Journal of Empirical Finance 8 (2001) 345–373
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/econbase

Eliminating look-ahead bias in evaluating persistence in mutual fund performance ☆

Jenke R. ter Horst [a,*], Theo E. Nijman [b], Marno Verbeek [c]

[a] *Department of Finance, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands*
[b] *Center and Department of Finance, Tilburg University, Tilburg, The Netherlands*
[c] *Erasmus University Rotterdam and Center for Economic Studies, K.U. Leuven, The Netherlands*

Accepted 8 June 2001

Abstract

Performance persistence studies typically suffer from ex-post conditioning biases. As stressed by Carhart [Carhart, M.M., 1997. Mutual Fund Survivorship, Working Paper, Marshall School of Business, U.S.C.] and Carpenter and Lynch [J. Financ. Econ. 54 (1999) 337.], standard methods of analysis on a survivorship free sample are subject to look-ahead biases. In this paper, we show how one can easily correct for look-ahead bias using weights based on probit regressions.

First, we model how survival probabilities depend upon historical returns, fund age and aggregate economy-wide shocks, using two samples of US based 'income' and 'growth' funds. Subsequently, we employ a Monte Carlo study to analyze the size and shape of the look-ahead bias in performance persistence that arise when a survivorship free sample is used with standard techniques. In particular, we show that look-ahead bias induces a spurious *U*-shaped pattern in performance persistence. Finally, we demonstrate how a weighting procedure based upon probit regressions can be used to correct for this bias. In this way, we obtain look-ahead bias-corrected estimates of abnormal performance relative to a one-factor and the Carhart [J. Finan. 52 (1997) 57.] four-factor model, as well as its persistence. The results suggest that in this sample, look-ahead bias is of minor importance and does not seriously affect estimates of persistence. Our bias-corrected results closely correspond to the findings of Carhart [J. Finan. 52 (1997) 57.], implying that there is no

☆ A previous version of this paper was entitled 'Eliminating Biases in Evaluating Mutual Fund Performance from a Survivorship Free Sample'.

* Corresponding author.
E-mail address: J.R.terHorst@kub.nl (J.R. ter Horst).

PII: S0927-5398(01)00032-9

evidence on a risk-adjusted basis for persistence in performance.

JEL classification: G11; G23; C34
Keywords: Mutual funds; Performance evaluation; Look-ahead bias

1. Introduction

A question that has intrigued researchers for decades is whether some investors systematically outperform others. Many authors (including e.g. Brown and Goetzmann, 1995; Gruber, 1996; Carhart, 1997a) have recently reported persistence in the performance of mutual funds, i.e. funds with an above average performance in the recent past are more likely to have above average future performance as well. However, as shown by Carhart (1997a), much of the persistence in fund returns can be explained by exposures to common factors like size, book-to-market and momentum, implying that on a risk-adjusted basis the performance of the best performing funds is not significantly different from zero. On the other hand, there is a wide belief among individual investors that mutual fund performance can be forecasted (Zheng, 1999) and many investors decide to invest more in funds that recently performed well (see, e.g. Sirri and Tufano, 1998).

Many estimates of persistence in mutual fund performance are based on data sets that only contain funds that are in existence at the end of the sample period (see, e.g. Grinblatt and Titman, 1992). It is well understood that such estimates are subject to a survivorship bias (see, e.g. Brown et al., 1992; Hendricks et al., 1997; Carhart, 1997b; Blake and Timmermann, 1998) and may spuriously indicate persistence or reversals in returns (Carpenter and Lynch, 1999). Survivorship bias is caused by the fact that poor performing funds are less likely to be observed in data sets that only contain the surviving funds because the survival probabilities depend on past performance (see, e.g. Brown and Goetzmann, 1995). Many studies (Malkiel, 1995; Elton et al., 1995; Wermers, 1997 among others) have argued that survivorship biases can be avoided by using standard techniques on a sample that contains data on each fund up to the period of disappearance.[1] Such data-sets are usually referred to as ‘survivorship free’.

Persistence studies typically suffer from a second ex-post conditioning bias, the so-called look-ahead bias, which is inherent in the methodology (see, e.g. Brown et al., 1992). In performance persistence studies, it is common to distinguish two periods. At the end of the first period, the ranking period, funds are ranked and

[1] A common way to avoid survivorship bias is to use a so-called ‘follow-the-money’ approach. In this approach, it is assumed that as soon as a fund disappears that the investor places his or her money in the average surviving fund.

assigned to portfolios on the basis of their past performance. In the second period, the evaluation period, the average performance of all portfolios is determined for the funds that survived this second period. The fact that funds disappear in a nonrandom way during the ranking or evaluation period causes a bias which is referred to as a look-ahead bias. This bias is not remedied even if a survivorship free data-base is used.

The contribution of this paper is twofold. First, we analyze the potential effect of look-ahead bias in persistence studies if there is no genuine persistence. Independently, this question has been addressed by Carpenter and Lynch (1999), whose study is based upon assumptions about the survival process that are not tested on observed behavior. The results indicate that look-ahead bias may lead to spurious persistence and reversals in mutual fund performance. Second, our paper suggests an estimation procedure that corrects for look-ahead bias in evaluating the persistence of mutual fund performance, employing knowledge of the empirical survival process estimated on the basis of a survivorship free sample. As stressed by Brown et al. (1995), an explicit model for survival is necessary to fully adjust for the effects of ex-post conditioning biases. Our correction method enables us to determine the size of look-ahead bias in an empirical study of performance persistence of US based open-end mutual funds for the period 1989–1994. The results suggest that in this sample look-ahead bias is of minor importance, and does not seriously affect estimates of persistence. Our bias-corrected results closely correspond to the findings of Carhart (1997a), implying that there is no evidence on a risk-adjusted basis for persistence in performance.

The remainder of this paper is organized as follows. In Section 2, we describe the sample of US based mutual funds that we employ. We show that the number of funds that leaves the sample is substantial and their average returns are substantially below those of surviving funds. This indicates the potential for biases and the need to correct for them. Following the micro-economics literature on sample selection (starting with Heckman, 1976, 1979), we model in Section 3 the process that determines attrition from the sample, and subsequently analyze it jointly with the model for performance persistence. The longitudinal probit model that we use to model attrition extends the model in Brown and Goetzmann (1995) by allowing for aggregate macro-economic shocks. We use this model in Section 4 to analyze the effects of look-ahead bias if persistence is absent using a Monte Carlo experiment. The empirical survival process induces a spurious pattern of performance persistence that is *U*-shaped, rather than *J*-shaped as in Hendricks et al. (1997). In Section 5, we show how the survival model can be used to construct weights that can be applied to correct for look-ahead bias. The empirical implementation of the correction approach is presented in Section 6, where we examine persistence in performance of US open end ‘income’ and ‘growth’ funds over the period 1989–1994, using a simple one-factor model and Carhart’s (1997a) four-factor model. By and large, our results support the findings of Carhart (1997a) and do not indicate the existence of a hot hands phenomenon in mutual fund performance. Section 7 summarizes the main results and presents some concluding remarks.

2. Stylized facts on survival of US equity funds

To examine the importance of look-ahead bias in performance persistence studies, we employ a data set selected from the Morningstar Mutual Funds Ondisc database (February 1995 edition). This database contains monthly information on more than 6000 US based open-end equity as well as fixed income mutual funds. This sample suffers from survivorship because only funds that existed at the end of the sample period (February 1995) are included. Many mutual funds have disappeared from the sample because they have merged with other funds or were closed down completely. A first step in obtaining results free of survivorship bias is the inclusion of attrited funds in the sample. We did so by extending the database with the mutual funds that disappeared between the first quarter of 1989 and the last quarter of 1994. This additional information was provided by Morningstar.[2] In the sequel, we will refer to this data set covering 1989–1994 as the combined 'survivorship free' sample.

Following previous studies, we concentrate on equity funds. During the period 1989/1–1994/4, we observe 2678 funds with their name, objective, the year of fund inception, and returns until the quarter of disappearance.[3] Table 1 presents the number of funds by inception year and the number of them that did not survive the period 1989–1994, aggregated to a yearly level. Note that we aggregated all funds with an inception date before 1977, which explains the relatively large number of 279 funds with inception year 1976. Table 1 shows that 498 of the mutual funds in our sample disappeared between 1989/1 and 1994/4 due to merger or liquidation, which corresponds with a yearly average of 5.3%. This estimate differs from Carhart (1997b), who reports a non-survival rate of 3.6% over the period 1962–1995, which increases to 4.6% for the period 1989–1994. The remaining difference with our yearly average is due to inclusion of types of funds in our sample that have relatively low survival rates (compare Table 3 below). Furthermore, looking at for instance the 173 funds that started in 1990, 29.5% already disappeared within the next 4 years. A similar pattern appears to hold for other years, so that a first conclusion would be that a large part of the defunct funds disappeared at a relatively young age, age being defined as the time

[2] Unfortunately, Morningstar was unable to provide information about funds that ceased to exist before 1989.

[3] In contrast to Carhart (1997b), we also included specialty funds (i.e. 273 sector funds), internationally diversified US based funds (490) and a number of funds which advertise as 'balanced fund' and invest less than 50% in fixed income securities (180).

Table 1
Number of non-survivors

Inception year	Total in	In sample end year	Out in year						Total out
			1989	1990	1991	1992	1993	1994	
≤ 1976	279	279	5	3	4	7	16	8	43
1977	14	293		1		1	2		4
1978	13	306			2	2			4
1979	11	317	1		1		1		3
1980	10	327					1		1
1981	26	353	2		1	1			4
1982	33	386	1		2	2	3	1	9
1983	58	444	4	3	3	3	3	2	18
1984	75	519		2	3	2	3	5	15
1985	98	617	2	5	5	3	6	4	25
1986	143	760	1	11	8	9	13	7	49
1987	162	922	3	6	9	7	16	11	52
1988	140	1062	2	2	7	13	14	10	48
1989	107	1148		3	5	4	12	7	31
1990	173	1285			11	6	18	16	51
1991	187	1411				8	7	23	38
1992	277	1620					17	15	32
1993	550	2025					13	40	53
1994	322	2180						18	18
Total	2678		21	36	61	68	145	167	498
Non-surv. rate (%/year)			1.98	3.14	4.75	4.82	8.95	8.25	5.31

The table reports the annual number of funds by inception year since 1976 and the annual number that ceased to exist between 1989 through 1994. The row labelled 'Non-surv. rate' contains the number of disappearing funds divided by the total number of funds at the beginning of the year. The column labelled 'In sample end year' reports the sample size at end of year.

elapsed since fund inception. Apparently, it is not only the case that the number of mutual funds has grown at an increasing rate over the last decade, but also that the relative number of funds that has closed down or merged has increased significantly. At a more disaggregated level (not reported in the table), it appears that in some quarters relatively many mutual funds leave the sample while in other quarters almost no funds disappear, indicating that common aggregate factors may play a role in fund disappearance as well.

It is often claimed that a bad record of fund returns is one of the main reasons that funds disappear from a sample (see, e.g. Elton et al., 1995). Low returns compared to other funds as well as low returns relative to some benchmark portfolio seem to be a reason for the management of the fund to close down the fund or to let the fund merge (see Brown and Goetzmann, 1995). Table 2 presents the average quarterly returns for the period 1989–1994 for the funds that survived until the end of 1994, for the funds that did not survive, and for the combined sample. For comparison, we also added the quarterly returns on the Standard and

Table 2
Average quarterly returns

Quarter	Surviving funds		Non-survivors		Combined mean return	Return S&P500	Return T-bill
	Mean return	Number	Mean return	Number			
1989/1	6.50	819	5.06	296	6.12	7.03	1.85
/2	6.36	830	5.28	294	6.08	8.80	2.19
/3	9.79	843	7.97	289	9.32	10.64	2.10
/4	0.34	858	−0.45	286	0.14	2.05	1.98
1990/1	−2.72	888	−2.78	298	−2.74	−3.02	1.79
/2	5.75	921	4.65	306	5.47	6.29	2.00
/3	−14.91	951	−13.35	300	−14.54	−13.78	1.95
/4	6.83	980	5.01	301	6.40	8.95	1.86
1991/1	15.16	1025	12.30	308	14.50	14.56	1.44
/2	−0.79	1058	−1.21	301	−0.88	−0.21	1.43
/3	6.91	1095	5.99	284	6.72	5.38	1.41
/4	7.24	1129	6.03	278	7.00	8.36	1.20
1992/1	−0.63	1184	−1.75	274	−0.84	−2.55	0.96
/2	−1.07	1213	−1.34	265	−1.12	1.97	0.92
/3	1.44	1298	0.93	254	1.36	3.10	0.83
/4	6.84	1374	5.63	242	6.66	5.10	0.75
1993/1	4.88	1515	3.59	241	4.70	4.28	0.71
/2	2.86	1603	2.10	253	2.75	0.51	0.71
/3	5.69	1732	4.51	164	5.59	2.56	0.75
/4	4.19	1871	2.36	150	4.05	2.31	0.70
1994/1	−3.13	2043	−3.46	136	−3.15	−3.82	0.73
/2	−2.03	2181	−2.61	115	−2.06	0.41	0.90
/3	5.69	2183	2.77	21	5.66	4.92	1.01
/4	−2.45	2184	–	–	−2.45	−0.03	1.20
Mean	2.86	–	–	–	2.70	3.08	1.31

The table reports the average quarterly returns for funds that survived until the end of 1994, average quarterly returns for funds that ceased to exist during 1989–1994, and average returns for the combined sample. Furthermore, we present quarterly returns for the Standard and Poor 500 index and Treasury bills. The columns labelled ‘Number’ contain the numbers of funds over which the averages are calculated.

Poor 500 index and the returns on a three-month Treasury Bill over the same period.

In accordance with other studies, like Malkiel (1995) and Blake and Timmermann (1998), it appears that in almost all quarters the surviving funds had a higher average return than the non-surviving funds, indicating that low returns increase the probability of disappearance. Furthermore, the average annual return over the period 1989–1994 for the sample containing the surviving funds is 0.64% higher than for the combined sample of funds, i.e. 11.44% vs. 10.80%. In contrast, Malkiel (1995) finds a difference of 1.50% over the period 1982–1991, Elton et al. (1995) even find a difference of 1.87% for 1976–1993, while Brown and

Table 3
Summary statistics by objective category

Group objective	Combined		Surviving funds		Drop out %
	Mean return	Number	Mean return	Number	
Aggressive growth	3.70	95	3.88	77	18.9
Growth/small companies	3.12	1052	3.22	904	14.1
Income/growth–income	2.49	540	2.59	434	19.6
Specialty	2.51	273	2.86	201	26.4
Foreign/world	2.04	490	2.18	429	12.4
Other	2.05	226	2.17	138	38.9

The table shows for six investment objectives categories the average quarterly returns for the combined sample and the average quarterly returns for the funds that survived until the end of 1994, as well as the number of funds in each category, the number of non-survivors and the corresponding drop-out percentage.

Goetzmann (1995) report a difference of 0.80% over the period 1977–1987. Note that all estimates for this survivorship effect in average returns are higher than the at most 0.40%, that was claimed by Grinblatt and Titman (1989).

In order to examine whether survival rates vary by investment objective, we broke down the sample by investment styles. While Morningstar reports investment objectives in thirteen different categories,[4] we chose to follow other studies (Malkiel, 1995; Brown and Goetzmann, 1995), and decided to split the sample, for ease of comparison, into six categories. The category ‘other’ represents the equity funds that could not be clearly assigned to any of the other five categories, so it contains, for instance, the funds that advertise as equity fund and invest less than 50% in fixed income securities as well as the funds with unknown investment objective. Table 3 presents the average quarterly return over the period 1989–1994 for all funds as well as for the subset of funds that survived until the end of 1994 for each of the six investment objectives.[5]

It appears that the categories ‘specialty’ and ‘other’ had the highest percentage of non-survivors. Moreover, the difference between the average annual returns for the ‘specialty’ category is 1.40%, which is much higher than the 0.64% for the aggregated sample of mutual funds. Because there is no a priori reason why survival would be described by the same model for each of the investment objectives, we shall estimate different survival models for different objectives. Because ‘growth’ and ‘income’ are the most popular investment styles, the sequel of the paper focusses on funds with these two styles only.

[4] For 79 funds, Morningstar did not report an investment objective. For 33 funds of this group, the investment objective was obvious from their name. The remaining 46 funds were classified as having an ‘other’ investment objective.

[5] Note that a fund’s investment objective is self-reported and can therefore easily lead to gaming to improve relative ex-post return rankings (see Brown and Goetzmann, 1997).

3. What determines mutual fund survival?

In the previous section, we noted that on average mutual funds that leave the sample have lower returns. Moreover, most of the disappearance occurs at a relatively young age, indicating that a bad record of returns in the first few years of its existence seriously decreases a fund's survival probabilities. It can also be noted that in particular months fund attrition is much larger than can be expected on the basis of observed returns. To account for this, we include a common time effect in our specification. Let y_{it} be an indicator variable that indicates whether or not fund i has an observed return in period t. Our specification describes the probability of fund survival $(y_{it}=1)$ using a longitudinal probit model, such that a fund survives if an underlying latent variable, y_{it}^* is positive. That is,

$$y_{it}^* = \alpha + \sum_{j=1}^{J} \gamma_{ij}\left(r_{i,t-j} - \theta\right) + \phi \text{age}_{i,t-1} + \lambda_t + \eta_{it} \tag{1}$$

$$y_{it} = 1 \quad \text{if fund } i \text{ is observed in quarter } t \ \left(y_{it}^* > 0\right)$$

$$y_{it} = 0 \quad \text{otherwise}$$

where $r_{i,t-j}$ is the return of fund i in quarter $t-j$, θ is an unknown constant, $\text{age}_{i,t-1}$ is the time in years since fund inception, and λ_t denote fixed time effects describing economy wide effects. Because the number of lagged returns in the specification is not the same for all funds at all times (some γ_{ij} may be zero, see below), the inclusion of θ prevents that the unconditional mean of Eq. (1) by construction depends upon the number of returns that is included. The inclusion of time dummies ensures a robust specification that does not require us to specify which economy wide variables are responsible for overall changes in survival rates. These time effects may thus incorporate financial variables, like the returns on one of more benchmark portfolios or a yield spread, as well as macro variables, like an indicator for the state of the business cycle, as long as they affect survival of each fund in a similar fashion. This way, we do not specify a priori whether absolute returns or returns in excess of some benchmark portfolio are determining fund survival. Moreover, the time effects incorporate the dependence in survival probabilities across different funds.

The error term η_{it} is assumed to be standard normally distributed, independently over funds and periods, i.e. $\eta_{it} \sim NID(0,1)$. The γ coefficients measure the impact of historical returns and, potentially, vary over funds and lags. To avoid dimensionality problems, we assume that the γ_{ij}'s can be described by a third-order polynomial in j. The advantages of a polynomial lag structure are that only a limited number of parameters have to be estimated, an increasing precision of the estimates, and a smooth pattern on the coefficients that is automatically imposed. To prevent that the model only applies to funds that have a return history of at least J quarters, and is thus conditional upon having survived these J quarters, we employ a flexible parametrization of the effects of lagged returns such that the model is conditional upon the observed return history only. Let m_{it} denote the number of lagged quarterly returns that is available for fund i in quarter t, with a maximum of J. Then we assume the following structure for the lagged quarterly returns coefficients[6]

$$\gamma_{ij} = \delta(m_{it}) \sum_{k=0}^{3} a_k j^k I(j \leq m_{it}), \tag{2}$$

where $\sum_{k=0}^{3} a_k j^k$ is a polynomial of degree three, and $I(\,)$ is the indicator function that equals 1 if j is smaller than or equal to m_{it} and 0 otherwise. The factor $\delta(m_{it})$ equals 1 if $m_{it} = J$ (all lagged returns are observed) and we expect it to be larger than unity if $m_{it} < J$. Experimenting with alternative specifications for $\delta(m_{it})$ did not lead to very different results. In the table we report results using $\delta(m_{it}) = 1 + \xi \ \ln(J + 1 - m_{it})$, where ξ is an unknown parameter to be estimated. Note that for mutual funds with a return record of more than J quarters, the lagged quarterly returns coefficients can simply be described by $\sum_{k=0}^{3} a_k j^k$. As it is implicitly assumed that further lags of the returns are irrelevant, we restrict the polynomial coefficients such that the hypothetical coefficient for lag $J + 1$ is zero. This gives an additional restriction that can be substituted into Eq. (1) and reduces the number of parameters describing variation in γ_{ij} to four. It should be noted, because of the presence of the time effects and the truncation of m_{it}, that both ξ and θ are only identified from information contained in funds that exist less than 12 quarters.

For both investment styles 'growth' and 'income' we estimate Eq. (1) with four ($J = 4$), eight ($J = 8$) and twelve ($J = 12$) lagged quarterly returns included, over the period 1989/01–1994/04. The three specifications are tested against each other and against more general alternatives. Note that the number of parameters in each of the models is the same. To test whether the inclusion of additional lags would improve the models, we applied variable addition tests to the three specifications. These tests are Lagrange Multiplier tests for the null hypotheses that the coefficients for one or more additional lags, added unrestrictedly to the model, are zero. More details about this and subsequent tests are provided in Appendix A.

Panel A of Table 4 presents the outcome of tests for the inclusion of additional lagged returns. Clearly, the specification with $J = 4$ is overly restrictive and has to be rejected against alternatives with additional lags. The specification with $J = 8$ is rejected at the 5% level for 'income' funds but not for 'growth' funds. For the

[6] For notational simplicity, the fact that the γ_{ij} coefficients vary over time as a function of m_{it} (for a subsample of the funds) is not reflected in their indices.

Table 4

Panel A: Variable addition tests

Specification	Additional variable(s)	*df*	Growth		Income	
			LM-stat	*p*-val	LM-stat	*p*-val
$J=4$	r_{t-5}	1	6.46	0.011	2.73	0.098
	$r_{t-5},\dots,r_{t-8}$	4	14.49	0.006	6.06	0.195
$J=8$	r_{t-9}	1	3.60	0.058	4.61	0.032
	$r_{t-9},\dots,r_{t-12}$	4	5.57	0.234	4.77	0.312
$J=12$	r_{t-13}	1	0.00	0.999	2.05	0.152
	r_{t-13}, r_{t-14}	2	0.10	0.951	2.07	0.355
	age_{t-1}^2	1	4.72	0.030	1.77	0.183
	$\sqrt{\text{age}_{t-1}}$	1	1.84	0.175	2.65	0.104

Panel B: Heteroskedasticity tests

Specification	Variable(s)	*df*	Growth		Income	
			LM-stat	*p*-val	LM-stat	*p*-val
$J=12$	$r_{t-1}\dots r_{t-12}$	3	1.64	0.650	1.13	0.770
	$r_{t-1}\dots r_{t-12}$, age_{t-1}	4	1.65	0.800	1.47	0.832

Panel C: Nonnormality test

Specification	*df*	Growth		Income	
		LM-stat	*p*-val	LM-stat	*p*-val
$J=12$	2	0.46	0.795	1.49	0.475

The table reports outcomes of Lagrange multiplier tests for omitted variables, heteroskedasticity and nonnormality in the probit model for fund survival.
All test statistics have an asymptotic null distribution that is Chi-squared with degrees of freedom given by *df*. A * indicates rejection at the 5% level; see Appendix A for details.

models with 3 years of quarterly returns ($J=12$), it cannot be rejected that further lags have zero coefficients. As it is well known that violation of the assumption of homoskedastic error terms typically leads to inconsistency of the maximum likelihood estimators in limited dependent variable models (see, e.g. Amemiya, 1985, p. 268 ff.), we also test the specification with $J=12$ against heteroskedastic alternatives, the error variance being functions of lagged returns and fund age. The results of the Lagrange Multiplier tests, presented in panel B of Table 4, do not cause any doubt on the validity of the homoskedasticity assumption. Another crucial assumption is that of normality, which we tested against the more general Pearson family of distributions, as described by Newey (1985). As shown in panel C of Table 4, we are not able to reject normality either. Finally, we performed a number of tests checking the linear way in which fund age is included against more general alternatives. This involves tests on the inclusion of nonlinear functions of age, which do not reject for the square root of age and only

Table 5
Estimation results

	Growth funds		Income funds	
	Estimate	Standard error	Estimate	Standard error
α	3.430	0.282	4.102	0.619
r_{t-1}	0.014	0.007	0.050	0.012
r_{t-2}	0.022	0.004	0.035	0.007
r_{t-3}	0.026	0.004	0.024	0.006
r_{t-4}	0.028	0.004	0.016	0.006
r_{t-5}	0.028	0.004	0.012	0.007
r_{t-6}	0.026	0.003	0.010	0.007
r_{t-7}	0.023	0.003	0.010	0.006
r_{t-8}	0.019	0.004	0.011	0.006
r_{t-9}	0.015	0.004	0.011	0.005
r_{t-10}	0.010	0.005	0.011	0.005
r_{t-11}	0.006	0.004	0.009	0.005
r_{t-12}	0.002	0.003	0.006	0.003
θ	4.932	0.641	10.168	3.811
ξ	−0.132	0.092	0.086	0.075
age_{t-1}	0.020	0.008	0.036	0.009

The table presents estimation results for probit specification (1) with 12 ($J=12$) lagged quarterly returns, age of the fund (in years) and 24 time dummies as explanatory variables. We do not report the estimates for the polynomial coefficients, but only report the implied estimates for the lagged quarterly returns under the condition that a fund has more than J quarterly returns available. Note that for funds with less than J historical returns available, the coefficients for the lagged return should be inflated by a factor (see main text). The total number of observations for funds with investment style 'growth' is 14,449, for the 'income' style it is 8055.

marginally reject (at the 5% level) for age squared in the sample of growth funds. Because it may be the case that past performance matters less for older funds (as suggested by Chevalier and Ellison, 1997), we also tested, by means of a likelihood ratio test, the current specification against one in which the γ_{ij} coefficients of older funds are different from those of young funds (that exist 5 years or less). For neither of the two fund types, the null hypothesis could be rejected.[7]

Table 5 reports the estimation results for the survival processes (with $J=12$) for funds with investment styles 'growth' and 'income'. The results indicate that lagged quarterly returns, age of the fund and the aggregate time effect have a significant effect on fund disappearance (at the 5% level). Low returns for a number of consecutive quarters significantly increase the probability of leaving the

[7] The values for the test statistics are 2.76 and 4.23, for 'income' and 'growth' funds, respectively, while the 5% critical value is 7.81.

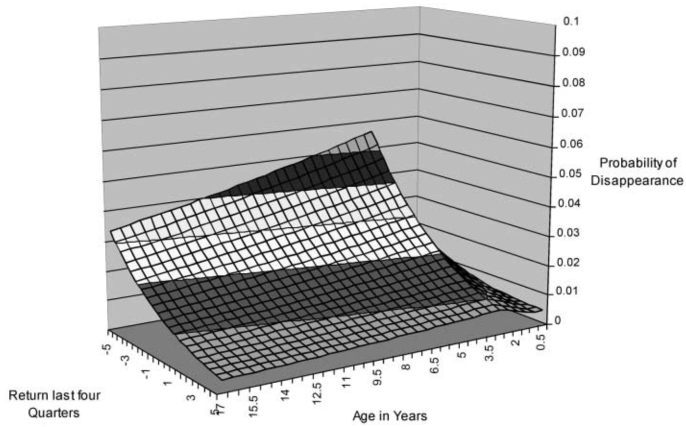

Fig. 1. Non-survival probabilities for growth funds. The figure shows the probability of disappearance for different years since fund inception (age) and different historical returns, according to the estimated survival model. The returns for each of the last four quarters, i.e. $r_{t-1},\ldots, r_{t-4}$, vary from -5% to $+5$%, while the returns of the eight preceding quarters ($r_{t-5},\ldots, r_{t-12}$) are fixed at 3% per quarter. The probabilities are averaged over the 24 different quarters.

sample. Interestingly, the pattern of estimated return coefficients for income funds is monotonically decreasing, while that for growth funds is hump-shaped, indicating that recent past performance is relatively more important for the survival of income funds than for growth funds. For both types of funds, the parameter ξ is not significantly different from zero. This indicates that the absolute weights of recent historical returns do not increase for funds with a short history. Put differently, the returns in, say, the last four quarters are equally important irrespective of whether the fund has a history of just these four or more than 12 quarterly returns. The positive coefficients for age indicate that, ceteris paribus, the older the fund, the more likely it is to survive. The estimates for the fixed time effects (not reported) indicated that the probability of fund disappearance varies significantly over the quarters, even if returns have not changed.

Using the estimates of the funds with investment style 'growth', we can compute the probability that an arbitrary growth fund will disappear in the next quarter given its past record of returns and its age.[8] In Fig. 1, we show the probability of disappearance for growth funds with different ages, where the past

[8] Because of the inclusion of the fixed time effects, probabilities are also conditional upon aggregate market movements. As these may not be observed in advance, the probabilities of survival do not correspond to ex-ante probabilities, as employed in Brown et al. (1995). Eliminating biases, which is an ex-post exercise, does not require knowledge of these ex-ante probabilities, but rather the survival probabilities conditional upon other observables.

record of returns varies from -5% to $+5\%$ for each of the last four quarters and the quarterly returns for the quarters $t-5$ through $t-12$ are fixed at 3.00%, corresponding to the average quarterly return over the period 1989–1994. The probabilities are averaged over the 24 different quarters.[9] It appears that, for instance, a 3-year-old fund with a return of -5% for each of the last four quarters has a probability of almost 5.5% to disappear in the next quarter, while a 16-year-old fund with a comparable return record only has a probability of almost 3% to disappear. The probability of attrition drops below 1% if the 16-year-old fund had a return of $+5\%$ for the last four quarters.

The signs we find for the estimated coefficients in the survival model are in accordance with the results of Brown and Goetzmann (1995). Our specification can be interpreted as a reduced form specification of their model that also includes the size of the fund and the expense ratio as explanatory variables. While we do not observe the size of a fund during the entire sample period, it has been well documented (see, e.g. Rockinger, 1995; Sirri and Tufano, 1998) that (relative) historical returns are key determinants of capital flows to mutual funds. In contrast to our reduced form model, Brown and Goetzmann (1995) did not include neither time effects nor test for their presence. It is important to allow for fixed-time effects to incorporate common aggregate shocks that affect the survival of all funds, such as bad returns on the stock market as a whole. Omitting the time effects, which may be correlated with the regressors, yields inconsistent parameter estimates (see, e.g. Baltagi (1995, p. 178 ff.)) and inappropriate bias corrections. The model with fixed time effects (or any other fund-invariant variables) can only be estimated consistently for the number of funds tending to infinity. We shall therefore assume that the cross-section of funds is sufficiently large.

4. Spurious persistence patterns due to look-ahead bias

In order to get an impression of the effect of look-ahead bias in a survivorship free sample, we consider an example where interest lies in the estimation of fund α's and their persistence. Our Monte Carlo simulation experiment follows the set-up of Brown et al. (1992). In particular, we generate quarterly returns from the following one factor model

$$r_{it} = r_{ft} + \beta_i(R_{mt} - r_{ft}) + u_{it} \tag{3}$$

where r_{ft} is the short term interest rate and $R_{mt} - r_{ft}$ is the quarterly excess return on the market portfolio. The idiosyncratic error term u_{it} is independent of the

[9] A similar picture can be drawn for the funds with investment style 'income'.

quarterly risk premium and is assumed to be normal with mean zero and variance σ_i^2, given by

$$\sigma_i^2 = k(1-\beta_i)^2. \tag{4}$$

This relationship is a rough approximation to the relationship between non-systematic risk and β that is often observed in mutual funds data.[10] We employ a set of parameter values in the return generating process that closely matches the first two sample moments of returns and beta. The quarterly excess return on a market portfolio is i.i.d. normal with mean 0.0215 and standard deviation 0.104, β is i.i.d. normal with mean 0.93 and standard deviation 0.37 and the value of k equals[11] 0.01997. Moreover, we assume that the short term interest rate can be described by an AR(1) model given by[12,13]

$$r_{ft} = \mu + \rho(r_{ft-1} - \mu) + \epsilon_t, \epsilon_t \sim NID(0,\sigma_\epsilon^2), \tag{5}$$

where ϵ_t is independent of u_{it} (for all i).

The simulation experiment proceeds as follows. We start with a 'random' number of funds such that an average 2% increase per quarter leads to 2500 mutual funds in the final quarter. This leads to a sample where the number of funds increases each quarter, while none of the funds drops out. Next, we apply the survival model of the previous section, as estimated for growth funds, to determine for each fund in each period the probability that it leaves the sample. This means that from the record of historical returns, the age of the fund and an aggregate time effect, a probability $\hat{p}_{it}$ of leaving the sample in the next quarter is determined. Then fund i leaves the sample in period t with a probability $\hat{p}_{it}$. Note that this assumes that η_{it} is independent of current returns, that is, the probability of survival only depends upon age and historical returns, and— conditional upon those—not upon current returns. Since the age of the fund, defined as the years since fund inception, is a significant factor in fund disappearance, we decided to draw a random age for the funds that already exist in the first quarter, closely corresponding to the observed age distribution in our sample of mutual funds, i.e. $\text{age}_1 \sim \text{abs}(N(0,16))$. The survival process used in our simulations is thus more complicated but also more realistic than the single or multi-period survival rules applied by Brown et al. (1992), Hendricks et al. (1997) and Carpenter and Lynch

[10] Since it can be expected that fund returns are highly correlated due to style effects and herding, we also conducted a simulation experiment where we take into account contemporaneous correlations across residual returns of funds. This experiment follows the set-up suggested by Brown et al. (1992) and Hendricks et al. (1997). The results are very similar to those reported here and therefore omitted.

[11] The value of k that we employ is based on the sample average of $\hat{\sigma}_i^2/(1-\hat{\beta}_i)^2$.

[12] An OLS estimation of the short rate AR(1) process over the period 1976–1994 yields $r_{ft} = 0.140 + 0.922 * r_{ft-1} + \upsilon_t$, with $\hat{\sigma}_\upsilon = 0.003$.

[13] In the simulations, the average T-bill return over 1976–1994 of 0.018 is used to start the process. Moreover, if a risk free rate smaller than zero happens to be generated, we set it equal to zero.

(1999), who simply remove, for instance, the worst performing 10% of the mutual funds in each period.

Note that the estimated time effects in specification (1) reflect aggregate shocks. We take the potential dependence between the time effects and observed aggregate variables in the model into account by running a number of regressions on variables such as the return on the S & P500 and the return on a three-month Treasury Bill over the period 1989–1994. It appears that the time effect is significantly correlated with the risk-free return on Treasury bills. Accordingly, the random effect that is used in the simulations is drawn from a truncated normal distribution[14] with a mean that is linearly dependent on the risk free rate. Note that a high risk-free rate leads to, on average, higher time effects, but the effect on the survival rate might be balanced by higher (nominal) returns. The numbers presented below refer to averages or standard deviations over 1000 replications. In order to prevent sensitivity to the starting conditions, fund returns are generated from quarter 1 onwards, while the survival process starts operating from quarter 13 and further.

We now construct two different samples. In the first sample, no funds are eliminated, so that all funds, entering at different points in time, have observed returns until the last period. This corresponds to a survival probability of 100% for all funds at all times. We will refer to this sample as the hypothetical full sample and we will only use it in a few cases. The second sample consists of the surviving funds completed with observations on those funds that left the sample before the last quarter. We refer to this sample as the survivorship free sample. Most recent empirical studies (e.g. Brown and Goetzmann, 1995; Carhart, 1997a; Wermers, 1997; Blake and Timmermann, 1998) employ such survivorship free samples.

Empirical studies by Brown and Goetzmann (1995), Malkiel (1995) and Carhart (1997a) examine whether 'winning' mutual funds, where winning is defined as exceeding the median fund return in a given period, are more likely to be winners in the next period. Studies of Brown et al. (1992) and Hendricks et al. (1997) show that survivorship bias in combination with look-ahead bias induces spurious persistence patterns if there is cross-sectional variation in expected returns or risk. Instead of hypothesizing a certain survival process, we use the empirical survival model that matches the appropriate subsample of US equity funds, to redo the calculations of Hendricks et al. (1997), who found a spurious *J*-shape pattern in performance persistence that was due to survivorship bias and look-ahead bias. As we generated fund returns such that any abnormal return is the result of (unpredictable) luck, and moreover, made use of a survivorship free

[14] The estimated mean that is employed is $2.20+0.47r_t^f$, while the variance is 0.36. The random effect is truncated at 1.7 (approximately two standard deviations below the mean) to prevent unrealistically high attrition rates.

sample, any regularity found in performance persistence measures is necessarily spurious and due to look-ahead bias.

The performance of the funds is evaluated by estimating Jensen's α from the one-factor model in Eq. (3), over four 3-year periods. We sort the funds on the basis of the estimated α's in each 3-year period into eight groups. For each octile group, we calculate the average Jensen's α in the subsequent 3-year period. Table 6 presents the average α for each group for the survivorship free sample, containing all funds until the period of disappearance, and the hypothetically full sample, not affected by attrition.

It appears that in our simulation experiment, the average α's in the two extreme octiles (i.e. octile one and eight) are biased upward by about 0.8% on an annual basis, indicating that the survivorship free sample exhibits a pattern of spurious persistence in performance. Clearly, the fact that the data are survivorship free does not imply that standard methods of analysis do not suffer from ex-post conditioning biases. Although the spurious persistence pattern is not exactly *J*-shaped, the simulation results more or less confirm the biases found by Hendricks et al. (1997), although their study also suffers from survivorship bias. The argument for such a pattern is a risk argument. Funds in one of the extreme ranks are more likely to be 'high risk' funds and thus less likely to survive. Conditional on the fact that they did survive in the second subperiod, they will have made better returns than average. Compared to Hendricks et al. (1997), we find an additional upward bias in Jensen's α for the lower octiles. The reason for this is that our survival process is dynamic. Funds with a low rank realized relatively bad returns in the first 12 quarters. As this will decrease a fund's probability to survive over the next 12 quarters, these funds must have made up for these bad returns given that they have survived. Apparently, with our parame-

Table 6
Simulated performance persistence

Initial period rank	Performance subsequent period			
	Survivorship free sample		Hypothetical full sample	
1 (losers)	0.196	(0.010)	−0.000	(0.006)
2	0.078	(0.007)	0.004	(0.004)
3	0.047	(0.005)	0.002	(0.003)
4	0.022	(0.004)	0.000	(0.002)
5	0.031	(0.004)	0.001	(0.002)
6	0.056	(0.005)	0.000	(0.003)
7	0.082	(0.007)	−0.001	(0.004)
8 (winners)	0.160	(0.009)	−0.000	(0.006)

The table presents the subsequent period performances for two types of samples. A survivorship free sample that contains the non-surviving funds until the quarter of disappearance and for the hypothetical full samples that is not affected by survivorship. Results are averages over 1000 replications. Monte Carlo standard errors in parentheses.

ter values, this effect is large enough to change the *J*-shape into a (more or less) *U*-shape. As expected, the hypothetical full sample that is not affected by survivorship does not exhibit a spurious persistence pattern.

5. Eliminating look-ahead bias

Knowledge of the survival process is a key to correcting for the effects of look-ahead bias. In Section 3, survival of a fund was modelled as a function of historical returns, age and an aggregate time effect. We will show, in this section, how inferences can be corrected for survivorship effects if it can be assumed that fund survival in period t is independent of the return in period t, after conditioning upon lagged returns, fund age and time.[15] Technically, this imposes that η_{it} in Eq. (1) is independent of r_{it}, as was assumed in the previous section. The corrections, based upon work of Fitzgerald et al. (1998), are relatively simple to apply and involve the use of weights. As these weights depend upon fund returns, they are endogenous and their use has implications for consistency of the employed estimators.

In general, let R_i denote a vector of returns for fund i that is used in an empirical analysis, for example in constructing a contingency table. Let $Y_i = 1$ if fund i is used in the analysis and 0 otherwise. While Y_i is determined by the researcher, we shall assume that it is a function of y_{it}'s only. The distribution of returns for the funds used in the analysis, conditional on some observed characteristics X_i, is described by $f(R_i | X_i, Y_i = 1)$, where f is generic notation for a density/probability mass function. Because Y_i is not independent of the returns, this distribution differs from the one unconditional upon survival $f(R_i | X_i)$, which is what we are interested in. The set of variables in X_i is chosen by the researcher and could reflect a fund's investment style, its return history, its relative historical performance or could be empty. Let Z_i denote observable fund characteristics that affect the probability of survival. Then it follows, using standard conditioning arguments (see Fitzgerald et al., 1998), that we can write

$$f(R_i, Z_i | X_i) = w_i f(R_i, Z_i | X_i, Y_i = 1), \tag{6}$$

where w_i is a weight factor given by

$$w_i = \frac{P\{Y_i = 1 | X_i\}}{P\{Y_i = 1 | R_i, X_i, Z_i\}}. \tag{7}$$

This weight equals the inverse of the normalized probability that fund i is kept in the sample for funds of type X_i. The left hand side of Eq. (6) provides the

[15] Econometric approaches of sample selection and attrition problems based upon the work of Heckman (1979) and Hausman and Wise (1979), assume that the model of interest is conditional upon the same set of variables, which is inappropriate in this case.

(un)conditional distribution of returns we are interested in. The right hand side is the conditional observable distribution of returns times a weight factor. If the weights w_i are known, any inference based upon the observed distribution of returns can directly be adjusted to reflect the unconditional distribution. For example, the expected returns of fund i satisfy

$$E[r_{it}] = \int R_i f(R_i, Z_i) \mathrm{d} R_i \mathrm{d} Z_i$$

$$= \int w_i R_i f(R_i, Z_i \mid Y_i = 1) \mathrm{d} R_i \mathrm{d} Z_i = E[w_i r_{it} \mid Y_i = 1], \quad (8)$$

which implies that the average of $w_i r_{it}$ rather than the average of r_{it} provides an unbiased estimate of the fund's mean return if r_{it} is observed if $Y_i = 1$ only. Similarly, a fund's α can be estimated as

$$\tilde{\alpha}_i = w_i \hat{\alpha}_i, \quad (9)$$

where $\hat{\alpha}_i$ is the usual (uncorrected) ordinary least squares estimate. This follows immediately from the fact that $\hat{\alpha}_i$ is a linear function of the returns of fund i.

Going back to our subsamples of US equity funds, suppose we are interested in performance as measured by α over a period of 12 quarters. This implies that we can only use funds in the analysis that have observed returns for 12 consecutive periods s to $s+11$, and we have that $Y_i = \prod_{t=s}^{s+11} y_{is}$. The probability that $Y_i = 1$ given the fund's returns R_i and characteristics Z_i, X_i is then described by our survival model, provided that X_i is included in the model (or can be assumed to have a zero coefficient), and provided that we assume that, conditional upon historical returns, Z_i and X_i, the probability of attrition in any given period does not depend upon (potentially unobserved) returns in that or future periods. In that case, we can write[16]

$$P\{Y_i = 1 \mid R_i, Z_i, X_i\} = \prod_{t=s}^{s+11} P\{y_{it} = 1 \mid r_{i,t-1}, \dots, \mathrm{age}_i, X_i\}. \quad (10)$$

If X_i denotes investment style, the latter probabilities are directly obtained from the survival model in Eq. (1) estimated over the appropriate subsample of funds. If X_i includes both investment style and historical performance, specification (1) can be used assuming that survival depends upon historical performance only through historical returns. The numerator in Eq. (7) reflects the probability of survival for a given X_i. When X_i denotes investment style, this can easily be estimated by the ratio of the appropriate number of funds that survived from period s to $s+11$ and number of funds that was in the sample in period $s-1$. If X_i includes a fund's initial performance ranking, this computation has to be done for each ranking separately. Together, this provides estimated weights $\hat{w}_i$ that are consistent for $N \to \infty$. Consequently, we can estimate the α of an individual fund asymptotically

[16] This assumes that η_{it} in Eq. (1) exhibits no autocorrelation.

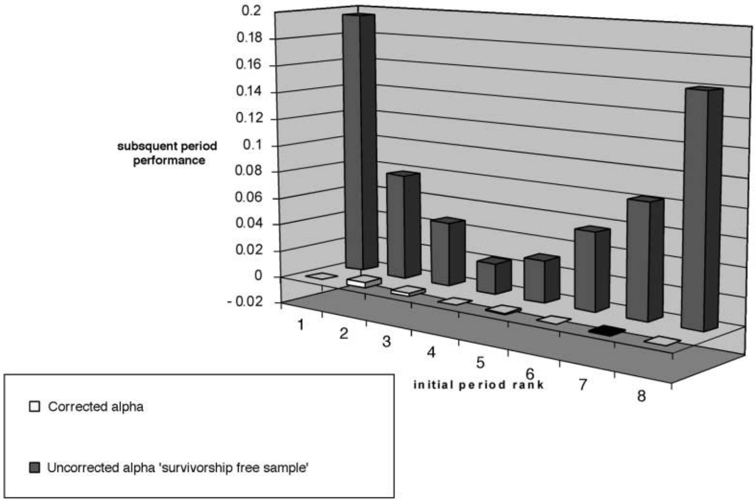

Fig. 2. Simulation results look-ahead bias correction on alpha. The figure shows the subsequent period performances for simulated survivorship free samples with performance correction for look-ahead bias (corrected α) as well as without correction for look-ahead bias (uncorrected α 'survivorship free sample').

unbiasedly[17] using Eq. (9) with $\hat{w}_i$ instead of w_i. Because the survival probabilities are necessarily conditional upon any fixed time effects, the above correction can also be applied in the case of conditional performance measurement, such as advocated by Ferson and Schadt (1996).

In order to illustrate the use of the proposed correction method, we apply it to the simulated samples generated in Section 4. Recall that in the performance persistence analysis, we found a spurious *U*-shape pattern in risk-adjusted returns of the formed octiles in the survivorship free sample. To correct Table 6 for this bias, we first estimate initial period α's and their ranking, using funds that have observations over 12 consecutive quarters, and correct these OLS estimates using weights based upon these 12 periods. Next, for the α's estimated over the second subperiod, which are conditional upon survival of another 12 quarters, we apply a weight correction based upon survival over these 12 periods. As interest lies in the distribution of α's conditional upon a fund's initial rank, the numerators in the weights correspond to survival probabilities for a given octile.

In Fig. 2 we present the average corrected Jensen's α in the subsequent 3-year period for each octile group, where octile one represents the worst performing funds of the initial period. For comparison, the figure also contains the uncorrected

[17] Asymptotically unbiased means that the expected value of the estimator equals the true value if the number of funds N goes to infinity.

results for the survivorship free sample, as given in Table 6. It is quite clear that the spurious persistence pattern that was present in the survivorship free sample of funds has disappeared. While not reported, the standard errors show that the average corrected Jensen's α's of the octile groups are not significantly different from zero anymore. Furthermore, while the Monte Carlo results show how the spurious persistence pattern can be eliminated if there is no genuine persistence, the correction with weights can also be applied to obtain estimates of performance persistence in observed mutual fund returns that do not suffer from biases.

6. Estimating persistence in US equity fund returns

A substantial number of empirical papers report persistence in the performance of mutual funds over 1 to 3-year horizons, see, e.g. Hendricks et al. (1997), Gruber (1996), Carhart (1997a) or Wermers (1997). Mostly these papers suggest that their results are free of survivorship bias and no attempts are made to correct for potential look-ahead biases. In this section, we address the question of short-term and medium-term predictability of mutual fund performance correcting for look-ahead bias using the methodology discussed in the previous section.

As our sample of attrited funds goes back to only the first quarter of 1989, we can estimate survival probabilities only over the period 1989/1–1994/4, and our survivorship bias free sample is limited to equity funds over this period. While this sample period is relatively short, we do not feel that this is a serious drawback. Given the huge increase in the number of mutual funds, and their holdings, over the last decade, it may be hard to argue that information from the 1960s and 1970s is relevant for current investment decisions. Thus, we prefer to work with a short but recent time period, with a fairly large cross-sectional dimension, as required by our methodology.

First of all, we examine persistence in fund returns over a short-term and a medium-term horizon, without any risk corrections. In the first case, we sort the funds in the ranking period on the basis of average returns over a 1-year period, while in case of a medium-term horizon the funds are ranked on the basis of average returns over a 3-year period. Next, we compute the average return over the next 1- or 3-year period for all funds in a given octile. These average returns are biased because they are conditional upon survival. To correct for look-ahead bias, we employ the estimated survival probabilities as described by model (1). Using the technique of Section 5, we correct the average return $\bar{r}_i$, over quarters s to $s+H$, using

$$\tilde{r}_i = \frac{\hat{q}_s}{\prod_{t=s}^{s+H} \hat{p}_{it}} \bar{r}_i, \tag{11}$$

where $\hat{p}_{it}$ is the estimated probability that fund i leaves the sample in period t, and $\hat{q}_s$ is the ratio of the number of funds that survived from period s to $s+H$

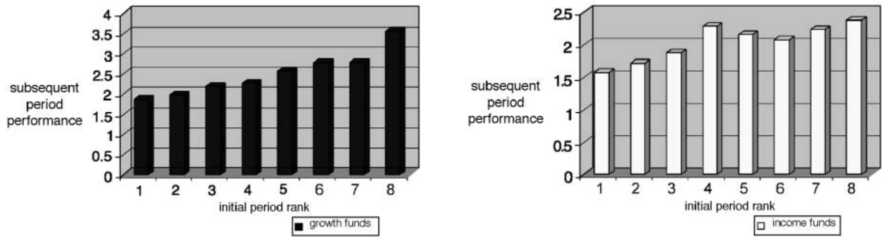

Fig. 3. Empirical results 1 year persistence in returns, corrected for look-ahead bias. The figures show the subsequent period performance for the funds with investment objective: growth and income. All results are corrected for look-ahead bias using the procedure outlined in the text.

(where $H=3$ for a 1-year horizon and $H=11$ for a 3-year horizon) and the number of funds that was in the sample at $s-1$. In the next step, we sort the funds into octiles on the basis of the corrected average return $\tilde{r}_i$. For the subsequent 1-year or 3-year period, we compute average returns, again correcting the estimates using Eq. (11), where $\hat{q}_s$ now corresponds to the proportion of survived funds in the corresponding octile. The results are summarized in Figs. 3 and 4.

It appears that for the funds with investment style growth a short-term persistence pattern is present, indicating that last year winners (octile eight) are more likely to have an above average return in the subsequent 1-year period than last year losers (octile one). During the period 1989/1–1994/4, a strategy of buying last year winners from the sample would have outperformed a strategy of buying last year losers with almost 1.69% on a quarterly basis (i.e. 6.76% annually). Note that the observed pattern strongly corresponds to the often found momentum effect in individual security returns (see, e.g. Chan et al., 1996). The persistence pattern in the funds with investment style income is less clear. Although a strategy of buying last year winners outperforms a strategy of buying

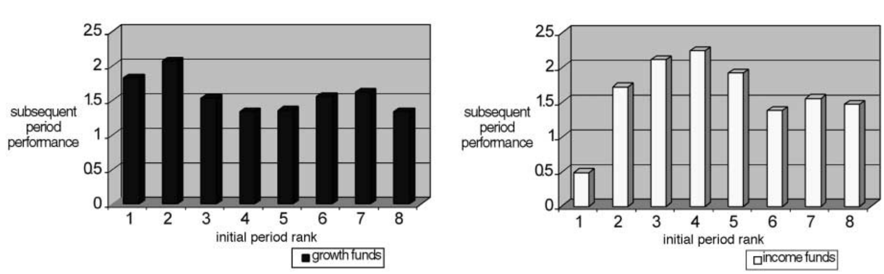

Fig. 4. Empirical results 3 year persistence in returns, corrected for look-ahead bias. The figures show the subsequent period performance for the funds with investment objective: growth and income. All results are corrected for look-ahead bias using the procedure outlined in the text.

last year losers with 0.8% on a quarterly basis, the results in intermediate octiles are indistinct. In case of a medium-term horizon, there is hardly evidence of performance persistence; see Fig. 4. For the sample of income funds, it is still the case that winners outperform losers, but for growth funds the opposite is found. Moreover, for both samples it is the case that intermediate octiles outperform the winners and losers. When we do not correct for look-ahead bias, we find similar results (not reported), indicating that look-ahead bias plays a minor role in the estimated short-term or medium-term persistence patterns in raw returns.

Next, we consider persistence in risk-adjusted returns (α's), defined in two different ways. First, we apply a simple one-factor model, given by

$$r_{i,t+1} - r_{f,t+1} = \alpha_i + \beta_i\left(r^m_{t+1} - r_{f,t+1}\right) + \varepsilon_{i,t+1}, \tag{12}$$

where $r_{i,t+1}$ is the return on mutual fund i in period $t+1, r^m_{t+1}$ is the return on the market portfolio in period $t+1$ and $r_{f,t+1}$ is the return on a risk free asset. Second, we use Carhart's (1997a) four-factor model given by

$$\begin{aligned} &r_{i,t+1} - r_{f,t+1} \\ &\quad = \alpha_i + \beta_{mi}\left(r^m_{t+1} - r_{f,t+1}\right) + \beta_{si} r^{\mathrm{smb}}_{t+1} + \beta_{hi} r^{\mathrm{hml}}_{t+1} + \beta_{pi} r^{\mathrm{pr1yr}}_{t+1} + \varepsilon_{i,t+1}, \end{aligned} \tag{13}$$

where r^{smb}_{t+1} is the difference between the returns on a portfolio of small stocks and one of big stocks, r^{hml}_{t+1} is the difference between the returns on a portfolio of high book-to-market and a portfolio of low book-to-market stocks and r^{pr1yr}_{t+1} is the difference between the return on a portfolio of stocks with the highest return over the previous year and a portfolio of stocks with the lowest return over the previous year.[18] We shall refer to α_i in Eqs. (12) and (13) as the Jensen's α.

The question we now try to answer is to what extent the ranking of a fund's α, over the subperiod 1989/1–1991/4, is informative about its α in the second subperiod 1992/1–1994/4, i.e. whether there is predictability on a medium-term horizon in risk-adjusted returns. We do so by first estimating Jensen's α's from Eqs. (12) and (13) over the initial 3-year period for all funds in the sample that survived these 12 quarters. Similar to the correction of average returns in Eq. (11), we correct the estimated α for look-ahead bias using

$$\tilde{\alpha}_i = \frac{\hat{q}_s}{\prod_{t=s}^{s+11} \hat{p}_{it}} \hat{\alpha}_i, \tag{14}$$

where $\hat{p}_{it}$ is the estimated probability that fund i leaves the sample in period t, and $\hat{q}_s$ is the ratio of the number of funds that survived from period s to $s+11$,

[18] We are very grateful to Mark Carhart for providing the data with returns on the market index, SMB portfolio, HML portfolio and PR1YR portfolio.

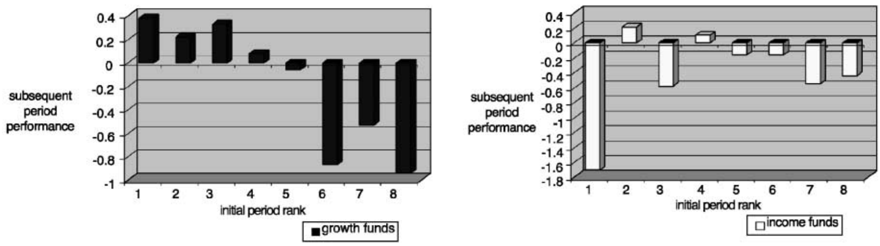

Fig. 5. Empirical results 3 year persistence in risk-adjusted returns, one factor model, corrected for look-ahead bias. The figure shows the subsequent period performance measured by a one-factor model for the funds with as investment objective: growth and income. All results are corrected for look-ahead bias using the procedure outlined in the text.

and the number of funds that was in the sample at $s-1$. In the next step, we sort the funds into octiles on the basis of the corrected Jensen's α. For the subsequent 3-year period, i.e. 1992–1994, we estimate α's, again correcting the OLS estimates using Eq. (14), where $\hat{q}_s$ now corresponds to the proportion of survived funds in the corresponding octile.

The results are summarized in Figs. 5 and 6. These figures present the average corrected Jensen's α in the subsequent 3-year period for each octile group of growth and income funds, where octile one represents the worst performing funds of the initial period. Fig. 5 is based on the one-factor model given in Eq. (12), while Fig. 6 corresponds to the four-factor model given in Eq. (13). Both pictures show the persistence patterns for the funds with investment objectives 'growth' and 'income'.

For the one-factor model, the sample of growth funds does not exhibit any positive persistence. Growth funds with a risk-adjusted return in the initial period below the median, have the highest Jensen's α in the subsequent period, and, on

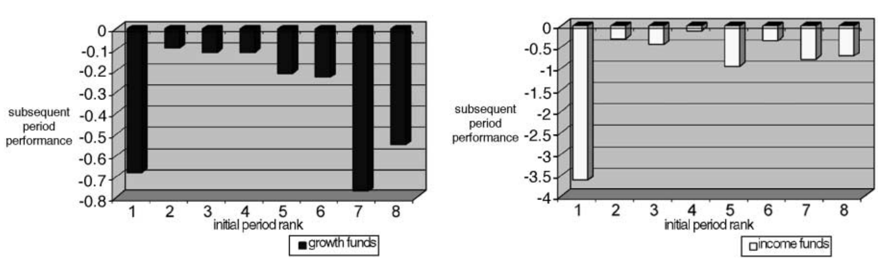

Fig. 6. Empirical results 3 year persistence in risk-adjusted returns, four-factor model, corrected for look-ahead bias. The figure shows the subsequent period performance measured by Carhart's four-factor model, for the funds with as investment objective: growth and income. All results are corrected for look-ahead bias using the procedure outlined in the text.

average, outperform the model by about 0.4% per quarter. On the other hand, the best performing growth funds of the initial period even have a negative average α in the evaluation period, corresponding to an underperformance of 0.93% per quarter. For the income funds, it is the case that the worst performing funds in the initial period persist in performing bad. However, the best performing funds of the initial period also have a negative risk-adjusted return in the evaluation period.

If we move away from the one-factor model and concentrate on Carhart's four-factor model, we find that the (reverse) persistence patterns have disappeared. There is no octile for any of the samples that significantly outperforms the model. One way to interpret this is that the four factors in Eq. (13) account for the reverse persistence pattern found in the one-factor model. It appears that the worst performing funds of the initial period are also the worst performing in the subsequent period, implying persistence of bad performance for this group of funds. For the funds with investment style 'growth' this means an underperformance of 0.7% per quarter, while the worst performing funds with style 'income' have an underperformance of more than 3.5% per quarter. The pattern we find for the four-factor model is in accordance with the one reported by Carhart (1997a) and do not support a hot hands phenomenon in mutual fund performance. However, in contrast to our findings, the best performing funds of Carhart's initial period have a slightly positive Jensen's α in the subsequent period. Moreover, the worst performing funds of Carhart's sample show less underperformance than the worst performing funds in the sample that we employ. There are two possible explanations for this difference in result. First of all, the difference in sample period, i.e. 1966–1993 vs. 1989–1994. Second, recalling the spurious persistence pattern we found in the simulation experiment, we found that the worst performing funds had a higher persistence bias than the best performing mutual funds. Since Carhart's methodology is not free from look-ahead bias, the difference in the two extreme octiles might be due to this effect. However, it appears that the uncorrected average Jensen's α's (not reported) are almost equal to the corrected average Jensen's α's, indicating once more that the effect of look-ahead bias in our sample is of minor importance.

7. Concluding remarks

In the recent literature, the presence of ex-post conditioning biases in empirical studies of performance persistence has been acknowledged. However, these studies were not able to determine the importance of these biases. This paper fills this gap by proposing and applying a fairly simple correction method that eliminates look-ahead bias. This correction method, which is based on the use of weights, requires information on the actual survival process that determines fund attrition. Consequently, in the first part of this paper we specified a longitudinal model for fund survival and estimated it for US equity funds with investment style 'income' and 'growth' over the period 1989/1–1994/4. This required us to extend the study of Brown and Goetzmann (1995), by examining which factors are important in determining mutual fund survival probabilities. From an extensive analysis of various specifications, it appeared that a specification with 12 lagged quarterly returns, fund age and aggregate time effects was to be preferred. The survival model was specified in such a way that it is not conditional upon the existence of a 3-year history of returns, so that it also models survival for 'young' funds.

In order to obtain insight in the effects of look-ahead bias in performance persistence measures, a number of Monte Carlo simulation experiments have been performed. By dropping funds from the sample based on the estimated survival probabilities, we analyzed the effect of survivorship on persistence in risk-adjusted returns, thus extending the analysis in Hendricks et al. (1997). Although the results are sensitive to the parameter values of the return generating process, we find that the sample that include returns of both survived and attrited funds, i.e. a survivorship free sample, is affected by look-ahead bias and generates a spurious persistence pattern. This is important, as it is often believed and suggested that such survivorship free samples are free of bias. With the dynamic survival model that was estimated, a spurious *U*-shape pattern was found in the persistence of risk-adjusted returns, similar to but different from the *J*-shape found in Hendricks et al. (1997).

Furthermore, we showed how inferences on mutual fund performance can be corrected for look-ahead bias using a simple weighting strategy, based upon the estimated survival model. A Monte Carlo experiment showed that the spurious pattern that arises in estimating performance persistence using traditional techniques, disappears with the correction that we propose. Finally, the approach was applied to equity funds with investment styles 'income' and 'growth' using a one-factor and Carhart's four-factor model. Within the latter model, we do not find any evidence for the period 1989–1994 that mutual funds that performed well in the past continue to perform well in the future. These results more or less confirm the results found by Carhart (1997a). Finally, it appears that the uncorrected results are of similar size, indicating that, for our data set, look-ahead bias is of minor importance.

Acknowledgements

Financial support by the Netherlands Organization for Scientific Research (N.W.O.) is gratefully acknowledged by the first author. We are grateful to Jon Danielsson, Wayne Ferson, Frank de Jong, Franz Palm (the editor) and two anonymous referees for helpful comments and suggestions, and to Mark Carhart for providing the data used in Section 6.

Appendix A. Misspecification tests in the probit model

In this appendix, we briefly indicate how the different misspecification tests for the probit model have been computed. In particular, we consider Lagrange multiplier (or conditional moment) tests for omitted variables, heteroskedasticity and nonnormality. More details can be found in, e.g. Newey (1985), Pagan and Vella (1989) or Verbeek (2000, Ch. 6, 7).

A.1. Variable addition tests

Let $\boldsymbol{x}_{it}$ denote the k-dimensional vector of explanatory variables in the probit model, including the time dummies. The loglikelihood function for the probit model is given by

$$\log L(\beta) = \sum_{i,t} y_{it} \log \Phi(\boldsymbol{x}'_{it}\beta) + \sum_{i,t}(1-y_{it})\log(1-\Phi(\boldsymbol{x}'_{it}\beta)), \tag{15}$$

so that the first order conditions can be written as

$$\sum_{i,t}\left[\frac{y_{it}-\Phi\left(\boldsymbol{x}'_{it}\hat{\beta}\right)}{\Phi\left(\boldsymbol{x}'_{it}\hat{\beta}\right)\left(1-\Phi\left(\boldsymbol{x}'_{it}\hat{\beta}\right)\right)}\phi\left(\boldsymbol{x}'_{it}\hat{\beta}\right)\right]\boldsymbol{x}_{it} = \sum_{i,t}\hat{\varepsilon}^{G}_{it}\boldsymbol{x}_{it} = 0, \tag{16}$$

where ϕ is the standard normal density function and Φ is the corresponding distribution function. The term in square brackets is referred to as the generalized residual (see Gouriéroux et al., 1987) and denoted $\hat{\varepsilon}^{G}_{it}$. The first order conditions thus imply that each explanatory variable is orthogonal to the generalized residual.

If r additional variables z_{it} were to be included in the model, it would not change the current estimates if the current estimates already satisfy the additional first order conditions. This means that if

$$\sum_{i,t}\hat{\varepsilon}^{G}_{it} z_{it} = 0 \tag{17}$$

then including z_{it} in the model would not change the current estimates. To test whether the lefthand side of Eq. (17) significantly differs from zero, we compute the Lagrange Multiplier test statistic as

$$\xi_{\mathrm{LM}} = \boldsymbol{\iota}'\boldsymbol{R}(\boldsymbol{R}'\boldsymbol{R})^{-1}\boldsymbol{R}'\boldsymbol{\iota} \tag{18}$$

where $\boldsymbol{R}$ is a matrix of individual gradients of the loglikelihood function, with typical row

$$\left(\hat{\varepsilon}^{G}_{it}\boldsymbol{x}'_{it}, \hat{\varepsilon}^{G}_{it} z'_{it}\right),$$

and $\boldsymbol{\iota}$ is a vector of ones. It can be shown that the Lagrange multiplier test statistic ξ_{LM} is asymptotically χ^2 distributed with r degrees of freedom, under the null hypothesis that z_{it} does not enter the model.

A.2. Testing for heteroskedasticity

Suppose that ε_{it} has a variance of

$$V\{\varepsilon_{it}\} = h_{it} = h(z_{it}'\gamma)$$

for some function $h > 0$ with $h(0) = 1$ (normalization condition), where z_{it} is of dimension r. Using this specification for the variance the loglikelihood changes to the following form

$$\begin{aligned}\log L(\beta,\gamma) \\ &= \sum_{i,t} y_{it} \log \Phi\left(\frac{\boldsymbol{x}_{it}'\beta}{\sqrt{h(z_{it}'\gamma)}}\right) + \sum_{i,t}(1-y_{it})\log\left(1-\Phi\left(\frac{\boldsymbol{x}_{it}'\beta}{\sqrt{h(z_{it}'\gamma)}}\right)\right).\end{aligned} \tag{19}$$

Now the first order conditions for γ, evaluated under the null hypothesis H_0: $\gamma = 0$ are

$$\sum_{i,t} \hat{\varepsilon}_{it}^G (\boldsymbol{x}_{it}'\beta) z_{it} = 0.$$

Consequently, it is easy to test $H_0 : \gamma = 0$ using the Lagrange Multiplier test statistic given in Eq. (18) using a matrix $\boldsymbol{R}$ that has typical row

$$\left(\hat{\varepsilon}_{it}^G \boldsymbol{x}_{it}', \hat{\varepsilon}_{it}^G \left(\boldsymbol{x}_{it}\hat{\beta}\right) z_{it}'\right).$$

A.3. Testing for non-normality

A test for normality can be derived by specifying an the alternative distribution function as $\Phi(x'\beta + \gamma_2(x'\beta)^2 + \gamma^3(x'\beta)^3)$ (compare Newey, 1985). The null hypothesis of normality corresponds to $\gamma_2 = \gamma_3 = 0$. This can be tested by using Eq. (18), where the matrix $\boldsymbol{R}$ now contains

$$\left(\hat{\varepsilon}_{it}^G \boldsymbol{x}_{it}', \hat{\varepsilon}_{it}^G \left(\boldsymbol{x}_{it}'\hat{\beta}\right)^2, \hat{\varepsilon}_{it}^G \left(\boldsymbol{x}_{it}'\hat{\beta}\right)^3\right).$$

References

Amemiya, T., 1985. Advanced Econometrics. Harvard Univ. Press, Cambridge.

Baltagi, B.H., 1995. Econometric Analysis of Panel Data. Wiley, Chichester.

Blake, C.R., Timmermann, A., 1998. Mutual fund performance: evidence from the UK. European Finance Review 2, 57–77.

Brown, S.J., Goetzmann, W.N., 1995. Performance persistence. The Journal of Finance 50, 679–698.

Brown, S.J., Goetzmann, W.N., 1997. Mutual fund styles. Journal of Financial Economics 43, 373–399.

Brown, S.J., Goetzmann, W.N., Ibbotson, R.G., Ross, S.A., 1992. Survivorship bias in performance studies. Review of Financial Studies 5, 553–580.

Brown, S.J., Goetzmann, W.N., Ross, S.A., 1995. Survival. The Journal of Finance 50, 853–874.

Carhart, M.M., 1997a. On persistence in mutual fund performance. The Journal of Finance 52, 57–82.

Carhart, M.M., 1997. Mutual Fund Survivorship, Working Paper, Marshall School of Business, U.S.C.

Carpenter, J.F., Lynch, A., 1999. Survivorship bias and attrition effects in measures of performance persistence. Journal of Financial Economics 54, 337–374.

Chan, L.C., Jegadeesh, N., Lakonishok, J., 1996. Momentum strategies. The Journal of Finance 53, 267–284.

Chevalier, J., Ellison, G., 1997. Risk taking by mutual funds as a response to incentives. Journal of Political Economy 105, 1167–1200.

Elton, E.J., Gruber, M.J., Blake, C.R., 1995. Survivorship bias and mutual fund performance. The Review of Financial Studies 9, 1097–1120.

Ferson, W.E., Schadt, R.W., 1996. Measuring fund strategy and performance in changing economic conditions. The Journal of Finance 51, 425–461.

Fitzgerald, J., Gottschalk, P., Moffitt, R., 1998. An analysis of sample attrition in panel data: the michigan panel study of income dynamics. Journal of Human Resources 33, 251–299.

Gouriéroux, C., Monfort, A., Renault, E., Trognon, A., 1987. Generalized residuals. Journal of Econometrics 34, 5–32.

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, 783–811.

Hausman, J.A., Wise, D., 1979. Attrition bias in experimental and panel data: the gary income maintenance experiment. Econometrica 47, 455–473.

Heckman, J., 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. The Annals of Economic and Social Measurement 5, 475–492.

Heckman, J., 1979. Sample selection bias as a specification error. Econometrica 47, 153–161.

Hendricks, D., Patel, J., Zeckhauser, R., 1997. The J-shape of performance Persistence given Survivorship Bias. The Review of Economics and Statistics 79, 161–170.

Malkiel, B.G., 1995. Returns from investing in equity mutual funds 1971 to 1991. The Journal of Finance 50, 549–572.

Newey, W.K., 1985. Maximum likelihood specification testing and conditional moment tests. Econometrica 53, 1047–1070.

Pagan, A., Vella, F., 1989. Diagnostic tests for models based on individual data: a survey. Journal of Applied Econometrics 4, S29–S59.

Rockinger, M., 1995. Determinants of Capital Flows to Mutual Funds, Working Paper, HEC, France.

Sirri, E.R., Tufano, P., 1998. Costly search and mutual fund flows. Journal of Finance 53, 1589–1622.

Verbeek, M., 2000. A Guide to Modern Econometrics. Wiley, Chichester.
Wermers, R., 1997. Momentum Investment Strategies of Mutual Funds, Performance Persistence, and Survivorship Bias, Working Paper.
Zheng, L., 1999. Is money smart? A study of mutual fund investors' fund selection ability. The Journal of Finance 54, 901–933.