the previous calendar year, shown in the last column of the table (see column 3). The extremely high birth rates at the beginning of the sample can be attributed to the small number of funds in existence at that time. While mutual funds have been around since the 1930s, the industry did not really take off until the 1970s. Subsequently, the birth rate stabilized. There are some significant variations, however: at a level close to 20% per year, birth rates were high during the bull markets of the mid-eighties but then fell to around 12% after the October 1987 stock market crash, and to a level well below 10% during the early 1990s. [5]

Our data set contains records on dead funds from 1977 onwards. The death rate, reported in column 5, is calculated as the number of funds that died during a given calendar year divided by the total number of funds in existence at the end of the previous calendar year. Death rates appear to be inversely correlated with birth rates, being low in the early seventies and early-to-mid-eighties only to increase

[5] There are strong similarities in the evolutions of the mutual fund industries in the UK and the US. High growth rates in the number of US equity mutual funds during the mid-eighties were also reported by Brown and Goetzmann (1995).

systematically after the October 1987 crash. Authorized unit trusts are allocated to one of 20 sectors specified by the Association of Unit Trusts and Investment Funds (AUTIF) and shown in Table 2. To be allocated to a particular sector, a trust must have at least 80% of its assets invested in that sector. [6]

For each sector, Table 2 shows the number of funds that survived until the end of the sample (column 2) or died within the sample (column 3). There are large numbers of funds in the domestic equity sector which has been split into the subcategories of growth, general, income, and smaller company funds. Likewise, there are many funds in international equity sectors such as international equity growth, North America, Europe, Japan and the Far East. In contrast, there are few funds in some of the more specialized sectors such as Australasia or investment trust units.

3. Hazard functions and durations

Fig. 1 presents a histogram of the age distribution of the 973 funds that died during the sample. Two funds died within the same month they were launched, and a total of 21 funds died within 6 months of inception. More than 70% of funds that died did so within a period of 3 to 15 years after launch. Because of the inevitable right-censoring of our sample, which ends in June 1995, these numbers are difficult to interpret, particularly at the long end of the distribution where censoring effects will be strongest. Nevertheless, the figure suggests that most funds are not closed down shortly after their inception. Fund management groups appear to give their new funds quite some time to establish track records before deciding on their future.

3.1. Nonparametric estimation of the hazard function

To analyze the duration of the funds included in our sample and in order to account formally for the right-censoring of the data, we utilized methods from the literature on economic duration data (see, e.g., Kalbfleisch and Prentice (1980), Kiefer (1988), Pudney (1989), Lancaster (1990), and, in particular, Han and Hausman (1990) and Tunali and Pritchett (1997) on Cox regressions). Let T be a random variable measuring the duration or age of a particular fund with a probability distribution $F(t) = \Pr(T < t)$ and a density function $f(t) = \mathrm{d}F(t)/\mathrm{d}t$. We will also be interested in estimating the survivor function, the probability that a fund has survived beyond a certain time horizon, defined as $S(t) = 1 - F(t) =$

[6] There are other conditions. For example, an income trust must invest in securities with yields exceeding 110% of the yield on the relevant index, while an equity trust can also have some of its portfolio invested in government bonds.

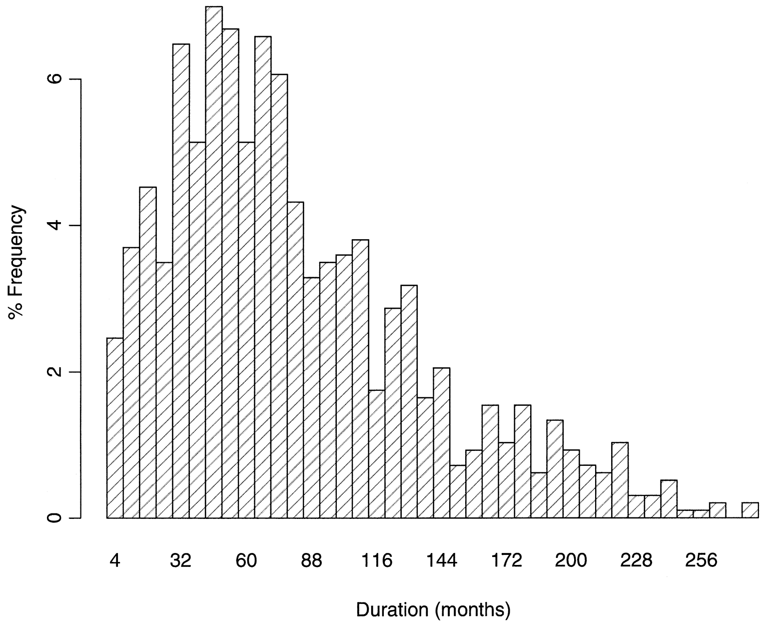


Fig. 1. The histogram shows the age distribution for the 973 funds that died during the course of the sample (1972–1995). The figures on the horizontal axis are the midpoints of the corresponding column.

$\Pr(T \geq t)$. Our duration data can be characterized in terms of the hazard function, that is, the conditional probability that the fund dies in a short time interval following period t, given that it survived up to period t:

$$\lambda(t) = f(t)/S(t). \tag{1}$$

Hypotheses concerning the probability that a fund is terminated as a function of the fund's age are naturally expressed in terms of the shape of this hazard function.

Initially, we estimated the hazard and survivor functions nonparametrically using the Kaplan–Meier (product–limit) estimator, which accounts for the right-censoring of the data set. Let $t_1, t_2, \dots, t_i, \dots, t_n$ denote either death or censoring times of the n funds in our sample. Suppose we order the discrete death times as $t_{(1)} < t_{(2)} <, \dots, < t_{(r)}$, and let h_j be the number of funds that died after $t_{(j)}$ months, while m_j is the number of funds censored between months $t_{(j)}$ and $t_{(j+1)}$. If $n_j = \sum_{i \geq t_{(j)}}^{t_{(r)}} (m_i + h_i)$, then the nonparametric Kaplan–Meier estimator for the funds' hazard rate per unit of time (month) is given by

$$\hat{\lambda}(t) = \frac{h_j}{\tau_j n_j}, \tag{2}$$

for $t_{(j)} \leq t < t_{(j+1)}$, where $\tau_j = t_{(j+1)} - t_{(j)}$. Eq. (2) provides an unconditional estimate of the funds' hazard rate at time $t_{(j)}$. A smoothed plot of Eq. (2) (with a

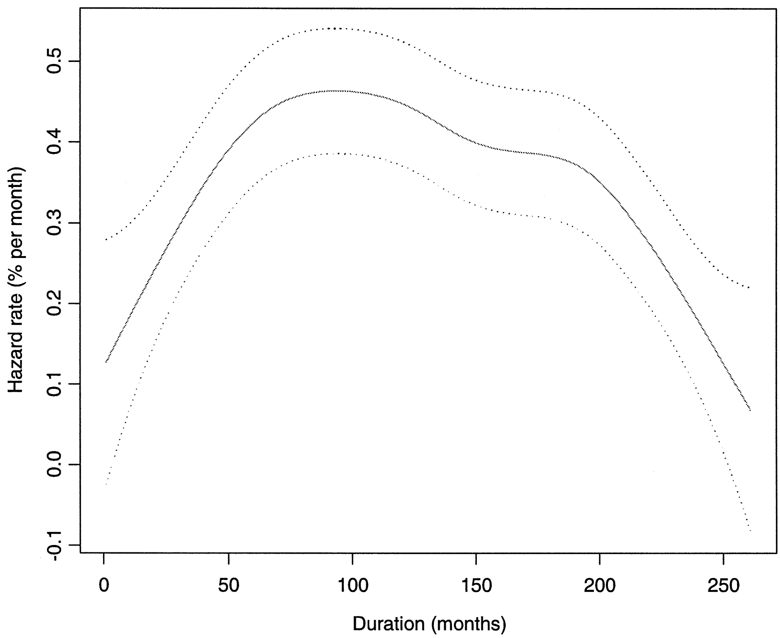


Fig. 2. This figure shows the smoothed hazard function for the UK equities total sector based on the Kaplan–Meier estimates. The 95% confidence interval for the smoothing spline was calculated from jackknife residuals.

95% jackknife confidence interval) [7] in the case of UK equity funds is presented in Fig. 2. [8] We use this set of funds as a basis for verifying that our findings are comparable with the results in subsequent sections, as well as with the vast literature on the performance of equity funds.

The hazard rate increases between months 1 and 80, peaks after 8 years or so and decreases gradually thereafter, most rapidly after 15 years. Thus, the hazard rate is smallest for young and very old funds (above 15 years) and it is reasonably well approximated by an inverse U-shaped function. At its peak, the hazard rate is around 0.5% per month or 6% annualized.

[7] A common problem when using standardized residuals to construct confidence intervals is their sensitivity to outliers. Jackknife residuals calculate the ith residual from the model fitted without observation i and hence provide a more robust method for computing standardized residuals.

[8] The estimated hazard function, $\hat{\lambda}(t)$, displays a high degree of monthly variability around a pronounced inverse U-shape. This variability is a common finding for discrete data such as ours and reflects the relatively small number of duration spells ending after a given number of months. To aid the visual interpretation of the data, we used a cubic spline to smooth the point estimates of the hazard function.

3.2. Semiparametric estimation of the hazard function

As noted in Section 2, the birth and death processes of individual funds do not appear to be independently distributed across funds. In particular, the death process appears to depend on the past level of returns within a given sector, high past returns seemingly associated with fewer deaths. Likewise, if, as has been found in a number of studies, [9] investors re-allocate money away from the poorest performing funds, it seems plausible that a fund's abnormal returns should influence the decision to close down the fund. To account for such effects we allow the fund hazard rate at a particular point in time to depend on the realization of a set of common and fund-specific time-varying covariates, with $x_{ij}(t)$ denoting the ith fund's jth covariate at time t.

We do not have any strong a priori reasons for imposing a particular functional form for the dependence of a fund's hazard rate on its age, and thus prefer to model this particular relationship nonparametrically. In contrast, the effects on the hazard rate of the time-varying covariates are modelled parametrically and the specification is checked against a variety of residual tests reported in the appendix. This leads to the semiparametric Cox regression model:

$$\lambda_i(t) = \lambda_0(t)\exp\left(\boldsymbol{\beta}'\mathbf{x}_i(t)\right), \tag{3}$$

where we have represented the ith fund's covariates by the column vector $\mathbf{x}_i(t) = (x_{i1}(t), \dots, x_{ip}(t))'$. Cox (1972) showed that the two components of Eq. (3) can be estimated separately in a two-step procedure where, first, $\boldsymbol{\beta}$ is estimated by maximum likelihood, and, second, the baseline hazard function, $\lambda_0(t)$ is estimated nonparametrically.

As a first step in the analysis, we transformed the sample data from calendar time to age of fund. For example, for a fund born in January 1983, we count this month as month 1, February 1983 will be month 2, etc. The funds are now aligned according to their age, so we can identify separately the effects on a fund's hazard rate of the time-varying covariates and the fund's age. Hence, we can determine the impact on the hazard rate of, say, a 5% underperformance by a 10-year-old fund. [10]

Suppose that data are available on n funds, among which there are r distinct death times and $n - r$ right-censored survival times. Also, assume that all funds have distinct termination and censoring times as would be the case if the processes were measured on a continuous scale. This assumption amounts to ruling out ties in the death and censoring processes. Recalling that $t_{(j)}$ denotes the jth ordered

[9] See, e.g., Berkowitz and Kotowitz (1993), Brown and Goetzmann (1995) and Gruber (1996).

[10] Although Eq. (3) implies that a certain level of underperformance (as measured by $x_i(t)$) has the same proportional effect on funds irrespective of their age, the effect on the absolute level of the hazard rate will vary. If, say, $\lambda_0(t)$ is highest for 5–10-year-old funds, then a given underperformance will increase the hazard rate the most for these funds.

death time, let $\mathscr{R}(t_{(j)})$ denote the *risk set at time* $t_{(j)}$, i.e., the set of all funds that are alive and uncensored just prior to $t_{(j)}$. The basic components of the likelihood function, $\mathscr{L}(\boldsymbol{\beta})$, are terms of the form

$$\frac{\Pr\left\{\text{fund with covariates } \mathbf{x}_i\left(t_{(j)}\right) \text{ dies at } t_{(j)}\right\}}{\Pr\left\{\text{one death at } t_{(j)}\right\}} = \frac{\lambda_i\left(t_{(j)}\right)}{\sum_{l \in \mathscr{R}(t_{(j)})} \lambda_l\left(t_{(j)}\right)}. \tag{4}$$

Because of the proportionality of $\lambda_i(t_{(j)})$ and $\lambda_0(t_{(j)})$, the baseline hazard will cancel out in this expression and Kalbfleisch and Prentice (1980) showed that the relevant 'likelihood function' for Eq. (3) is given by [11]

$$\mathscr{L}(\boldsymbol{\beta}) = \prod_{j=1}^{r} \frac{\exp\left(\boldsymbol{\beta}' \mathbf{x}_j\left(t_{(j)}\right)\right)}{\sum_{l \in \mathscr{R}(t_{(j)})} \exp\left(\boldsymbol{\beta}' \mathbf{x}_l\left(t_{(j)}\right)\right)}. \tag{5}$$

The summation in the denominator of Eq. (5) is over all funds that are still alive and hence, included in the risk set at time $t_{(j)}$. [12] Using a censoring indicator, δ_i, which equals zero if the ith fund is right-censored, and unity otherwise, we can write the log-likelihood function as follows: [13]

$$\log \mathscr{L}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \delta_i \left\{ \boldsymbol{\beta}' \mathbf{x}_i(t_i) - \log \sum_{l \in \mathscr{R}(t_i)} \exp\left(\boldsymbol{\beta}' x_l(t_i)\right) \right\}. \tag{6}$$

Maximum likelihood estimates of $\boldsymbol{\beta}$ can be found by numerical methods.

The assumption of no ties is not, in practice, satisfied, since our sample only measures the month in which a fund is terminated. To account for ties, the above set-up needs to be slightly modified. Suppose that there are d_j deaths at time $t_{(j)}$ and define a vector of sums of the covariates relating to the funds that die at time $t_{(j)}$:

$$\mathbf{s}\left(t_{(j)}\right) = \left(\sum_{k=1}^{d_j} x_{k1}\left(t_{(j)}\right), \sum_{k=1}^{d_j} x_{k2}\left(t_{(j)}\right), \ldots, \sum_{k=1}^{d_j} x_{kp}\left(t_{(j)}\right) \right), \tag{7}$$

[11] Since this likelihood function does not use all information on the actual censored and uncensored survival times, it is not a true likelihood function. Cox (1975) called it a *partial likelihood function* and, like a marginal likelihood, its purpose is to allow inference in the presence of nuisance parameters (in this case, the parameters determining the shape of the baseline hazard).

[12] Notice that the product in Eq. (5) is taken over the funds with recorded death times. While censored survival times do not contribute to the numerator of the likelihood function, they enter through the summation over the risk sets at death times that occur before a censoring time.

[13] Recall that t_i denotes the death or censoring time of the ith fund, whereas we defined $t_{(j)}$ as the jth ordered death time.

where $x_{kh}(t_{(j)})$ is the value of the kth fund's hth covariate, for some k belonging to the set of d_j funds that die on the jth termination date. Using these sums, Breslow (1974) proposes the following approximation to the likelihood (5): [14]

$$\log \mathscr{L}(\boldsymbol{\beta}) = \sum_{j=1}^{r} \left\{ \boldsymbol{\beta}'\mathbf{s}(t_{(j)}) - d_j \log \sum_{l \in \mathscr{R}(t_{(j)})} \exp\left(\boldsymbol{\beta}'\mathbf{x}_l(t_{(j)})\right) \right\}. \tag{8}$$

A nonparametric estimate of the baseline hazard at $t_{(j)}$, $\lambda_0(t_{(j)})$, is then given by (cf. Kalbfleisch and Prentice, 1973)

$$\hat{\lambda}_0(t_{(j)}) = 1 - \hat{\xi}_j, \tag{9}$$

where $\hat{\xi}_j$ is the solution to the equation

$$\sum_{l \in \mathscr{D}(t_{(j)})} \frac{\exp\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_l(t_{(j)})\right)}{1 - \hat{\xi}_j^{\exp(\hat{\boldsymbol{\beta}}'\mathbf{x}_l(t_{(j)}))}} = \sum_{l \in \mathscr{R}(t_{(j)})} \exp\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_l(t_{(j)})\right) \tag{10}$$

and $\mathscr{D}(t_{(j)})$ indexes the set of funds dying at time $t_{(j)}$. $\hat{\boldsymbol{\beta}}$ is the maximum likelihood estimate obtained in the previous step.

3.3. Empirical results

There are many reasons why mutual funds are terminated. One is that a fund is launched, but never reaches a critical mass in terms of market capitalization and is subsequently closed down. Another is related to rationalization within the fund management group: a new manager might decide that the group has two funds with too similar objectives and so decides to merge the poorer performing fund into the more successful fund. A different explanation, but with similar effect, is takeover and merger activity within the unit trust industry: this can result in duplication of funds with similar objectives and, again, a poorer performing fund is likely to be merged into a more successful one. As a final reason, the manager might decide to close down the poorest performing funds in order to raise the average performance of the group as a whole. While there is a wide range of reasons for funds closing, they all appear to be related to fund performance. [15] For this reason, we attempt to explain fund deaths principally in terms of previous return performance.

The covariates used in the analysis were constructed as follows. As a measure of the level of returns within a sector, we simply used the past average return on the funds that were in existence during the included months. Many of the existing sectors simply do not have any good alternative indices measuring the overall performance of the sector, and the purpose of our calculation is to isolate a common component in a sector's returns. Evidence cited in Section 2 suggests

[14] We also examined the approximation due to Efron (1977). The two approximations gave almost identical results so we report only the computationally less demanding Breslow approximation.

[15] Even in the case of a new fund that does not reach critical mass and is closed down, the main explanation for this is likely to be disappointing initial performance.

that, within a given sector, fund birth and death rates depend on past returns in the sector, so it is important to account for this common component in the hazard specification – otherwise, the residuals used in the analysis would no longer be independent and the resulting likelihood function would be misspecified. To construct a measure of a fund's relative performance, we used a simple peer-group adjustment, deducting from a fund's past average return the average market return over the same period. Using a sample of Canadian funds, Berkowitz and Kotowitz (1993) found that a fund's market share only responds relatively slowly to previous performance with weights on past performance that increase up to a 3-year-lag. Following this work, and in order to get a sense of the investment horizon adopted by investors to assess fund performance, we report results separately for 12- and 36-month horizons. [16]

Comparisons between a given fund's returns and the average return within the fund's sector were also made by Brown and Goetzmann (1995) and can be justified by recent studies (e.g., Brown et al., 1996; Chevalier and Ellison, 1997) which point to the importance in the assessment of fund managers' skills of their relative performance against a peer-group index. To illustrate, managers of commodity and energy stocks are far more likely to be judged against their peers than against some overall market index, on the grounds that fund managers from this sector face similar sets of objectives and constraints. Such peer-group comparisons are particularly important for our data set since the sectors have formal restrictions on their choice of assets as described in Section 2.

Table 3 reports the estimation results by sector. In 19 out of 20 sectors, a negative performance by a fund relative to its peer-group is associated with a higher hazard rate. Of these sector estimates, eight were statistically significant at the 5% level when performance was judged over the previous year, while 11 estimates were significant based on relative performance over the previous 3 years. The coefficients on lagged relative performance also tend to be largest over the 3-year performance measurement period, indicating that a fund's performance over the relatively long-term matters the most for its subsequent survival. There is also a large dispersion in estimated values across sectors ranging from -1.52 to 0.07 at the 12-month horizon. Perhaps this is not a surprising finding since many sectors have both relatively few funds and hence imprecisely determined coefficients.

There was also a negative coefficient on the sector return in 13 out of 20 sectors at the 12-month evaluation horizon and in 15 out of 20 sectors at the 36-month horizon. In line with the observations in Section 2, the UK equity

[16] We report results separately for these periods rather than including several lags of performance over, say, the previous 12, 24, and 36 months in a single regression. This is because our sample is not a balanced panel (since certain funds close before building up a sufficiently long track record), so we cannot always compute past returns at several lag lengths. If a fund closes prior to building up, say, a 36-month track record, we simply use the mean performance during the fund's lifetime.

Table 3
Estimated effects of the covariates on the hazard rates: Cox semiparametric regressions

Sector	Performance measured over preceding 12 months		Performance measured over preceding 36 months	
	Abnormal return	Market return	Abnormal return	Market return
UK equity growth	**−0.649** (−5.500)	**−0.215** (−3.071)	**−0.984** (−6.390)	**−0.234** (−2.108)
UK equity general	**−0.975** (−5.669)	−0.038 (−0.427)	**−1.247** (−5.567)	0.116 (0.753)
UK equity income	**−0.498** (−4.049)	**−0.152** (−2.375)	**−0.547** (−3.218)	**−0.193** (−2.218)
UK smaller companies	−0.139 (−1.130)	−0.082 (−0.812)	−0.448 (−1.349)	−0.234 (−1.648)
UK gilt and fixed interest	−0.387 (−1.261)	−0.001 (−0.006)	−0.268 (−0.561)	0.050 (0.134)
UK balanced	**−0.801** (−2.464)	0.050 (0.286)	**−1.544** (−3.327)	−0.218 (−0.779)
Financial and property	**−1.015** (−3.104)	−0.102 (−0.638)	**−1.577** (−3.451)	−0.291 (−1.021)
Investment trust units	−0.378 (−0.521)	−0.468 (−1.843)	−2.064 (−1.301)	**−0.619** (−3.346)
Commodity and energy	−0.159 (−1.459)	0.114 (1.629)	**−0.382** (−2.315)	0.115 (0.280)
International equity growth	−0.265 (−1.755)	−0.090 (−1.139)	**−0.986** (−4.382)	−0.110 (−0.636)
International equity income	−0.446 (−1.177)	0.044 (0.246)	−0.459 (−0.888)	−0.255 (−0.746)
International fixed interest	−1.522 (−1.281)	−0.956 (−1.380)	2.437 (1.018)	−0.382 (−0.367)
International balanced	−0.201 (−0.804)	−0.275 (−1.648)	−0.529 (−1.165)	−0.223 (−0.753)
Fund of funds	−0.279 (−0.530)	−0.092 (−0.371)	−0.838 (−0.951)	−0.426 (−1.057)
North America	**−0.346** (−4.220)	0.002 (0.029)	**−0.491** (−3.897)	−0.117 (−0.880)
Europe	**−0.411** (−4.618)	−0.062 (−0.925)	**−0.555** (−5.500)	**−0.209** (−2.714)
Japan	**−0.485** (−2.073)	0.144 (1.500)	**−0.645** (−2.022)	0.181 (1.146)
Far East including Japan	0.069 (0.259)	−0.012 (−0.133)	−0.189 (−0.461)	0.025 (0.111)
Far East excluding Japan	−0.368 (−1.139)	0.054 (0.370)	**−0.814** (−2.114)	−0.122 (−0.482)
Australasia	−0.216 (−0.939)	0.155 (1.449)	−0.625 (−1.727)	−0.262 (−0.882)

Columns two to five report the semiparametric estimates of the coefficients on the time-varying covariates in the hazard function. A fund's abnormal return is measured as the fund's average return over the previous 12 or 36 months relative to the average return within the fund's sector during the same period. Market returns are measured as the average returns within the sector during the previous 12 or 36 months.
The figures in parentheses are *t*-statistics and parameter estimates written in boldface are significant at the 5% critical level.

sectors all produced negative coefficients on the sector return covariate, but only for UK equity growth and UK equity income were these statistically significant. None of the positive estimates was statistically significant.

3.4. Risk-adjusted results for UK equities

Our usage of peer-group adjusted relative returns in the above calculations has the benefit of reflecting UK mutual fund industry practice, but is only one possible method for assessing a fund's performance within a given sector. An alternative, more conventional measure of performance is obtained by adjusting fund returns for their exposure to multiple risk factors. In the absence of a complete set of benchmarks for the returns on all the asset categories included in Table 3, we restrict our analysis to the UK equity sectors for which good external benchmarks are available; these sectors collectively cover 31% of the UK mutual fund industry by number of funds. We computed abnormal returns by regressing the funds' excess returns (relative to a short risk-free rate, $r_{f,t}$) on a constant and the excess returns on the stock market index, $r_{m,t} - r_{f,t}$, small-cap stocks (over the market index), $r_{s,t} - r_{m,t}$, and 5-year government bonds (over the risk-free rate), $r_{5,t} - r_{f,t}$:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{m,i}(r_{m,t} - r_{f,t}) + \beta_{s,i}(r_{s,t} - r_{m,t}) + \beta_{5,i}(r_{5,t} - r_{f,t}) + \epsilon_{i,t} \tag{11}$$

As our measure of the market index, we used returns on the FT-A All Share Index, while the Hoare–Govett Small-Cap Index compiled by the London Business School was used to measure returns on small-cap stocks. A 3-month T-bill rate was used as the risk-free rate. This specification is based on the 'four-index' risk-adjustment procedure suggested by Elton et al. (1993) and Gruber (1996), but differs from theirs in some respects. [17]

Table 4 presents the results from the semiparametric estimation of the hazard function using these alternative covariates. [18] To assess more fully the importance

[17] There is no good long-running index for returns on large-cap stocks in the UK, so we simply use the difference between returns on small-caps and the market portfolio to capture a small-size factor. There are also no commonly used growth and income equity indices in the UK, so we exclude a growth minus income factor. Finally, we use returns only on government bonds, and exclude corporate bonds, because of the dominance of government bonds in the UK bond market.

[18] To compare the specifications based on the two alternative performance measurement procedures (i.e., peer-group adjusted performance vs. three-factor risk-adjusted performance), we computed the partial log-likelihood function for the 20 sectors. We used the average return of a given sector less the risk-free rate as a proxy for the market index. We found that the two procedures gave very similar results for most sectors in the sense that the resulting partial log-likelihood functions were very similar. Only in the UK equity growth, UK equity general, fund of funds and North America sectors did the risk-adjusted covariate seem to provide a better fit, while the peer-group adjusted covariate provided the better fit for the international equity growth and Europe sectors.

Table 4
Effects of covariates on the hazard rate

	UK equity growth	UK equity general	UK equity income	UK smaller companies	UK equity total
Performance measured over preceding 12 months					
Abnormal return	**−0.665** (−5.038)	**−1.329** (−7.069)	**−0.378** (−3.024)	0.243 (1.482)	**−0.533** (−6.922)
Market return	**−0.224** (−3.068)	−0.015 (−0.158)	**−0.134** (−2.030)	−0.095 (−0.931)	**−0.124** (−3.180)
Performance measured over preceding 24 months					
Abnormal return	**−1.024** (−4.571)	**−1.397** (−6.621)	**0.361** (2.360)	0.302 (1.678)	**−0.660** (−7.021)
Market return	**−0.224** (−2.093)	0.112 (0.683)	−0.152 (−1.810)	−0.244 (−1.848)	**−0.154** (−2.852)
Performance measured over preceding 36 months					
Abnormal return	**−1.211** (−7.082)	**−1.455** (−6.644)	**−0.404** (−2.494)	0.283 (1.497)	**−0.744** (−4.621)
Market return	**−0.264** (−2.296)	0.153 (0.884)	−0.163 (−1.852)	−0.239 (−1.683)	**−0.161** (−2.649)
Performance measured over preceding 60 months					
Abnormal return	**−1.403** (−7.346)	**−1.374** (−6.189)	**−0.384** (−2.259)	0.340 (1.709)	**−0.759** (−7.093)
Market return	−0.226 (−1.638)	0.116 (0.624)	−0.128 (−1.293)	**−0.392** (−2.306)	**−0.153** (−2.354)

UK equities, risk-adjusted returns.
A fund's abnormal return is measured as the average over the preceding period of the risk-adjusted excess return computed using a 'three-factor' model similar to Elton et al. (1993). Market returns were measured as the average returns over the preceding period within a given sector.
The figures in parentheses are t-statistics and parameter estimates written in boldface are significant at the 5% critical level.

of the length of the performance measurement interval, and to exploit the substantial number of funds comprising the UK equity sectors, we report results for 12-, 24-, 36- and 60-month intervals. Three of the four equity sectors, as well as the total UK equity sector, produced negative coefficients for the abnormal return covariate, with all of the coefficients being statistically significant at the 5% level. Only the UK smaller companies sector generated a (statistically insignificant) positive coefficient on this covariate. However, this result is most likely explained by the very small number of non-surviving funds (30) in this sector. The coefficient estimates on lagged risk-adjusted performance increase in size up to a return horizon of around 36 months and then flatten out. This is in line with the finding of Berkowitz and Kotowitz (1993) that investors take a relatively long view when assessing a fund's performance.

To assess the economic significance of the coefficient estimates, we consider the results for the total UK equities sector. The coefficient estimate on the abnormal return over the previous 36 months, at -0.74, is almost five times larger than the estimate obtained on the market return (-0.16). These estimates imply that the cumulative effect of a 1% abnormal underperformance in a given month is to more than double a fund's hazard rate relative to the scenario of zero relative underperformance. Similarly, the cumulative effect of a decrease in the return on UK equities of 1% in a given month is to increase the hazard rate by 17%. These estimates suggest that both relative and absolute performance are important determinants of the UK equity funds' hazard rate, with relative performance being particularly important. [19]

Fig. 3 shows a smoothed plot of the nonparametric estimates of the baseline hazard, computed according to Eqs. (9) and (10) and measuring abnormal performance according to Eq. (11). Adjusting for the effects on the hazard rate of the time-varying covariates is clearly important: it lowers the overall level of the hazard rates, and the peak of the hazard rate is moved from a duration of 5 to 10 years to a duration of 10 to 15 years. At its peak the baseline hazard rate is around 0.20% per month. As in the case of the plot in Fig. 2 of the Kaplan–Meier estimate of the unconditional hazard rate, the adjusted baseline hazard is lowest for young and very old funds, and the variation in the hazard rate is very considerable. At its peak between 10 and 15 years, the baseline hazard rate is around four times higher than its level for the youngest and oldest funds in the sample.

[19] A referee suggested that funds might be punished for errors in tracking the market or peer-group index and conjectured that this might be reflected in a lower R^2 statistic in regressions of non-surviving funds' excess returns on the peer-group excess return relative to the R^2 statistic produced by the surviving funds. To investigate this possibility, we computed the R^2 statistic for all 20 sectors and found that in 18 out of 20 sectors, the mean R^2 of the dead funds was smaller than that of the surviving funds. Across sectors, the mean R^2 statistic was 0.69 for dead funds and 0.79 for surviving funds.

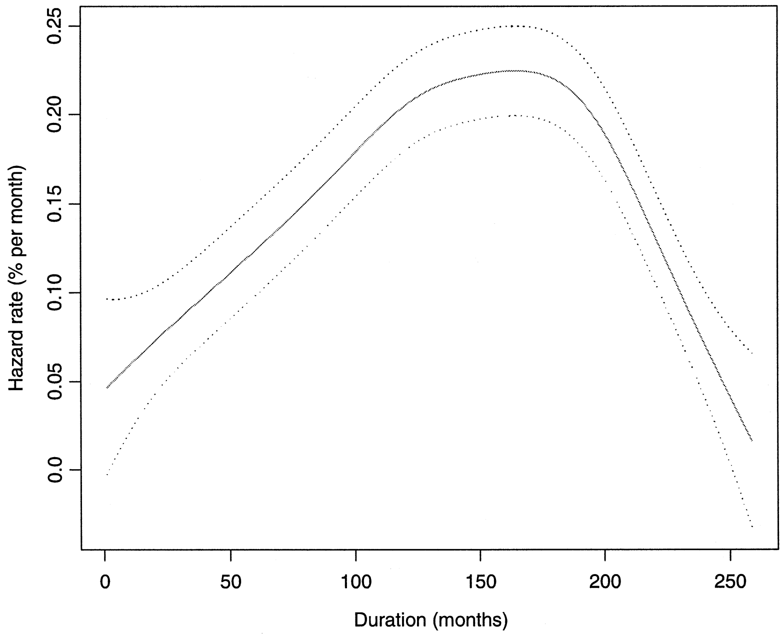


Fig. 3. This figure shows the smoothed baseline hazard function for the UK equities total sector based on the semiparametric Cox regression. The 95% confidence interval for the smoothing spline was calculated from jackknife residuals.

Our analysis of the funds' hazard function provides estimates of the probabilities of closure as a function of a fund's risk-adjusted performance and the performance of the sector in which the fund is operating. A closely related question is whether there exists a critical value such that abnormal performance below this point almost inevitably leads to closure. To investigate this, we define an event to be fund closure and the event period (denoted closure period) to be the 36 months preceding closure, while the non-event period (denoted continuation period) consists of non-overlapping 36-month intervals prior to the event period. We then computed histograms of the mean abnormal performance for the closure and continuation periods and compared these. Table 5 reports fractiles of the two distributions as well as the mean and standard deviation of the average monthly performance during the 3-year intervals in the case of UK equity funds. Irrespective of the way abnormal returns are calculated, the distribution of performance has a much lower mean in the closure period than in the continuation period and, most importantly, also has a much fatter left tail. For example, the 25% fractile of the non-survivors corresponds to the 5–10% fractile of the surviving funds. While this suggests that there is no mechanical rule for deciding when mutual funds are

Table 5
Distribution of fund performance in closure and continuation periods

	Risk-adjusted return		Peer-group adjusted return	
	Closure period	Continuation period	Closure period	Continuation period
5%	−1.449	−0.871	−1.230	−0.657
10%	−1.154	−0.681	−0.905	−0.466
25%	−0.759	−0.375	−0.491	−0.210
50%	−0.423	−0.123	−0.215	0.001
75%	−0.103	0.133	0.015	0.237
90%	0.257	0.441	0.284	0.519
95%	0.607	0.695	0.561	0.816
Mean	−0.399	−0.084	−0.243	0.038
Standard deviation	0.821	0.805	0.541	0.531

The table reports fractiles of the distribution of mean abnormal return in the 36-month period prior to the event of a fund's closure and in non-overlapping 36-month continuation periods. Abnormal return is measured either as the risk-adjusted excess return computed using a three-factor model similar to Elton et al. (1993) or by subtracting the peer-group return from an individual fund's return in a given month.

closed down, it confirms that the probability of closure conditional on past abnormal performance varies significantly across different levels of performance.

4. Discrete choice models

In this section we compare the Cox regression approach with the probit model adopted by Brown and Goetzmann (1995) in the only other study to date of mutual fund closure. Brown and Goetzmann (1995) construct a series of binary variables, $Y_i(1),\dots,Y_i(t)$, for each fund i. Each variable, $Y_i(t_a)$, is assigned the value 0 if the fund lives through the time period $[t_{a-1},t_a)$ and the value 1 if the fund is terminated within this interval. Traditional probit models are then estimated from

$$\Pr\big(Y_i(t_a)=1\big)=\Phi\big(\alpha t_{a-1}+\boldsymbol{\beta}'\mathbf{x}_i(t_{a-1})\big),\quad a=1,\dots,A_i, \tag{12}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal variate. This is a conditional probability for the risk of termination in the interval $[t_{a-1},t_a)$, given that this interval is reached in the first place, so the full notation for the *continuation probability* would be $\Pr(Y_i(t_a)=1|Y_i(t_{a-1})=\dots=Y_i(t_1)=0)$. This probability is distinct from the unconditional probability of termination in the interval $[t_{a-1},t_a)$ which is given by

$$\begin{aligned}\Pr\big(Y_i(t_a)=1,Y_i(t_{a-1})=\dots=Y_i(t_1)=0\big)\\ =\Phi\big(\alpha t_{a-1}+\boldsymbol{\beta}'\mathbf{x}_i(t_{a-1})\big)\\ \times\prod_{s=1}^{a-1}\big(1-\Phi\big(\alpha t_{s-1}+\boldsymbol{\beta}'\mathbf{x}_i(t_{s-1})\big)\big).\end{aligned} \tag{13}$$

The contribution to the log-likelihood function of the probit model resulting from the ith fund's observed lifetime path is

$$\begin{aligned}\mathscr{L}\left(Y_i(t_{A_i}) = 1, Y_i(t_{A_i-1}) = \ldots = Y_i(t_1) = 0\right) \\ = \Phi\left(\alpha t_{A_i-1} + \boldsymbol{\beta}'\mathbf{x}_i(t_{A_i-1})\right) \\ \times \prod_{a=1}^{A_i-1} (1 - \Phi(\alpha t_{a-1} + \boldsymbol{\beta}'\mathbf{x}_i(t_{a-1})) \\ = \prod_{a=1}^{A_i} \Phi(\alpha t_{a-1} + \boldsymbol{\beta}'\mathbf{x}_i(t_{a-1}))^{Y_i(t_a)} \\ \times (1 - \Phi(\alpha t_{a-1} + \boldsymbol{\beta}'\mathbf{x}_i(t_{a-1})))^{1-Y_i(t_a)}.\end{aligned} \qquad (14)$$

The closure probability $\Pr(Y_i(t_a) = 1 | Y_i(t_{a-1}) = 0, \ldots, Y_i(t_1) = 0)$, assumed by Brown and Goetzmann to be of the probit type, is in fact a particular parameterization of the discrete-time hazard function. To establish this point, we first draw the connection between the continuous-time hazard specification

$$\lambda(t|\mathbf{x}) = \lambda_0(t)\exp(\boldsymbol{\beta}'\mathbf{x}(s)) \qquad (15)$$

and its discrete-time equivalent. The need for discrete-time models arises when the failure time cannot be observed continuously, but is known to lie between consecutive observations. Such *grouped duration* models (or interval censored models, see Kalbfleisch and Prentice, 1980; Kiefer, 1990; Fahrmeir and Tutz, 1994) are not widely used in finance so we briefly outline their main characteristics. First, the discrete-time hazard function is given by

$$\lambda^{\mathrm{d}}(t_a|\mathbf{x}_i) = \Pr(T < t_a | T \geq t_{a-1}, \mathbf{x}_i), \quad a = 1,2,\ldots,A_i, \qquad (16)$$

which is the conditional probability of failure in the interval $[t_{a-1}, t_a)$ given that this interval is reached. The discrete-time survivor function for the interval $[t_{a-1}, t_a)$ can be written as

$$S^{\mathrm{d}}(t_a|\mathbf{x}_i) = \Pr(T > t_a|\mathbf{x}_i) = \prod_{s=1}^{a}(1 - \lambda^{\mathrm{d}}(t_s|\mathbf{x}_i(s-1))), \qquad (17)$$

where $\mathbf{x}_i \equiv (x_i(t_1)', \ldots, x_i(t_{a-1})')'$. Hence, the unconditional probability of failure in $[t_{a-1}, t_a)$ is

$$\begin{aligned}\Pr(t_{a-1} \leq T < t_a|\mathbf{x}_i) &= \lambda^{\mathrm{d}}(t_a|\mathbf{x}_i(t_{a-1}))S^{\mathrm{d}}(t_{a-1}|\mathbf{x}_i) \\ &= \lambda^{\mathrm{d}}(t_a|\mathbf{x}_i(t_{a-1}))\prod_{s=1}^{a-1}(1 - \lambda^{\mathrm{d}}(t_s|\mathbf{x}_i(s-1))).\end{aligned} \qquad (18)$$

Suppose again that $Y_i(t_{A_i}) = 0$ indicates that the ith fund is censored at t_{A_i}. Then Eq. (14) obtains when the discrete-time hazard function is of the probit type and there is a one-to-one mapping between a discrete-time duration model and the discrete choice model used by Brown and Goetzmann (1995). We can now derive the connection between the continuous- and discrete-time duration models:

$$\begin{aligned}
\lambda^{\mathrm{d}}(t_a|\mathbf{x}_i) &= \frac{\Pr(t_{a-1} \le T < t_a)}{\Pr(T \ge t_{a-1})} \\
&= \frac{F(t_a) - F(t_{a-1})}{S(t_{a-1})} \\
&= \frac{S(t_{a-1}) - S(t_a)}{S(t_{a-1})} \\
&= 1 - \frac{\exp\left(\int_0^{t_a} \lambda(s)\mathrm{d}s\right)}{\exp\left(\int_0^{t_{a-1}} \lambda(s)\mathrm{d}s\right)} \\
&= 1 - \exp\left(\int_{t_{a-1}}^{t_a} \lambda(s)\mathrm{d}s\right) \\
&= 1 - \exp\left(\int_{t_{a-1}}^{t_a} \lambda_0(s)\exp(\boldsymbol{\beta}'\mathbf{x}_i(s))\mathrm{d}s\right) \\
&= 1 - \exp\left(\exp(\boldsymbol{\beta}' x_i(t_{a-1}))\int_{t_{a-1}}^{t_a} \lambda_0(s)\mathrm{d}s\right) \\
&= 1 - \exp\left(\exp\left(\tilde{\lambda}_{0a} + \boldsymbol{\beta}'\mathbf{x}_i(t_{a-1})\right)\right), \quad \text{for } t_{a-1} \le t < t_a
\end{aligned} \tag{19}$$

where

$$\tilde{\lambda}_{0a} = \ln\int_{t_{a-1}}^{t_a} \lambda_0(s)\mathrm{d}s = \ln\left(\int_0^{t_a} \lambda_0(s)\mathrm{d}s - \int_0^{t_{a-1}} \lambda_0(s)\mathrm{d}s\right). \tag{20}$$

It follows that a proportional hazard model corresponds to assuming a Type-I extreme value cumulative probability of closure for each interval. Seen this way, months are no longer the smallest possible time unit and consequently the baseline hazard $\lambda_0(t)$ is not identified by the data. Only the change in the cumulative hazard between the interval limits can be identified. In this strict sense, the Cox regression is not a semiparametric model but instead corresponds to estimating a piecewise linear baseline hazard. The large number of intervals implies that the baseline hazard contains 281 parameters. Our assumption that months are a continuous time unit allows use of the Cox regression [20] which provides the

[20] Ryu (1995) analyzes the potential bias induced by estimating continuous-time proportional hazard models using discrete duration data. He finds that the bias decreases with the interval length. The implication is that the interval length should be a small fraction of the average duration. In the present data set this ratio is smaller than 1/116, which is about 10 times smaller than the size Ryu suggests.

estimated baseline hazard: $\tilde{\lambda}_{0a}$, $a = 1,\ldots,A_i$. The only difference is that the Cox regression estimates the parameters in two steps, whereas the discrete approach uses simultaneous estimation.

In general, the discrete-time model for the ith fund can be written as

$$\Pr\left(T < t_a \middle| T \geq t_{a-1}, \mathbf{x}_i\right) = \lambda^{\mathrm{d}}\left(t_a \middle| \mathbf{x}_i\right) = \mathscr{F}\left(\tilde{\lambda}_{0a} + \boldsymbol{\beta}' \mathbf{x}_i(t_{a-1})\right), \; a = 1,\ldots,A_i. \tag{21}$$

The model can then be estimated parametrically by specifying the $\mathscr{F}(\cdot)$ function. There are numerous ways in which this can be done. [21] The most widely applied functional forms are logit and probit models (Brown and Goetzmann, 1995), so we will compare these with the extreme value link derived from the Cox model. That is, we estimate the following three specifications:

$$\begin{aligned} \mathscr{F}^{\mathrm{L}}(z_a) &= \frac{\exp(z_a)}{1+\exp(z_a)}, \\ \mathscr{F}^{\mathrm{P}}(z_a) &= \Phi(z_a) \text{ and} \\ \mathscr{F}^{\mathrm{EV}}(z_a) &= 1 - \exp(-\exp(z_a)), \end{aligned} \tag{22}$$

for five different choices of z_a, namely

Base: $z_a = \boldsymbol{\beta}' \mathbf{x}(t_{a-1})$,
Trend: $z_a = \alpha(t_{a-1}) + \boldsymbol{\beta}' \mathbf{x}(t_{a-1})$ (Brown and Goetzmann's choice),
Piecewise: $z_a = \tilde{\lambda}_{0a} + \boldsymbol{\beta}' \mathbf{x}(t_{a-1})$, with $\tilde{\lambda}_{0a}$ constant over 50, 15, and 10 months.

Using again the UK equity total sector and the three-factor risk-adjusted performance and sectoral performance measured over the preceding 36 months as covariates, the estimated parameters of the discrete-time models are reported in Table 6 and the values of their log-likelihoods appear in Table 7. Recall that the corresponding parameter estimates and t-statistics from the Cox regression are -0.744 (-7.37) and -0.161 (-2.82). The first point to note is that the qualitative results of the effects of the covariates on the closure probability obtained in the Cox regression are unaffected by discrete-time estimation. Second, irrespective of which specification of $\mathscr{F}$ is used, Table 7 shows that the log-likelihood continues to increase as we shorten the interval over which the $\tilde{\lambda}_{0a}$'s are constant. Furthermore, the estimates of $\boldsymbol{\beta}$ from the extreme value and logit specifications approach their values from the Cox regression as the interval over which the age effect stays constant is narrowed. This clearly suggests that the piecewise-linear model for the baseline hazard together with the Cox regression is the appropriate modelling framework since it circumvents the above-mentioned nuisance parameter problem.

[21] See Aldrich and Nelson (1984) for some alternatives from the discrete-choice literature.

Table 6
Effects of covariates in the discrete-choice models: UK equities total

	Abnormal returns			Market returns		
	Extreme value	Logit	Probit	Extreme value	Logit	Probit
Base	**−0.330** (−6.639)	**−0.337** (−6.593)	**−0.120** (−6.191)	−0.065 (−1.741)	−0.066 (−1.740)	−0.0229 (−1.692)
Trend	**−0.351** (−6.885)	**−0.358** (−6.824)	**−0.136** (−6.311)	−0.073 (−1.907)	−0.074 (−1.907)	0.0254 (−1.831)
Piecewise 1	**−0.424** (−7.513)	**−0.434** (−7.454)	**−0.154** (−6.459)	**−0.125** (−2.984)	**−0.128** (−3.059)	**−0.0424** (−2.776)
Piecewise 2	**−0.466** (−7.652)	**−0.482** (−7.448)	**−0.163** (−6.474)	**−0.138** (−3.158)	**0.147** (−3.198)	**−0.0463** (−2.932)
Piecewise 3	**−0.578** (−8.389)	**−0.601** (−8.114)	**−0.192** (−7.118)	**−0.147** (−3.127)	**−0.154** (−3.150)	**−0.0463** (−2.932)

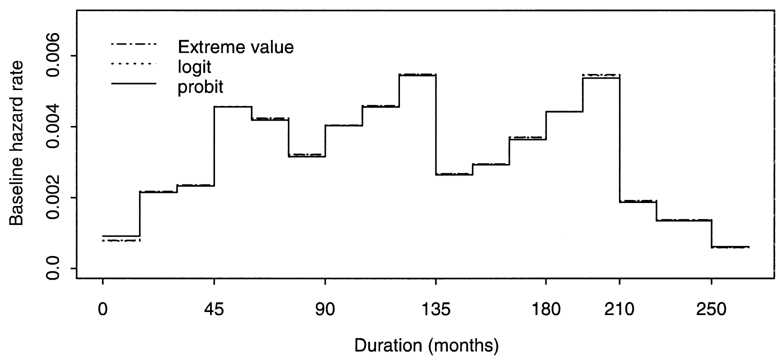
The figures in parentheses are *t*-statistics and parameter estimates written in boldface are significant at the 1% critical level.
The Base model includes only abnormal returns and market returns as regressors.
The Trend model adds fund age linearly, while Piecewise 1–3 fit a piecewise linear spline for the age effect based on intervals of 50, 15 and 10 months, respectively.

Table 7
Maximum value of the log-likelihood functions for the estimated discrete-choice models: UK equities total

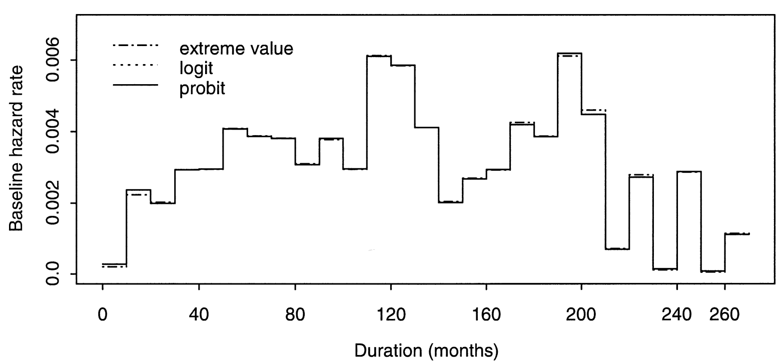
	Extreme value	Logit	Probit
Base	−1895.76	−1895.66	−1894.50
Trend	−1893.98	−1893.86	−1892.87
Piecewise 1	−1869.37	−1869.22	−1870.06
Piecewise 2	−1861.53	−1861.34	−1863.08
Piecewise 3	−1849.25	−1848.98	−1851.82

See Table 6 for explanations of the models.

It is important to recognize that the effect of the covariates is biased towards zero when the duration dependence is misspecified and Table 6 demonstrates the



(a) 15 month dummies



(b) 10 month dummies

Fig. 4. Estimated baseline hazard rates for the discrete choice model, $\mathscr{F}(\tilde{\lambda}_{0a})$ for $a=1,\ldots, A_i$, where $\mathscr{F}(\cdot)$ is the extreme value, logit and probit specification. Panel (a) has the baseline hazard rate piecewise linear within 17 15-month intervals, and panel (b) has the baseline hazard rate piecewise linear within 27 10-month intervals.

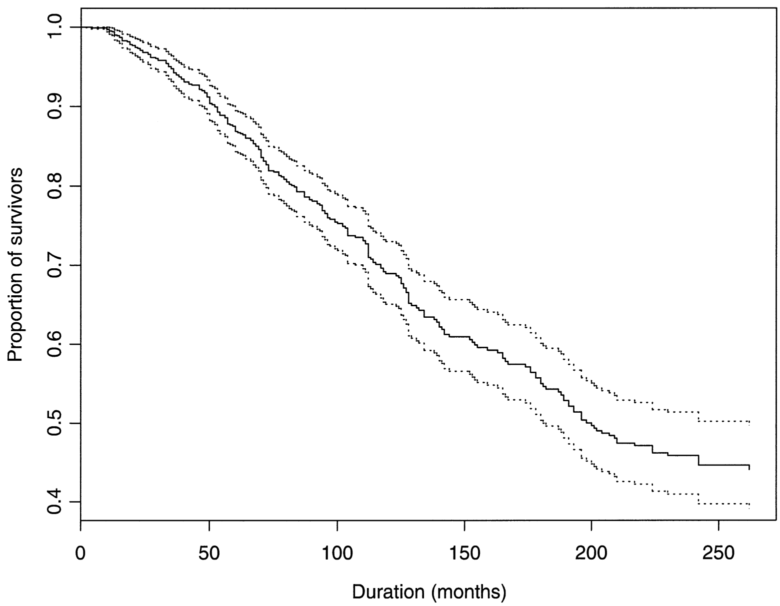


Fig. 5. Semiparametric estimate of the survivor function for UK equities total with a 95% confidence interval.

significance of using the correct specification in terms of our data set. In particular, we do not find evidence in favor of linear duration dependence. In Fig. 4 the discrete-time analogs of the baseline hazard are shown, with $\mathscr{F}(\tilde{\lambda}_{0a})$ constant over 15 and 10 months. This figure is consistent with the inverse U-shaped figure derived from the Cox regression. Hence, there are important benefits from using a nonparametric specification for duration dependence compared with the linear specification adopted by Brown and Goetzmann (1995).

5. Mutual funds' expected survival times

The funds' hazard rate modeled in the previous section provides information on the probability that, at a given point in time, a fund will close down the following month. To get a complete picture of the fund attrition process, we need to consider the proportion of funds that have survived up to a given age, i.e., the fund survivor function. The survivor function captures the effects on the cumulated hazard rate of both the funds' age and underperformance. It is based on the cumulative hazard function which can be estimated from the sequence of hazard rates:

$$\hat{\Lambda}_j(t) = \int_0^{t_{(j)}} \exp\left[\hat{\boldsymbol{\beta}}' \mathbf{s}(u)\right] \hat{\lambda}_0(u) \mathrm{d}u \tag{23}$$

for $t_{(j)} \leq t < t_{(j+1)}$, $j = 1, \ldots, r-1$. This estimate depends on integrals which are generally difficult to compute. We resolve this problem by using stepfunctions over the covariates and death times and compute the cumulative hazard as

$$\hat{\Lambda}_j(t) \approx \sum_{k=1}^{j} \left(t_{(k)} - t_{(k-1)}\right) \exp\left[\hat{\boldsymbol{\beta}}' \mathbf{s}\left(t_{(k-1)}\right)\right] \hat{\lambda}_0\left(t_{(k-1)}\right) \tag{24}$$

for $t_{(j)} \leq t < t_{(j+1)}$, $j = 1, \ldots, r-1$.

In turn, an estimate of the survival function can be obtained as

$$S_j(t) = \exp\left\{-\hat{\Lambda}_j(t)\right\} \text{ for } t_{(j)} \leq t < t_{(j+1)}, \; j = 1, \ldots, r-1. \tag{25}$$

Fig. 5 shows our estimate of the survivor function for the UK equity funds. The plot indicates that 25% of all funds have been terminated after 8 years, while 50% have been terminated after 15 years. That the location of the survivor curve is quite precisely determined is apparent from the narrowness of the 95% confidence interval plotted in Fig. 5. The shape of the survivor function clearly reflects the lower hazard rates faced by very young and very old funds.

6. Fund attrition and persistence of performance

Our findings in Section 3 suggest that a fund's relative performance is an important predictor of its probability of closure in the subsequent period. This finding would be difficult to rationalize unless relative performance persists over time. Indeed, if a fund's past performance did not help predict its future performance, then there would be no reason for investors to withdraw their money from funds with poor track records. Results reported by Lehmann and Modest (1987), Grinblatt and Titman (1992), Hendricks et al. (1993), Wermers (1996) and Carhart (1997) suggest that there does exist a group of persistently underperforming mutual funds in the US.

To explore the conjectured relationship between persistence of performance and fund attrition, we used a variant of the persistence measurement approach adopted by Hendricks et al. (1993). We sorted the set of UK equity funds into quartiles based on their peer-group-adjusted performance over the previous 36 months. [22] A 36-month assessment period was used because of the evidence in Section 3 that investors take a long view in assessing mutual fund performance. Only funds in existence at the time of the sorting were considered. Four equal-weighted portfolios were held over the subsequent 36-month period, after which the sorting and portfolio formation procedure was repeated, based on the new set of funds in

[22] We chose this measure because it avoids estimation of betas in the risk-adjustment procedure. This is an important consideration in our analysis since many of the (non-surviving) funds have relatively short return histories and hence, mean reversion in the $\boldsymbol{\beta}$ estimates is likely to induce substantial measurement errors in estimates of performance persistence.

existence.[23] Transition probabilities linking abnormal performance in the pre- and post-sorting periods were used as a statistical measure of persistence of performance and we measure the economic significance of past performance by computing the mean abnormal returns on the four past-performance-sorted portfolios.

Since we have a data set that includes both surviving and non-surviving funds, we can carry out these experiments both for the entire set of funds (comprising survivors and non-survivors) and for the set of surviving funds only. A comparison between the two sets of results will then shed light on the effect of fund closure on measured persistence of performance.

Table 8 reports the outcome from this exercise. First, consider the estimated transition probabilities for the full set of surviving and non-surviving funds (Panel A). If there was no persistence of performance, the estimated transition probabilities should all equal 0.25. This hypothesis is clearly rejected by the data: the probability that the historically worst-performing funds will remain in the bottom quartile of performers is estimated at 0.332. Likewise, the probability of repeated relative performance in the top quartile is 0.355.

Turning next to the surviving funds, the evidence of persistence in relative performance is weaker. For example, the transition probability estimates associated with repeat performance in the bottom and top quartiles are now 0.284 and 0.317, down by about five and four points for the worst and best performers, respectively, compared with the full set of UK equity funds.[24] In both cases, the decline in persistence can be attributed to drawing the funds from a more homogenous set of funds (funds observed conditional on survival) as opposed to drawing funds from a mixture of non-survivors and survivors, each with very different centering points of their return distribution (cf. Table 5).

A similar result emerges from our analysis of the economic significance of persistence. The estimates of mean abnormal performance in panel B show that the portfolio comprising the historically worst-performing funds paid a mean return of −0.11% per month. This compares with an abnormal performance of 0.11% for the historically best-performing funds, producing a difference of 0.21% per month. In comparison, the set of historically worst-performing surviving funds generated mean returns of 0.02% per month, while the historically best-performing surviving funds generated mean returns of 0.05%. The performance differential is thus of the order of 0.03%, or six times smaller than the differential return recorded for the full set of funds that included non-survivors.

[23] If a fund closed during the 36-month holding period, this fund was dropped from the portfolio in the subsequent month and a new equal-weighted portfolio excluding the dead fund was formed.

[24] A standard chi-squared test of significance of the individual cell counts showed that the upper left and lower right cell probabilities were statistically significantly different from 0.25 at the 1% critical level when the full set of funds was considered but only remained significant for the best performers in the set of surviving funds.

Table 8
Persistence of peer-group adjusted returns: UK equities total

(A) Transition probabilities

Past performance	Future performance			
	I	II	III	IV
All funds				
I (worst)	0.332	0.251	0.212	0.205
II	0.224	0.267	0.288	0.221
III	0.203	0.297	0.281	0.219
IV (best)	0.242	0.184	0.219	0.355
Surviving funds				
I (worst)	0.284	0.240	0.221	0.255
II	0.225	0.277	0.280	0.218
III	0.221	0.303	0.266	0.210
IV (best)	0.269	0.181	0.232	0.317

(B) Mean returns on funds sorted according to previous performance

	All funds	Surviving funds
I (worst)	−0.107	0.019
II	−0.039	0.042
III	−0.003	0.031
IV (best)	0.105	0.052

Every 36 months, the funds were sorted into quartiles based on their performance over the previous 3-year period. For each of these quartiles, the proportion of funds that fall into a given quartile based on performance over the subsequent 3-year period was recorded and is reported as transition probabilities.
Quartile I comprises funds with the worst relative performance while quartile IV consists of the best performing funds.
Panel B reports the mean returns on equally-weighted portfolios formed by sorting the set of funds based on past performance over the previous 3-year period. Only funds that were in existence at the time the sort was performed are included in the portfolios.
The ‘All funds’ portfolio includes both survivors and non-survivors while the ‘surviving funds’ portfolio consists exclusively of funds that were active at the end of the sample.

In a model where excess returns are genuinely serially uncorrelated, Brown et al. (1992) demonstrate how survivorship bias can introduce spurious persistence in standard measures of performance. In their model, there is no persistence when returns are computed across the full set of surviving and non-surviving funds in existence. Excluding funds with the worst performance from the sample will introduce a potentially serious bias in the estimated persistence for the surviving funds. A very different effect operates in our sample which is dominated by funds with persistently poor performance records and a higher than normal closure probability. Since these funds have higher persistence than the funds that survive, excluding them from the analysis leads to a *decline* in the persistence estimate.

We conclude from these results that the mutual fund attrition process can have an important effect on standard measures of performance persistence. While there is only weak evidence of persistence in the sample comprising funds that survived to the end of the sample, inclusion of non-surviving funds introduces stronger evidence of performance persistence.

7. Conclusion

Our finding of an inverse U-shaped hazard function for mutual funds may be consistent with a variety of theoretical models of investor behavior and it has an interesting theoretical precedence in the labor literature. Jovanovic (1979) proposes a model where workers and firms have to learn gradually the quality of their job match and this implies an inverse U-shape for the job tenure hazard function. Our findings can possibly be explained by a similar learning process in which investors extract information concerning fund performance. Investors are unlikely to know ex ante which funds will outperform in the future and hence have to learn this gradually as a fund's track record is established. After sufficient data has been accumulated, investors may recognize that certain funds underperform, withdraw their money, thereby causing the funds to close. Funds that survive this treatment are more likely to be perceived as having a good track record, thus, possibly explaining the subsequent decline in the hazard rate observed in our data.

In our set-up, the data from which investors infer a mutual fund's performance is its published returns record. Returns data are notoriously volatile and noisy so the process through which investors attempt to learn the true skills of a mutual fund manager is likely to be rather slow. This matches our findings of a 'honeymoon period', i.e., that mutual funds are typically not closed down early on after their inception. It is also consistent with the finding that a fund's performance over the previous 3 years matters more for its closure probability than its performance over the previous year.

A Bayesian updating story of fund closure also suggests that the hazard rate declines in periods where the market is noisy and inference is slowed down. If, as seems likely, genuinely good managers are distinguished from bad managers in that they observe information that is only weakly correlated with market returns, then luck rather than investment skill is likely to be more important to relative return performance in a very volatile market with a lower signal-to-noise ratio. To investigate whether hazard rates decline when market volatility is high, we repeated the Cox regression analysis for the UK equity funds, using both the lagged means and standard deviations of abnormal performance and sector performance as covariates. We found that while the standard deviation of individual funds' abnormal performance did not affect the hazard rate, the coefficient on the standard deviation of lagged sectoral return, at -0.29, was highly statistically significant. Consistent with the updating story, hazard rates thus appear to decline during volatile markets.

Acknowledgements

The first author would like to thank the Center for Nonlinear Modelling in Economics, University of Aarhus for financial support. Further, we would like to thank Rob Engle, Tony Lancaster, Lars Muus, Jeff Russell, Pravin Travedi and seminar participants at UCSD for discussions and comments on this paper. We also thank John Richardson of Micropal, for providing us with the data set used in the analysis. Two anonymous referees and the Editor (Franz C. Palm) provided many valuable suggestions that greatly improved the paper.

Appendix A: Analysis of residuals from the Cox regression

To assess the parametric specification in the Cox regression, we computed martingale residuals and score residuals. An estimate of the ith fund's cumulative hazard function is given by

$$\hat{\Lambda}_i(t) = \int_0^t \exp\left[\hat{\boldsymbol{\beta}}' \mathbf{x}_i(u)\right] \hat{\lambda}_0(u) \mathrm{d}u. \tag{26}$$

The martingale residual for the ith fund is defined as

$$r_{\mathrm{M}_i}(t_i) = \delta_i - \hat{\Lambda}_i(t_i). \tag{27}$$

where δ_i is zero if the ith fund is right-censored and unity otherwise. These residuals have properties similar to those possessed by residuals from linear regression analysis, although they are not symmetrically distributed about zero. The residual can be interpreted as the difference between the observed and expected number of deaths for the ith fund in the interval $(0, t_i)$. In Fig. 6, the martingale residuals for the UK equity funds are plotted against abnormal and market returns prior to death or censoring. The residuals in the interval $(0, 1)$ are the ones associated with the dead funds, while the residuals below zero correspond to censored funds. Consistent with the conclusion from the semiparametric regression analysis, dead funds appear to have worse underperformance than censored funds. Apart from this regularity, the plots do not display any systematic patterns. The clustering of points in the common component plot simply reflects the censoring of a group of funds at the end of the study. Funds in the same sector register the same value of the common component at this time. Significantly, very few extreme values of the residuals are encountered, indicating that the functional representation of the covariates is adequate.

The score residual for the ith fund's jth covariate is based on the score process:

$$r_{\mathrm{S}_{ij}}(\hat{\boldsymbol{\beta}}, t) = \delta_i\left\{x_{ij}(t_i) - \bar{x}_j(\hat{\boldsymbol{\beta}}, t_i)\right\} - \int_0^t \left\{x_{ij}(s) - \bar{x}_j(\hat{\boldsymbol{\beta}}, s)\right\} \mathrm{d}\hat{\Lambda}_i(s) \tag{28}$$

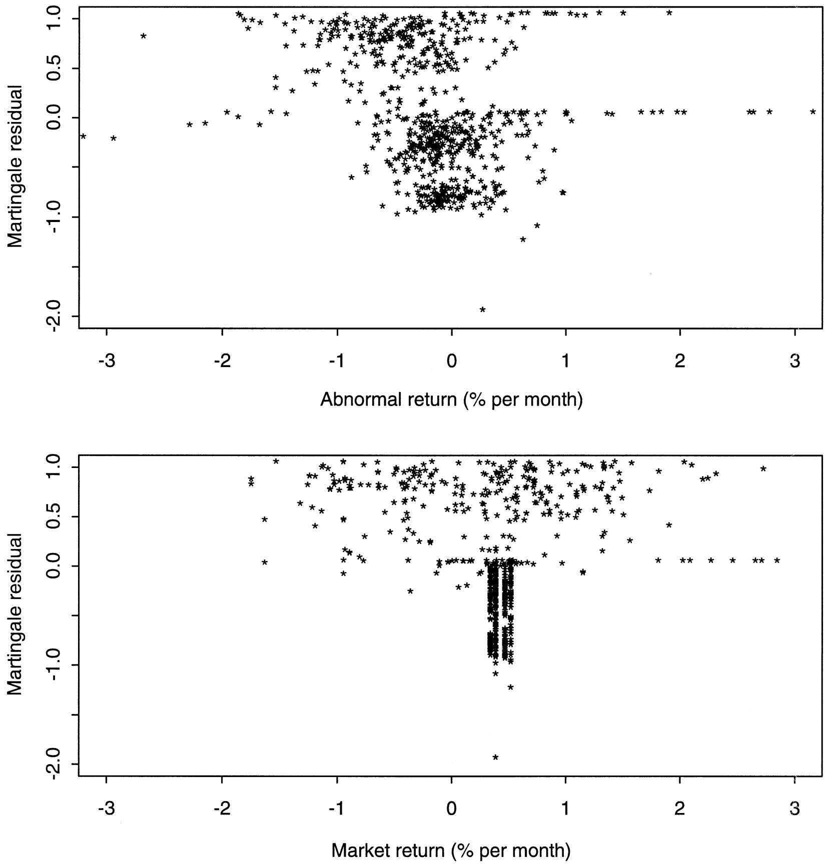


Fig. 6. Using the UK equity total sector, the figure plots martingale residuals from the semiparametric regression with time-varying covariates against the abnormal and market returns corresponding to either the death or censoring time for each fund.

where $\bar{x}_j$ is the weighted mean of the covariates of those funds that are still at risk at time t,

$$\bar{x}_j(\hat{\boldsymbol{\beta}},t) = \frac{\sum_{l \in \mathscr{R}(t)} x_{lj}(t) \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_l(t))}{\sum_{l \in \mathscr{R}(t)} \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_l(t))}. \tag{29}$$

The score residuals provide an estimate of the ith component of the efficient score for the model's jth parameter. These residuals can be used to examine the leverage of individual funds by computing the approximate change in $\hat{\boldsymbol{\beta}}$ if that observation was dropped. In Fig. 7, the score residuals for the UK equity funds are plotted against the survival time of the corresponding fund. The plot for the abnormal performance coefficient gives a very weak indication that funds with the

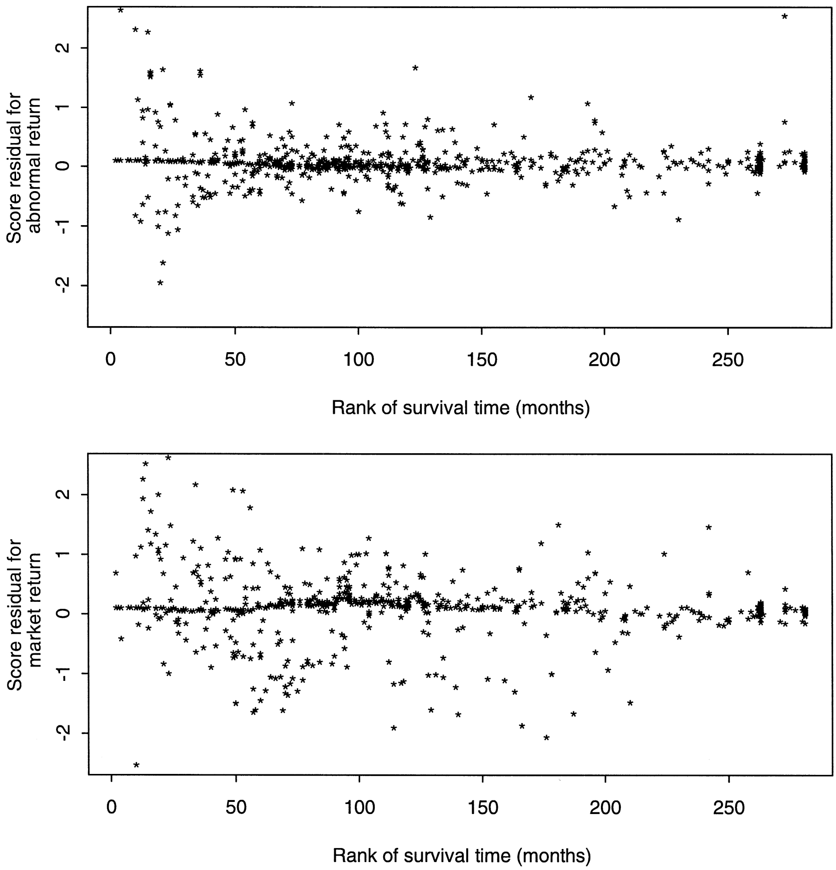


Fig. 7. Using the UK equity total sector, the figure plots score residuals for the abnormal and market return against a fund's age.

shortest survival times influence the coefficient estimates the most. There are no apparent anomalies in the plot for the common component coefficient.

We conclude from the analysis of these residuals that the functional specification of the included covariates is satisfactory.

References

Aldrich, J.H., Nelson, F.D., 1984. Linear probability, logit, and probit models, Discussion Paper 07-045. Sage University.

Berkowitz, M.K., Kotowitz, Y., 1993. Incentives and efficiency in the market for management services: a study of Canadian mutual funds. Canadian Journal of Economics 26 (4), 850–866.

Breslow, N.E., 1974. Covariance analysis of censored survival data. Biometrics 30, 89–100.

Brown, S.J., Goetzmann, W., 1995. Performance persistence. Journal of Finance 50, 679–698.
Brown, S.J., Goetzmann, W., Ibbotson, R.G., Ross, S.A., 1992. Survivorship bias in performance studies. Review of Financial Studies 5 (4), 553–580.
Brown, K.C., Harlow, W.V., Starks, L.T., 1996. Of tournaments and temptations: an analysis of managerial incentives in the mutual fund industry. Journal of Finance 51, 85–110.
Carhart, M.M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.
Chevalier, J., Ellison, G., 1997. Risk taking by mutual funds as a response to incentives. Journal of Political Economy 105 (6), 1167–1200.
Cox, D.R., 1972. Regression models and life-tables (with discussion). Journal of the Royal Statistical Society B 34, 187–220.
Cox, D.R., 1975. Partial likelihood. Biometrika 62 (2), 269–276.
Efron, B., 1977. The efficiency of Cox's likelihood for censored data. Journal of the American Statistical Association 76, 557–565.
Elton, E.J., Gruber, M.J., Das, S., Hlavka, M., 1993. Efficiency with costly information: A reinterpretation of evidence from managed portfolios. Review of Financial Studies 6, 1–22.
Fahrmeir, L., Tutz, G., 1994. Multivariate Statistical Modelling based on Generalized Linear Models, Springer, New York.
Grinblatt, M., Titman, S., 1989. Mutual fund performance: An analysis of quarterly portfolio holdings. Journal of Business 62, 393–416.
Grinblatt, M., Titman, S., 1992. The persistence of mutual fund performance. Journal of Finance 47, 1977–1984.
Gruber, M.J., 1996. Another puzzle: The growth in actively managed mutual funds. Journal of Finance 51, 783–810.
Han, A., Hausman, J.A., 1990. Flexible parametric estimation of duration and competing risk models. Journal of Applied Econometrics 5 (1), 1–28.
Hendricks, D., Patel, J., Zeckhauser, R., 1993. Hot hands in mutual funds: Short-run persistence of relative performance, 1974–88. Journal of Finance 48, 93–130.
Jovanovic, B., 1979. On the determination of the public debt. Journal of Political Economy 87, 972–990.
Kalbfleisch, J.D., Prentice, R.L., 1973. Marginal likelihood based on Cox's regression and life model. Biometrika 60 (2), 267–278.
Kalbfleisch, J., Prentice, R., 1980. The Statistical Analysis of Failure Time Data, Wiley, New York.
Kiefer, N.M., 1988. Economic duration data and hazard functions. Journal of Economic Literature 26, 646–679.
Kiefer, N.M., 1990. Econometric methods for grouped duration data. In: Hartog, J., Ridder, G., Theeuwes, J. (Eds.), Panel Data and Labor Market Studies, Elsevier, Amsterdam.
Lancaster, T., 1990. The Econometric Analysis of Transition Data, Econometric Society Monographs No. 17, Cambridge Univ. Press, Cambridge.
Lehmann, B.N., Modest, D.M., 1987. Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. Journal of Finance 42, 233–265.
Malkiel, B.G., 1995. Returns from investing in equity mutual funds 1971 to 1991. Journal of Finance 50, 549–572.
Pudney, S., 1989. Modelling Individual Choice: The Econometrics of Corners, Kinks and Holes, Blackwell, Oxford.
Ryu, K., 1995. Analysis of a continuous-time proportional hazard model using discrete duration data. Econometric Reviews 14 (3), 299–313.
Tunali, I., Pritchett, J.B., 1997. Cox regression with alternative concepts of waiting time: The New Orleans yellow fever epidemic of 1853. Journal of Applied Econometrics 12 (1), 1–25.
Wermers, R., 1996. Momentum investment strategies of mutual funds, performance persistence, and survivorship bias, Discussion paper, University of Colorado, Boulder, CO.