Review of Finance (2007) 11: 605–632
doi: 10.1093/rof/rfm012
Advance Access publication: 8 May 2007

Fund Liquidation, Self-selection, and Look-ahead Bias in the Hedge Fund Industry*

JENKE TER HORST[1] and MARNO VERBEEK[2]
[1] *Tilburg University;* [2] *RSM Erasmus University and Netspar*

Abstract. A wide range of empirical biases hampers hedge fund databases. In this paper we focus upon survival-related biases and disentangle look-ahead biases due to self-selection of funds and due to fund termination. Self-selection arises because funds voluntarily report their information to data vendors and may decide to stop doing so. By extending existing methodology, we analyze persistence in hedge fund performance over the period 1994–2000, taking into account the above biases. The results show that look-ahead biases due to liquidation and self-selection enforce each other and may lead to overestimating expected returns by as much as 8% per year. Overall, the results are consistent with positive persistence in hedge fund returns at horizons of two and four quarters.

JEL Classification: G11, G23, G14

1. Introduction

During the last decade, the hedge fund industry has grown enormously. Hedge funds differ from mutual funds and other investment vehicles by their lack of regulation[1], with limited transparency and disclosure, and by their internal structure (see, e.g., Fung and Hsieh, 1997). For example, most hedge funds try to achieve an absolute return target, irrespective of global market movements, while hedge fund managers typically have incentive-based contracts. Accordingly, hedge funds have a broad flexibility in the type of securities they hold and the type of positions they take. On the other

* The authors gratefully acknowledge financial support by the Institute for Quantitative Investment Research Europe (INQUIRE). We thank the editor (Josef Zechner), an anonymous associate editor and two anonymous referees for helpful comments and suggestions.
[1] U.S. hedge funds are defined by their freedom from regularity controls of the Investment Company Act of 1940.

© The Author 2007. Published by Oxford University Press on behalf of the European Finance Association. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org

hand, investors in hedge funds are often confronted with lockup periods and redemption notice periods. Such restrictions on withdrawals imply smaller cash fluctuations, and give fund managers more freedom in setting up long-term or illiquid positions.

Their nonstandard features make hedge funds an interesting investment vehicle for investors with potential diversification benefits. A wide range of academic papers examines hedge fund performance and its persistence.[2] Both for the mutual fund industry (Sirri and Tufano, 1998) and the hedge fund industry (Agarwal et al., 2004; Baquero and Verbeek, 2006), it is reported that money flows chase past performance. Berk and Green (2004) present a theoretical model that explains that persistence can be competed away by investors rationally shifting their money in search for superior investments. In the hedge fund industry, however, the presence of liquidity restrictions, which prevent investors to quickly shift their money from one fund to the other, may result in genuine (short-run) persistence even if investors allocate their money according to past performance (see Baquero et al., 2005).

In this paper we analyze the persistence in hedge fund performance taking into account a number of potentially important biases that are present in hedge funds databases (see Fung and Hsieh, 1997; Ackermann et al., 1999) or are induced by the employed methodology. Survivorship bias (see, e.g., Brown et al., 1992; Brown et al., 1999) arises if information on defunct funds is unavailable and only the performance of surviving funds is investigated. For hedge funds, this bias is more severe than in the mutual fund industry because of the much higher attrition rate (about 14% per year for hedge funds versus about 5% per year for mutual funds); see Malkiel and Saha (2005). Survivorship bias can be avoided by using a "survivorship-bias-free database" (see, e.g., Elton et al., 1996), which also includes (historical) information on funds that are no longer active. However, when investigating performance or its persistence over multiple periods, this imposes the condition that fund returns are available over a number of consecutive periods. For example, when investigating whether fund performance over two historical years persists in the subsequent 2 years, only funds that survived and reported over a 4 year period are used in the empirical analysis. This typically introduces a multiperiod sampling bias or look-ahead bias (see, e.g., ter Horst et al., 2001), even if the database is "survivorship-bias-free". This paper focuses upon the importance of look-ahead bias and its underlying causes.

Since hedge funds are not allowed to advertise publicly, hedge fund databases (like Trading Advisors Selection System (TASS), Managed Account Reports

[2] See, e.g., Brown et al., 1999; Agarwal and Naik, 2000; Bares et al., 2003; Boyson and Cooper, 2004; Capocci and Hübner, 2004; Baquero et al., 2005; Jagannathan et al., 2006.

(MAR), and Hedge Fund Research (HFR)) serve as important distribution channels. However, reporting to these databases is voluntary. Accordingly, hedge funds may decide to start reporting after some initial successful period, and enter a database with an instant history. This process may lead to an upward bias in performance measures, referred to as incubation bias and backfilling bias (see, e.g., Posthuma and van der Sluis, 2003). Further, funds may stop reporting for several reasons. Besides liquidation of the fund, fund managers may stop reporting because underperformers do not wish to make their performance known because funds that performed well have less incentive to report to data vendors to attract potential investors, or because funds may decide to change their reporting from one database to another. There is no consensus in the literature on the terminology for these biases (see, e.g., Jagannathan et al., 2006). We will refer to the bias due to an endogenous decision of fund managers to stop reporting as self-selection bias (or, more precisely, look-ahead bias due to self-selection). When nonreporting is related to fund liquidation, we refer to it as liquidation bias.

A recent study of persistence in performance of hedge funds by Baquero et al. (2005) finds that look-ahead bias due to liquidation seriously affects the results and that correcting for liquidation bias is essential. For instance, without correcting for endogenous fund liquidation, average raw returns (within a given ranking decile) might be overestimated by as much as 5% when persistence is analyzed at an annual level. However, in Baquero et al. (2005) it is assumed that self-selection is exogenous. If self-selection would be mainly driven by good performing funds that are closed to new investment, this may have a compensating impact upon performance and persistence measures, such that liquidation bias and self-selection bias offset each other. On the other hand, if self-selection is negatively related to past performance, correcting for self-selection bias may exacerbate the liquidation bias corrections and thus strengthen the reported persistence patterns in hedge fund performance. Consequently, it is an interesting question to separately identify the impact of liquidation bias and self-selection bias in hedge fund persistence.

In this paper we analyze the persistence in hedge fund performance taking into account both liquidation and self-selection bias. The question whether past performance is indicative of future performance has been extensively studied for mutual funds. The results are somewhat mixed, but in general it can be concluded that there is little evidence of performance persistence of mutual funds.[3] For hedge funds, recent studies show some evidence of short-term performance persistence (see, e.g., Agarwal and Naik, 2000; Bares

[3] See, e.g., Carhart, 1997; ter Horst and Verbeek, 2000; Wermers, 2003; Bollen and Busse, 2005; Huij and Verbeek, 2007.

et al., 2003), while at longer horizons the results are more ambiguous (see, e.g., Brown et al., 1999; Brown and Goetzmann, 2003; Kosowski et al., 2007). None of these studies corrects for the possibility of look-ahead bias. Baquero et al. (2005) correct for look-ahead bias due to fund liquidation and find exacerbated persistence patterns in hedge fund performance, particularly at the annual horizon.

In this paper we make a number of contributions to the hedge funds literature. First, we empirically examine the factors that affect self-selection bias by identifying variables from reports supplied by data vendors. Interestingly, past performance appears to have a significant and negative impact upon the probability that a fund decides to stop reporting. That is, poorly performing funds are more likely to disappear from the TASS database at their own request. Second, we propose a method that will correct for self-selection bias separately from the look-ahead bias due to fund liquidation. Finally, while disentangling the effects of liquidation bias and self-selection bias, we analyze the persistence in hedge fund performance over various horizons, using the TASS database of hedge funds over the period 1994–2000. The results indicate that, in addition to liquidation bias, correcting for self-selection bias is important. Both biases work in the same direction and their combined impact may result in overestimating expected returns within a given decile by as much as 7.7% per year. As a result, the finding of persistence in hedge fund performance is strengthened once one corrects for both liquidation and self-selection biases. At the annual horizon, the expected excess return on a winner minus loser portfolio, based upon previous year returns, is approximately 10% when both biases are corrected for, while it is only 4.3% without corrections.

The structure of this paper is as follows. In Section 2 we discuss the TASS database, analyze fund attrition, and relate it to liquidation and self-selection. Moreover, we estimate probit specifications for both the liquidation and the self-selection decisions. Section 3 explains how one can correct for look-ahead bias due to liquidation and self-selection when analyzing persistence in hedge fund performance and how these two biases can be disentangled. Empirical results concerning persistence at different horizons are presented in Section 4, while Section 5 contains some robustness checks. Finally, Section 6 concludes.

2. Liquidation and Self-selection

The data used in this paper are from TASS Management Limited and contain information of 1797 hedge funds over the period 1994–2000, where we restrict attention to funds reporting in US$. Although the TASS database contains information of hedge funds since 1979, we focus on the period 1994–2000 for several reasons. First, information on "dead" funds is available only for funds that disappeared since 1994, and second, the number of funds before 1994 is very small. As mentioned above, whether or not we observe returns for a given fund depends upon two main issues. First, the fund may be liquidated. Second, if the fund is not liquidated, its management may prefer to not report returns and other information to TASS. We refer to this second decision as self-selection. Another potential problem is backfilling bias or instant history bias (see, e.g., Posthuma and van der Sluis, 2003), which arises because when funds are included in the TASS database for the first time, they come with a history of several quarters. We take backfilling bias into account by only using information on a fund once its age exceeds 1 year.

TASS provides some qualitative information on the reason why a fund disappeared from their sample. For example, we find explanations like "This fund liquidated end of August 1996." "Due to the Fund Managers request this fund has been taken off the database.", or "This fund is closed for new investments." In order to disentangle the different selection processes, we first construct a variable that contains a code for the disappearance reason. We used the following codes and scanned the reported disappearance records for the following expressions:

1. liquidated or closed down,
2. at fund manager's request,
3. closed to new investors,
4. closed (unknown),
5. matured.

From the 612 hedge funds that have disappeared from the TASS database during 1994–2000, about half falls in category 1 and we classify this group as "liquidated". Cases (42 funds) with reasons 3 and 5 are classified as "self-selection". For the other reasons, classification as either liquidated or self-selected is less obvious. Some of the funds with disappearance reason 2 use a negative formulation to motivate their removal from the database. For instance "This fund has been removed at fund manager's request and will not be providing performance tables as the fund is too small.", or "This fund was requested to be taken off by the fund manager because they felt that TASS clients were not taking an interest in their fund!". Accordingly, some of the funds with disappearance reason 2 may stop reporting the final periods leading up to their liquidation (see Ackermann et al., 1999) and are more appropriately classified as "liquidated" or "terminated" rather than "self-selected".[4] In order to classify these funds, and also to classify funds that

[4] The resulting bias is sometimes referred to as "end-of-life bias", see Malkiel and Saha (2005).

disappeared due to an unknown reason (reason 4), we evaluate the money flows of these funds over the year preceding their attrition. We estimate money flows following the procedure used by Agarwal et al. (2004). Assuming that flows take place at the end of quarter $t + 1$, flows are measured as the growth in total assets under management of a fund between the start and end of quarter $t + 1$, in excess of the investment return during the quarter. Subsequently, we aggregate these money flows over the four last available quarters.

In general, it is difficult to pin down what exactly self-selection means for these hedge funds, and a variety of reasons may underlie a fund's decision to stop reporting. We prefer to be transparent and accountable in our classification and classify all funds that disappear (according to TASS) "at the fund manager's request" (reason 2) or because of an unknown reason (reason 4), by the sign of their cash flows during the last year. Funds in this category experiencing negative cash flows are considered "liquidated", while those with positive cash flows are considered "self-selected".[5] Note that what we identify empirically as "self-selection" refers only to the decision of a hedge fund to stop reporting to the TASS database, conditional upon the fact that the fund started reporting at some point. It does not reflect the initial decision of a fund to report to TASS or another database vendor (e.g. HFR), or to not report at all.

From the 612 cases disappearing between 1994 and 2000, we classify 464 (76%) as being liquidated, and the remaining 148 as "self-selected" out of the TASS database. The average quarterly return, in the last year of reporting, is −0.10% and −0.51%, respectively, while the average net flows are −5.21% for the liquidated funds and +24.10% for the self-selected funds. Somewhat surprisingly, funds that stop reporting voluntarily experience positive average cash flows in their final year but relatively low average returns. The estimate of −0.51%, however, is not very precise. More detailed information is provided in Table I, where we report the average quarterly returns and the average quarterly net money flows for funds that self-select, liquidate, or survive during the sample period 1994–2000. For example, the first row indicates that, in the first quarter of 1994, the average return of funds that liquidated before the end of our sample period is −2.33%, while it is −1.83% for funds that survived until 2000. The table clearly shows that funds that liquidate during the sample period have substantial lower average returns and net money flows than funds that self-select or survive. The average return for funds that liquidate is about 0.95% per quarter, while funds that self-select or survive have an average quarterly return of 1.62% and 3.59%, respectively. The average quarterly net money flows exhibit a similar pattern. Funds that liquidate have

[5] We explore and discuss alternative classifications in Section 5 below.

Table I. Average quarterly returns and net money flows

The table shows the average quarterly returns and net money flows of U.S. hedge funds in the TASS database that self-select, liquidate, or survive during the sample period 1994–2000. The row labeled "average (weighted)" reports the weighted averages (weighted by the number of funds per quarter) over the sample period, while the row "average (unweighted)" reports the unweighted averages.

	Self-selected		Liquidated		Survivors	
Quarter	Return	Net flow	Return	Net flow	Return	Net flow
94-I	−1.54%	1.22%	−2.33%	9.73%	−1.83%	6.33%
94-II	0.26%	2.23%	2.26%	8.27%	0.97%	8.71%
94-III	1.19%	10.94%	0.07%	3.98%	2.27%	8.54%
94-IV	−1.15%	−1.22%	−1.32%	1.35%	−0.73%	2.40%
95-I	2.35%	4.10%	2.91%	5.46%	4.47%	12.92%
95-II	2.44%	19.39%	1.84%	1.31%	5.11%	6.83%
95-III	3.51%	2.37%	1.19%	0.73%	4.71%	7.18%
95-IV	3.12%	−1.30%	2.57%	3.79%	3.64%	1.60%
96-I	4.00%	8.05%	1.37%	1.16%	3.32%	8.39%
96-II	4.79%	4.66%	5.64%	13.07%	6.06%	11.40%
96-III	1.84%	3.59%	0.92%	−1.18%	2.25%	14.04%
96-IV	5.01%	4.12%	3.76%	8.76%	6.68%	14.63%
97-I	3.43%	33.03%	4.60%	11.87%	4.32%	13.79%
97-II	3.09%	6.97%	3.57%	2.99%	5.35%	12.14%
97-III	4.14%	16.56%	6.14%	11.30%	7.83%	15.64%
97-IV	−2.56%	−0.94%	−2.38%	−5.00%	−0.50%	5.75%
98-I	2.94%	18.32%	1.12%	3.03%	5.73%	19.48%
98-II	−5.73%	47.18%	−3.21%	−2.09%	−0.66%	9.76%
98-III	−3.14%	11.63%	−5.64%	0.82%	−4.98%	2.63%
98-IV	−2.93%	23.97%	0.65%	0.29%	5.85%	9.14%
99-I	−0.10%	87.28%	−1.25%	−9.88%	3.70%	2.96%
99-II	2.06%	0.16%	2.30%	−4.15%	8.45%	9.48%
99-III	−0.99%	11.38%	−1.75%	−10.11%	0.69%	−3.29%
99-IV	12.85%	22.14%	−0.25%	−3.67%	13.36%	17.20%
00-I	–	–	–	–	5.96%	0.81%
Average (unweighted)	1.62%	13.99%	0.95%	2.16%	3.59%	9.07%
Average (weighted)	1.58%	12.01%	1.30%	2.85%	3.98%	9.14%

an average quarterly net money flow of only 2.16%, while funds that self-select or survive have an average net money flow of 13.99% and 9.07%, respectively. Combining the three subsamples gives an average quarterly return and net money flow of 3.06% and 7.78%, respectively (not reported). Note that the average returns of the surviving funds are about 2.1% (per annum) higher than the average return of the combined samples. This number is usually referred to as survivorship bias (see, e.g. Malkiel, 1995; Liang, 2000). The liquidation bias is defined as the difference between the average returns of the combination of the subsamples of self-selected and surviving funds and the combination of all three subsamples. This bias is about 2.0% (per annum) (see, e.g. Baquero et al., 2005). Finally, we can define self-selection bias as the difference between the subsamples of liquidated and surviving funds and the combination of all three subsamples. This bias is about 0.8% (per annum).

Because our interest lies in persistence at horizons of at least one quarter, we aggregate all information to quarterly levels. This has the advantage of reducing the impact of return smoothing due to the possibility that a hedge fund invests in securities that are not actively traded (see Getmansky et al., 2004). Consequently, we also analyze liquidation and self-selection at the quarterly level. In the remainder of the paper liquidation will be denoted by an indicator variable L, such that $L_{it} = 0$ if fund i has liquidated in quarter t ($L_{it} = 1$ otherwise). Given that a fund is not liquidated, returns may not be available due to self-selection, and we let $S_{it} = 0$ if fund i attrited the database because of self-selection ($S_{it} = 1$ otherwise). This implies that a return r_{it} is observed only if $L_{it} S_{it} = 1$.

For both decisions we specify a binary choice model. First, the liquidation decision is modeled by means of a binary probit model, with latent variable equation

$$L_{it}^{*} = \alpha_1 + \sum_{j=1}^{6} \gamma_{1j} r_{i,t-j} + x_{it}' \beta_1 + \varepsilon_{1,it}, \tag{1}$$

where x_{it} denotes a vector of fund characteristics that affect liquidation. The observed indicator satisfies $L_{it} = 0$ (liquidation) if $L_{it}^{*} < 0$ and 1 otherwise. The specification allows fund returns up to six quarters ago to affect the liquidation decision. It is assumed that $\varepsilon_{1,it}$ is $IIN(0, 1)$, independent of the explanatory variables. We expect that $\gamma_{1j} > 0$ for several of the lags, so that the better performing funds are, *ceteris paribus*, less likely to liquidate.

Similarly, we specify a process for the self-selection decision as a probit model based on

$$S_{it}^{*} = \alpha_2 + \sum_{j=1}^{6} \gamma_{2j} r_{i,t-j} + x_{it}' \beta_2 + \varepsilon_{2,it}, \tag{2}$$

with $S_{it} = 0$ (self-selection) if $S_{it}^{*} < 0$ and 1 otherwise. While the set of conditioning variables x_{it} in both equations is in principle the same, a priori exclusion restrictions may be imposed.

In Table II we present some summary statistics of the fund-specific variables (x_{it}) that were included in the liquidation and self-selection models. Most of these variable also appear in related specifications of Brown et al. (2001) and

Table II. Summary statistics

The table reports the summary statistics of fund-specific variables. (20,138 fund/period observations)

Variable	Mean	Std.dev	Min	Max
offshore	0.59	0.49	0	1
Incentive fees	15.87	7.82	0	50
Mng. fees	1.63	1.08	0	8
Underwater	0.14	0.34	0	1
ln(NAV)	16.73	1.79	7.58	23.30
ln(Age)	3.80	0.66	2.56	5.62
$\ln(\text{Age})^2$	14.89	5.09	6.58	31.55
StDev	0.08	0.08	0.00	1.63

Baquero et al. (2005). Summary statistics are based on 20,138 fund/period observations. It appears that 59% of the observations are from offshore hedge funds. These funds, while reporting in US$, are located in tax havens like the Virgin Islands. The average incentive fee of the fund manager is about 16%, but can be as high as 50% of realized performance. Note that these incentive fees are obtained only when the fund has recovered past losses (high watermark). The annual management fee varies between 0% and 8% (of net asset value) and has an average of 1.6%. The underwater indicator is equal to 1 if a fund has a negative cumulative return over the past eight quarters[6], which occurs in 14% of the cases. The age of the funds varies between 13 months and 275 months (about 23 years), while the average age is about 45 months. The average size of the hedge funds, measured by their log net asset value corresponds to about 18 million US$. Total risk is measured by the standard deviation of the previous six quarterly returns (StDev).

Fund size (NAV) is not available for each quarter for all funds in our sample. Therefore, we use the most recent observation of net asset value available from the TASS database. However, there remain some observations for which NAV is missing and cannot be imputed. This occurs in 7% of the cases. Because we do not want to eliminate these observations, we estimated the liquidation and self-selection models using two specifications, one including size (based on 20,138 fund/period observations) and one excluding size (based on 21,297 fund/period observations). Following Baquero et al. (2005) we employ the TASS investment style classification, which is closely related to

[6] The cumulative return is determined over at least five quarters with a maximum of eight quarters.

Table III. Estimation results liquidation model

The table reports the estimation results of the liquidation model, including net asset value (size); 20,138 fund/period observations.

Parameters	Estimate	Std.error	Parameters	Estimate	Std. Error
Intercept	1.984	0.763	Offshore	−0.136	0.051
$r(-1)$	1.182	0.214	Incentive Fees	−0.008	0.003
$r(-2)$	0.802	0.222	Mng. Fees	−0.012	0.024
$r(-3)$	1.058	0.223	Underwater	−0.245	0.066
$r(-4)$	0.220	0.227	Emerging markets	−0.154	0.080
$r(-5)$	0.029	0.207	Equity market neutral	−0.191	0.092
$r(-6)$	0.103	0.221	Event driven	0.083	0.107
ln(NAV)	0.155	0.015	Fixed income arbitrage	−0.180	0.202
StDev	1.182	0.370	Global macro	−0.164	0.176
ln(Age)	−0.617	0.379	Long/short equity	−0.160	0.075
$\ln(\text{Age})^2$	0.087	0.050	Managed futures	−0.075	0.071
Loglikelihood: −17,29.338 pseudo R^2 : 0.148			Chi-squared test: 600.28($DF = 42$) ($p = 0.000$)		

the nine commonly used Tremont hedge fund style indices.[7] In Table III we report the estimation results based on 20,138 fund/period observations for the probit specification for liquidation versus nonliquidation (including size). Note that nonliquidation means that it is still possible that the fund self-selected during the sample period. Therefore, we subsequently remove all the fund/period observations where a fund liquidated (417 fund/period observations) and estimate the probit specification to explain self-selection versus survival (including size), where survival implies that the fund did not liquidate and still prefers to report their performance to the data vendor. The estimation results are reported in Table IV. The estimation results for both specifications excluding size are reported in the Appendix (Table IX and Table X). All models include investment style dummies, while time dummies are included to capture aggregate shocks to the probabilities of liquidation and self-selection. The investment styles "convertible arbitrage" and "dedicated short bias" contain very limited numbers of funds. Accordingly, no dummies are included for these styles, and these funds are implicitly allocated to the general hedge fund index (reference category). The coefficient estimates for the time dummies are available upon request.

In specification Equation (1), the first three past quarterly returns have a significant impact on the liquidation decision, while in specification Equation

[7] Tremont has the following nine hedge fund investment styles: Convertible Arbitrage, Dedicated Short Bias, Emerging Markets, Equity Market Neutral, Event Driven, Fixed Income Arbitrage, Global Macro, Long/Short Equity, Managed Futures, and Hedge Fund Index.

Table IV. Estimation results self-selection model

The table reports the estimation results of the self-selection model, including net asset value (size); 19721 fund/period observations.

Parameters	Estimate	Std.error	Parameters	Estimate	Std. Error
Intercept	2.915	1.163	Offshore	0.093	0.076
$r(-1)$	1.393	0.357	Incentive Fees	−0.013	0.005
$r(-2)$	1.289	0.366	Mng. Fees	−0.035	0.036
$r(-3)$	0.445	0.340	Underwater	0.083	0.116
$r(-4)$	0.732	0.374	Emerging markets	−0.068	0.133
$r(-5)$	−0.130	0.326	Equity market neutral	0.023	0.148
$r(-6)$	0.657	0.389	Event driven	0.118	0.150
ln(NAV)	0.093	0.023	Fixed income arbitrage	0.019	0.366
StDev	1.414	0.622	Global macro	0.202	0.357
ln(Age)	−0.719	0.591	Long/short equity	−0.148	0.108
$\ln(\text{Age})^2$	0.102	0.078	Managed futures	0.058	0.111
Loglikelihood: −658.876 pseudo R^2 : 0.093			Chi-squared test: 135.78 ($DF = 42$) ($p = 0.000$)		

(2) only the first two past quarterly returns are statistically significant. The cumulative impact of the first three lagged returns in both models is about the same. In the liquidation model, positive coefficients indicate that higher historical returns imply a lower probability to liquidate. In the self-selection model, they indicate that funds with high historical returns are more likely to survive, where surviving means that the fund did not liquidate and decides to keep reporting to TASS. However, in the liquidation model an additional impact of historical returns is captured by the underwater indicator, which is highly significant. The negative coefficient implies that if a fund has a negative aggregate return over the most recent eight quarters, it is significantly more likely to liquidate. That is, if a fund is underwater, implying that the manager will not receive the incentive fee, the probability of liquidation increases substantially, potentially due to excessive risk-taking (compare Goetzmann et al., 2003). For self-selection, the impact of being underwater seems negligible, both economically and statistically. Overall, the results in Tables III and IV indicate that the impact of historical returns is somewhat stronger for the liquidation decision than for the self-selection decision. The impact of size (ln(NAV)) is significantly positive, i.e. smaller funds have a higher probability to liquidate or self-select than larger funds (ceteris paribus).

The results for the self-selection model are clearly at odds with the idea that good performing funds are more likely to stop reporting because they no longer wish to attract new investors. Also, it does not appear to be the case that large funds, which may suffer most from decreasing returns to scale of their investment strategies and are more likely to voluntarily stop reporting because managers are unwilling to attract new money. Total risk, as measured by the standard deviation over the past six quarters, significantly affects the liquidation and self-selection decision positively, implying that high-risk hedge funds have a higher probability to survive. This seems a counterintuitive result, and apparently contradicts the findings of Brown et al. (2001) who find that high-risk funds have a higher probability to liquidate. However, our results suggest that high-risk funds experience a somewhat lower liquidation and self-selection probability, given the return history and fund size. Accordingly, they indicate that high-risk funds are allowed to have more extreme negative returns than low-risk funds before they decide to liquidate or self-select (see Baquero et al., 2005).

In both specifications, age has a significant nonlinear impact, indicating that old funds with past poor performance are less likely to disappear than young funds with a similar poor performance. This finding corresponds to the results of Boyson and Cooper (2004), who perform unconditional and conditional survival tests, and finds that age and manager ability are positively related to the likelihood of a manager's survival. Offshore funds have a larger probability to liquidate than onshore funds, while given that the fund did not liquidate, being an offshore fund does not significantly affect the self-selection decision. The impact of the incentive fee on the nonliquidation probability or survival probability is significantly negative, i.e. a higher incentive fee, *ceteris paribus*, increases the probability that a fund will liquidate or self-select. The significant coefficient estimates for the investment style dummies "equity market neutral" and "long/short equity" in the liquidation model indicate, *ceteris paribus*, that these styles have a higher probability to liquidate. The impact of investment styles on the self-selection decision is insignificant, although most of the coefficient estimates are positive in contrast to the estimates for the liquidation model.

Our results show that most of the factors in our specifications affect the liquidation and self-selection decisions in the same direction. Given the importance of past performance, we conclude that self-selection is not exogenous. Consequently, it can be expected that look-ahead biases due to self-selection and liquidation will not offset, but even strengthen each other, and correcting for both biases will be necessary. In the next section, we describe how both biases can be disentangled and how persistence analyses can be corrected for these biases. This can be achieved using an extension of the methodology reported in ter Horst et al. (2001).

3. Disentangling Look-ahead Bias and Self-selection Bias

Suppose interest lies in analyzing fund performance over the period $t+1$ to $t+s+1$, conditional upon a given information set Ω_t. In some applications, this information set may be empty. In others, Ω_t will contain indicators for the fund's investment style and its previous performance (e.g., its performance decile during a ranking period). This means that interest lies in the conditional distribution of returns $r_{i,t+1}, \ldots, r_{i,t+s+1}$ given Ω_t, which we denote by

$$f(r_{i,t+1}, \ldots, r_{i,t+s+1} | \Omega_t), \tag{3}$$

where f is generic notation for a (conditional) density function. Empirically, we can only obtain full information about this joint distribution if the fund has not liquidated or self-selected during the period $t+1$ to $t+s+1$. Let us denote this by $Y_{it} = 1$. This means we can empirically identify

$$f(r_{i,t+1}, \ldots, r_{i,t+s+1} | \Omega_t, Y_{it} = 1). \tag{4}$$

If Equations (3) and (4) are identical, liquidation and self-selection is exogenous and no biases arise if the sample selection process is ignored. However, as we have seen in the previous sections, it is likely that both liquidation and self-selection are determined by historical performance and other characteristics that may have a relation with returns during the period $t+1$ to $t+s+1$. For example, funds that have high levels of (idiosyncratic) risk are more likely to have extreme returns and are typically less likely to survive (see Brown et al., 1992; Hendricks et al., 1997). The difference between Equations (3) and (4) drives the look-ahead bias in performance measures.

The distribution of interest in Equation (3) can be derived from the joint distribution of $r_{i,t+1}, \ldots, r_{i,t+s+1}$ and z_{it}, conditional upon Ω_t and $Y_{it} = 1$, where z_{it} denotes a vector of observable fund characteristics and other variables that are relevant for fund liquidation and self-selection from $t+1$ to $t+s+1$. The role of z_{it} will become clear below. First, note that

$$\begin{aligned} f(r_{i,t+1}, \ldots, .r_{i,t+s+1}, z_{it} | \Omega_t, Y_{it} = 1) \\ &= \frac{f(r_{i,t+1}, \ldots, r_{i,t+s+1}, z_{it}, Y_{it} = 1 | \Omega_t)}{P\{Y_{it} = 1 | \Omega_t\}} \\ &= \frac{P\{Y_{it} = 1 | r_{i,t+1}, \ldots, r_{i,t+s+1}, z_{it}, \Omega_t\} f(r_{i,t+1}, \ldots, r_{i,t+s+1}, z_{it} | \Omega_t)}{P\{Y_{it} = 1 | \Omega_t\}} \\ &= \frac{f(r_{i,t+1}, \ldots, r_{i,t+s+1}, z_{it} | \Omega_t)}{w_{it}}, \end{aligned} \tag{5}$$

where

$$w_{it} = \frac{P\{Y_{it} = 1|\Omega_t\}}{P\{Y_{it} = 1|r_{i,t+1}, \ldots, r_{i,t+s+1}, \Omega_t, z_{it}\}} \tag{6}$$

is a weight factor. Accordingly, it follows that

$$f(r_{i,t+1}, \ldots, r_{i,t+s+1}|\Omega_t) = \int_z w_{it} f(r_{i,t+1}, \ldots, r_{i,t+s+1}, z_{it}|\Omega_t, Y_{it} = 1) dz_{it}, \tag{7}$$

where the integral is over the support of z_{it}. In this expression, the weight factor w_{it} indicates how the distribution, conditional upon $Y_{it} = 1$, can be adjusted to recover the distribution of returns unconditional upon $Y_{it} = 1$, which is what we are really interested in. If we are willing to assume that the denominator of Equation (6) does not depend upon $r_{i,t+1}, \ldots, r_{i,t+s+1}$ directly, but only through z_{it} (which may contain historical returns and other fund characteristics), the weights can be identified. In this case, the weights reduce to

$$w_{it} = \frac{P\{Y_{it} = 1|\Omega_t\}}{P\{Y_{it} = 1|\Omega_t, z_{it}\}}. \tag{8}$$

Note that in general these weights are endogenous, as z_{it} will not be independent of $r_{i,t+1}, \ldots, r_{i,t+s+1}$. This approach to solve selection bias problems assumes that a set of (observable) explanatory variables z_{it} can be chosen such that, conditional upon z_{it}, selection is independent of current and future, potentially unobserved, returns. This is referred to as "selection upon observables" and is employed in, e.g., Fitzgerald et al. (1998) to correct for attrition bias from the Panel Study of Income Dynamics, and in ter Horst et al. (2001) to eliminate look-ahead bias in evaluating persistence in mutual fund performance.

In the current application, $Y_{it} = 1$ implies that the fund has not liquidated during $t + 1$ to $t + s + 1$ and has not stopped reporting because of self-selection. Let us refer to these two conditions as $Y_{1,it} = 1$ and $Y_{2,it} = 1$, respectively, so that $Y_{it} = Y_{1,it}Y_{2,it}$. Referring to the two binary choice models specified above, it holds that

$$Y_{1,it} = \prod_{\tau=t+1}^{t+s+1} L_{i\tau}$$

and

$$Y_{2,it} = \prod_{\tau=t+1}^{t+s+1} S_{i\tau}.$$

Then $Y_{1,it} = 0$ says that fund i is not used in the analysis in period t because of fund liquidation, while $Y_{2,it} = 0$ says that it is not used because of self-selection. To disentangle the impact of these two processes, we first rewrite Equation (8) as

$$\begin{aligned} w_{it} &= \frac{P\{Y_{it} = 1|\Omega_t\}}{P\{Y_{it} = 1|\Omega_t, z_{it}\}} \\ &= \frac{P\{Y_{2,it} = 1|\Omega_t, Y_{1,it} = 1\}}{P\{Y_{2,it} = 1|\Omega_t, z_{it}, Y_{1,it} = 1\}} \times \frac{P\{Y_{1,it} = 1|\Omega_t\}}{P\{Y_{1,it} = 1|\Omega_t, z_{it}\}} \\ &= w_{2,it} w_{1,it}. \end{aligned} \tag{9}$$

If $w_{2,it} = 1$ for all i, t, then self-selection is exogenous and does not lead to look-ahead bias in measures for performance (persistence). In this case, liquidation implies look-ahead bias if $w_{1,it} \neq 1$ and this is the case analyzed by Baquero et al. (2005). In this paper, we disentangle the two sources of bias by identifying both sets of weights and applying corrections with one weight or their product. The correction for self-selection is conditional upon the fund not liquidating.[8] The application of the above correction weights allows us to determine to what extent we get different results if we correct only for selection bias due to liquidation, assuming self-selection is random.

To identify the weights (and to derive Equation (8)) we need to assume that the probabilities do not depend upon future, potentially unobserved returns. Further, we assume that self-selection and fund liquidation are mutually exclusive events, and both describe "absorbing states". That is, once a fund stops reporting to TASS, it will not return in the database at a later stage. Then the denominator of $w_{1,it}$ can be determined from the binary choice model as

$$\begin{aligned} &P\{Y_{1,it} = 1|\Omega_t, z_{it}\} \\ &P\{L_{i,t+1} = 1|r_{it}, r_{i,t-1}, \ldots, x_{i,t+1}\} \ldots \\ &P\{L_{i,t+s+1} = 1|r_{i,t+s}, r_{i,t+s-1}, \ldots, x_{i,t+s+1}\} \end{aligned} \tag{10}$$

and similarly for $w_{2,it}$. The right-hand side probabilities are described by the probit model in Equation (1) provided the appropriate functional form (and conditioning variables) are chosen in x_{it}. Consequently, consistent estimation of the binary choice models for liquidation and self-selection allows us to obtain consistent estimators for the two sets of weights, which enables us to

[8] Statistically, Equation (9) also holds with the role of $Y_{1,it}$ and $Y_{2,it}$ reversed, so that the correction for liquidation bias would be conditional upon the fund not stopping reporting. Given that existing literature (Baquero et al., 2005) assumes self-selection is exogenous $(w_{2,it} = 1)$, the most natural ordering is employed here.

correct for look-ahead bias due to these two processes and separate their effects upon performance measures and their persistence. To estimate the numerator in Equation (9) when Ω_t takes on a limited number of different values (e.g., past performance decile), it is most convenient to use a simple nonparametric approach (see below).

4. Persistence in Hedge Fund Performance

Empirical studies on the behavior of investors in hedge funds have shown that money-flows chase past performance (see, e.g., Agarwal et al., 2004; Baquero and Verbeek, 2006). Moreover, several recent studies document some evidence of persistence in hedge fund performance at quarterly horizons (see, e.g., Agarwal and Naik, 2000; Bares et al., 2003; Boyson and Cooper, 2004), while at longer horizons the results are more ambiguous (see, e.g., Brown et al., 1999; Brown and Goetzmann, 2003; Kosowski et al., 2006). Apparently, if investors take into account that persistence is mainly a short-term phenomenon, looking at past performance provides potentially valuable information for investing in hedge funds. However, while all the above-mentioned studies on performance persistence of hedge funds control for the effects of survivorship bias, none of them corrects for look-ahead bias due to fund liquidation or self-selection. Baquero et al. (2005) correct for look-ahead bias due to liquidation and find positive persistence at horizons of one and four quarters, although the statistical significance is weak. In the previous section we extended the methodology of the latter paper to allow us to disentangle the effects of look-ahead bias and self-selection bias. In this section we will apply this method on analyzing performance persistence of hedge funds, and examine whether look-ahead biases due to liquidation and self-selection offset each other.

First, we will examine performance persistence in raw returns by examining whether winning funds over the last two or four quarters are more likely to be top performers in the future. We follow a procedure similar to Baquero et al. (2005), by distinguishing a ranking period of two or four quarters and an evaluation period of two or four quarters. The ranking based on past performance is broken down into ten deciles (decile 10 contains the past winners), and in the subsequent evaluation period we calculate the average returns of each of these deciles. The procedure is repeated over the entire sample period, moving forward by one quarter at a time and adjusting the sample by including the funds that have a sufficiently long return history. Fund-of-funds are excluded to avoid double-counting. To avoid backfilling bias, returns are used in this exercise only if the fund has a history of at least four quarters.

First of all, in order to prevent spurious performance persistence patterns that are due to look-ahead bias (see, e.g. Carpenter and Lynch, 1999), we apply the correction method as introduced by ter Horst et al. (2001). Basically, we repeat the analysis of Baquero et al. (2005) by multiplying the performance measure (e.g., average return over the ranking period) with a estimated weight factor $\hat{w}_{1,it}$, which is the ratio of an unconditional nonliquidation probability and a conditional nonliquidation probability. The latter probability can be obtained from our estimated liquidation process reported in Section 3. Let

$$\begin{aligned}\hat{p}_{it} &= \hat{P}\{L_{it} = 1 | r_{i,t-1}, r_{i,t-2}, \ldots, x_{it}\} \\ &= \Phi\left(\hat{\alpha}_1 + \sum_{j=1}^{6} \hat{\gamma}_{1j} r_{i,t-j} + x'_{it}\hat{\beta}_1\right) \end{aligned} \tag{11}$$

denote the estimated (conditional) probability that fund i does not liquidate in period t, where Φ denotes the standard normal distribution function. Then the denominator of $w_{1,it}$ in Equation (9) is estimated as

$$\prod_{\tau=t+1}^{t+s+1} \hat{p}_{i\tau}, \tag{12}$$

where $s = 4$ (quarters) in case of annual persistence. The unconditional probability is equal to the ratio of funds that were not liquidated during the ranking period and the number of funds present at the beginning of the period. For the evaluation period, we compute average returns within each decile, again weighted by $\hat{w}_{1,it}$, where the numerator now corresponds to the proportion of survived funds in the corresponding decile.

Next, we correct for self-selection bias by multiplying the performance measure with a second weight factor $w_{2,it}$. This factor is the ratio of the conditional probability of non-self-selection (conditional upon not being liquidated) and an unconditional non-self-selection probability (conditional upon not being liquidated). The conditional probability can be obtained from the estimated self-selection process of Section 3. The unconditional probability is now equal to the ratio of the number of funds that were not self-selected minus those that were liquidated during the ranking period, and the number of funds present at the beginning of the ranking period minus those that were liquidated during the ranking period. Similarly, we correct the average returns over the evaluation period once more, but adjusting for the fact that the unconditional probabilities are now conditional upon the fund's decile during the ranking period.

Table V. Persistence estimates (raw returns)

Each quarter, funds are sorted into ten rank portfolios based on their previous two-quarters. Next, average returns over the next two quarters are computed, for each decile. Using returns from 1994–2000, this produces a time-series for each decile of 21 average two-quarter returns. The numbers in the table are the annualized time-series averages and their standard errors in parentheses. The standard errors are corrected for autocorrelation based on the Newey–West approach. The corrected figures employ a weighting procedure to eliminate liquidation bias, and the double corrected employ a weighting procedure to eliminate liquidation and self-selection bias.

Average performance (raw returns)				
	Two quarters			
Decile	(1) Non corrected	(2) Corrected	(3) Double corrected	(1) − (3)
1 (losers)	0.1284	0.1139	0.1090	0.0193
	(0.0794)	(0.0932)	(0.0987)	(0.0229)
2	0.1278	0.1253	0.1250	0.0028
	(0.0335)	(0.0374)	(0.0380)	(0.0067)
3	0.1308	0.1345	0.1356	−0.0048
	(0.0291)	(0.0300)	(0.0298)	(0.0046)
4	0.1302	0.1322	0.1333	−0.0031
	(0.0242)	(0.0246)	(0.0246)	(0.0029)
5	0.1230	0.1297	0.1314	−0.0085
	(0.0218)	(0.0249)	(0.0253)	(0.0058)
6	0.1492	0.1510	0.1503	−0.0011
	(0.0255)	(0.0258)	(0.0242)	(0.0061)
7	0.1602	0.1650	0.1694	−0.0091
	(0.0296)	(0.0286)	(0.0310)	(0.0068)
8	0.1848	0.1892	0.1891	−0.0042
	(0.0416)	(0.0450)	(0.0445)	(0.0072)
9	0.2112	0.2090	0.2095	0.0017
	(0.0529)	(0.0512)	(0.0518)	(0.0054)
10 (winners)	0.2448	0.2278	0.2283	0.0165
	(0.0810)	(0.0798)	(0.0800)	(0.0108)
winners − losers	0.1164	0.1139	0.1193	−0.0028
	(0.0918)	(0.0984)	(0.1027)	(0.0247)

The results of the above exercises are provided in Tables V and VI for the two-quarter and four-quarter horizon, respectively, and summarized in Figures 1 and 2. We report the empirical persistence of raw returns at bi-quarterly and annual horizons, without any correction (raw returns), with a correction for look-ahead bias due to liquidation (corrected returns) and with a correction for look-ahead bias due to both liquidation and self-selection (double corrected returns). All estimates are based on the full sample of hedge funds, excluding fund-of-funds.

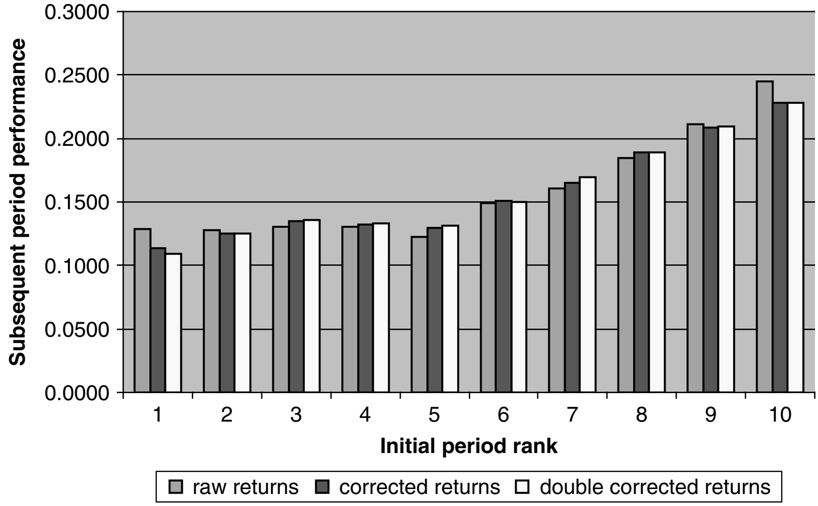

Figure 1. Two-quarterly persistence in raw returns.

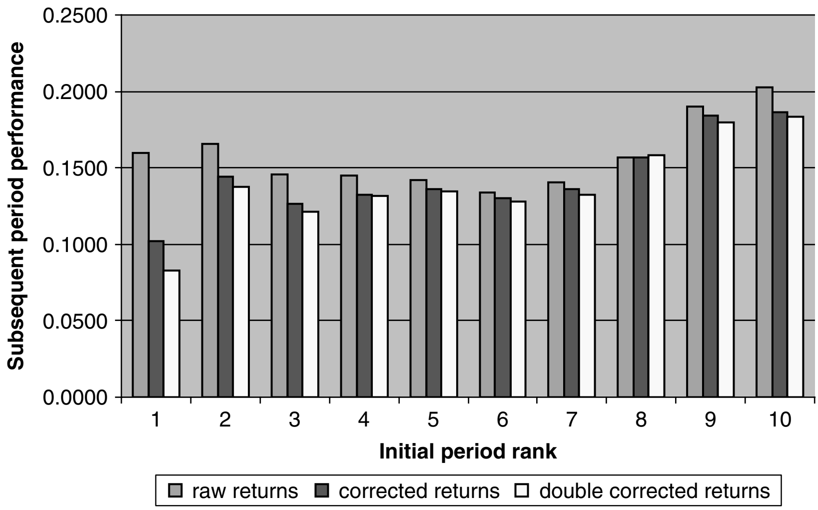

Figure 2. Four-quarterly persistence in raw returns.

Table VI. Persistence estimates (raw returns)

Each quarter, funds are sorted into ten rank portfolios based on their previous four-quarter returns, respectively. Next, average returns over the next four quarters are computed, for each decile. Using returns from 1994–2000, this produces a time-series for each decile of 17 (overlapping) average four-quarter returns. The numbers in the table are the annualized time-series averages and their standard errors in parentheses. The standard errors are corrected for autocorrelation based on the Newey–West approach. The corrected figures employ a weighting procedure to eliminate liquidation bias, and the double corrected employ a weighting procedure to eliminate liquidation and self-selection bias.

	Average performance (raw returns)			
	Four quarters			
Decile	(1) Non corrected	(2) Corrected	(3) Double corrected	(1) − (3)
1 (losers)	0.1601	0.1018	0.0832	0.0769
	(0.0911)	(0.0946)	(0.0956)	(0.0207)
2	0.1658	0.1446	0.1374	0.0284
	(0.0600)	(0.0545)	(0.0526)	(0.0141)
3	0.1459	0.1262	0.1216	0.0243
	(0.0455)	(0.0404)	(0.0401)	(0.0134)
4	0.1451	0.1328	0.1318	0.0132
	(0.0387)	(0.0370)	(0.0351)	(0.0081)
5	0.1418	0.1359	0.1345	0.0073
	(0.0252)	(0.0246)	(0.0271)	(0.0054)
6	0.1342	0.1304	0.1280	0.0062
	(0.0318)	(0.0349)	(0.0366)	(0.0090)
7	0.1403	0.1361	0.1327	0.0076
	(0.0381)	(0.0398)	(0.0397)	(0.0099)
8	0.1565	0.1566	0.1580	−0.0015
	(0.0325)	(0.0337)	(0.0337)	(0.0085)
9	0.1900	0.1840	0.1798	0.0102
	(0.0468)	(0.0491)	(0.0461)	(0.0094)
10 (winners)	0.2029	0.1862	0.1838	0.0191
	(0.1016)	(0.0979)	(0.1001)	(0.0120)
Winners − losers	0.0428	0.0844	0.1006	−0.0578
	(0.0693)	(0.0724)	(0.0713)	(0.0186)

The results are remarkable. First of all at a bi-quarterly as well as an annual horizon we observe evidence of a persistence pattern which is without corrections (raw returns) slightly J-shaped. As discussed in Hendricks et al. (1997) and ter Horst et al. (2001) such a pattern could be explained by look-ahead bias. If we correct for look-ahead bias, the expected returns on the lower deciles are significantly reduced and the J-shape flattens. As a result, the relationship between past and future performance becomes more monotonic. The expected returns on a zero investment portfolio that is long in winners (decile 10) and short in losers (decile 1) is approximately 10.0% at the annual horizon, while it is only 4.3% if no correction for look-ahead bias is applied. The difference is statistically significant with a t-value of -3.11. At the bi-quarterly horizon, the winner–loser portfolio has an expected return of 11.9%, and the correction for look-ahead bias has very little impact. Brown et al. (1999), Agarwal and Naik (2000), and Bares et al. (2003), also find evidence of a persistence pattern at a short-term horizon, while the pattern is less strong at longer horizons. However, note that these studies do not correct for look-ahead bias, and that without corrections, average returns may be overestimated by as much as 7.7% (decile 1, annual horizon). For additional discussion on correcting for look-ahead bias we refer to Baquero et al. (2005).

The most striking result is that the correction for look-ahead bias due to self-selection works in the same direction as the one for liquidation bias, and thus strengthens the positive persistence patterns. For most of the deciles we see that the two biases have the same sign. Apparently, the majority of funds that stop reporting voluntarily have lower returns than those that continue to report, consistent with Tables I and IV. Accordingly, there is no counterbalancing effect that reduces the liquidation bias owing to relatively successful funds that close for new investments and stop reporting. Thus, while Ackermann et al. (1999) argue that positive and negative survival-related biases may offset each other, this does not apply to look-ahead bias due to liquidation and self-selection. Because this conclusion may be related to the way we define liquidation and self-selection cases, the next section provides some robustness checks.

5. Robustness Checks

As mentioned before, it is difficult to pin down what exactly self-selection means for hedge funds. For example, a fund may decide to stop reporting because it is too small and subsequently decide to liquidate several quarters later. Our classification of attrition cases so far was determined by the qualitative information provided by TASS and, when necessary, by the sign of the fund's cash flows in the last available year. We have experimented with alternative classification procedures, resulting in qualitatively similar results when we analyze performance persistence. To be precise, the post-ranking expected returns for all deciles after correcting for both liquidation and self-selection bias are highly robust, although the individual magnitude of the two biases depends upon the exact classification scheme. The latter finding can be attributed to the problem that it is not unambiguous whether a given situation should be classified as self-selection or liquidation.

In this section we report the results when we classify disappearance reasons 2 (“at fund manager’s request”) and 4 (“unknown”) into positive and negative events by using additional information on fund flows and returns. In particular, we classify a case as “liquidation” when a fund experiences negative cash flows in its last year, has negative cumulative returns over the last year, and has negative returns for each of the last two quarters of its reporting period. All other cases are treated as “self-selection”. Because empirical studies report that flows are highly linked to past performance (Agarwal et al., 2004; Baquero and Verbeek, 2006), this may allow us to better identify funds that are expected to liquidate in the future (because of large negative flows after their reporting period). Compared to the previous classification, 19 cases move from the self-selection to the liquidation bin.

Because the classification of cases is based on past returns, this obviously has some impact upon the coefficient estimates for the lagged returns in the binary probit models explaining liquidation and self-selection. For the liquidation model, the cumulative estimated impact of past returns increases slightly from 3.394 (Table III) to 3.775, not taking into account the underwater dummy. For the self-selection model, the cumulative impact reduces from 4.386 (Table IV) to 3.364. Apart from this, the estimation results for these two models, with and without net asset value (available upon request), are not noticeably different from before. Note that, even with this new classification, the probability of self-selection is negatively (and statistically significantly) related to past performance.

We use these new models to correct the persistence analysis for look-ahead bias due to liquidation and self-selection. As before, look-ahead bias has limited impact at the two-quarter horizon, reported in Table VII, but substantial impact at the four-quarter horizon, reported in Table VIII. Given that poor past returns are used to classify a case as liquidation, it is not surprising to see that the bias due to fund liquidation has increased in magnitude, at the cost of the self-selection bias. However, the expected returns for each of the deciles, after correcting for both sources of bias, are virtually the same as those reported in Tables V and VI. To some extent this reflects the ambiguity in classifying a case as either “liquidation” or “self-selection”. Two conclusions are clear, however. First, no matter how we define it, self-selection has no offsetting impact upon the liquidation bias. Second, it is inappropriate to only correct for liquidation bias while at the same time restricting attention to those cases that are classified by TASS as “liquidated”.

Table VII. Persistence Estimates (Raw returns), alternative classification

Each quarter, funds are sorted into ten rank portfolios based on their previous two-quarters. Next, average returns over the next two quarters are computed, for each decile. Using returns from 1994–2000, this produces a time-series for each decile of 21 average two-quarter returns. The numbers in the table are the annualized time-series averages and their standard errors in parentheses. The standard errors are corrected for autocorrelation based on the Newey–West approach. The corrected figures employ a weighting procedure to eliminate liquidation bias, and the double corrected employ a weighting procedure to eliminate liquidation and self-selection bias. This table is based on an alternative classification of cases as "liquidated" or "self-selected" (see text).

	Average performance (raw returns)			
	Two quarters			
Decile	(1) Non corrected	(2) Corrected	(3) Double corrected	(1) − (3)
1 (losers)	0.1284	0.1094	0.1086	0.0198
	(0.0794)	(0.0967)	(0.0990)	(0.0233)
2	0.1278	0.1252	0.1248	0.0028
	(0.0335)	(0.0375)	(0.0280)	(0.0066)
3	0.1308	0.1344	0.1354	−0.0046
	(0.0291)	(0.0301)	(0.0297)	(0.0045)
4	0.1302	0.1318	0.1332	−0.0030
	(0.0242)	(0.0246)	(0.0257)	(0.0028)
5	0.1230	0.1290	0.1310	−0.0080
	(0.0218)	(0.0252)	(0.0259)	(0.0062)
6	0.1492	0.1510	0.1516	−0.0024
	(0.0255)	(0.0235)	(0.0238)	(0.0064)
7	0.1602	0.1658	0.1682	−0.0080
	(0.0296)	(0.0303)	(0.0310)	(0.0069)
8	0.1848	0.1890	0.1886	−0.0038
	(0.0416)	(0.0448)	(0.0447)	(0.0072)
9	0.2112	0.2078	0.2100	0.0012
	(0.0529)	(0.0506)	(0.0516)	(0.0061)
10 (winners)	0.2448	0.2288	0.2288	0.0160
	(0.0810)	(0.0806)	(0.0799)	(0.0109)
Winners − losers	0.1164	0.1194	0.1204	−0.0040
	(0.0918)	(0.1018)	(0.1031)	(0.0250)

6. Concluding Remarks

When analyzing hedge fund performance and its persistence, a multiperiod sampling bias or look-ahead bias may arise if funds attrite from the available databases due to reasons that relate to their performance. In this paper, we consider two important reasons why funds may disappear from hedge fund

Table VIII. Persistence Estimates (Raw returns), alternative classification

Each quarter, funds are sorted into ten rank portfolios based on their previous four-quarter returns, respectively. Next, average returns over the next four quarters are computed, for each decile. Using returns from 1994–2000, this produces a time-series for each decile of 17 (overlapping) average four-quarter returns. The numbers in the table are the annualized time-series averages and their standard errors in parentheses. The standard errors are corrected for autocorrelation based on the Newey–West approach. The corrected figures employ a weighting procedure to eliminate liquidation bias, and the double corrected employ a weighting procedure to eliminate liquidation and self-selection bias. This table is based on an alternative classification of cases as "liquidated" or "self-selected" (see text).

	Average performance (raw returns)			
	Four quarters			
Decile	(1) Non corrected	(2) Corrected	(3) Double corrected	(1) − (3)
1 (losers)	0.1601	0.0922	0.0826	0.0775
	(0.0911)	(0.0960)	(0.0957)	(0.0210)
2	0.1658	0.1424	0.1370	0.0288
	(0.0600)	(0.0542)	(0.0528)	(0.0142)
3	0.1459	0.1235	0.1213	0.0246
	(0.0455)	(0.0408)	(0.0400)	(0.0134)
4	0.1451	0.1323	0.1321	0.0129
	(0.0387)	(0.0365)	(0.0350)	(0.0081)
5	0.1418	0.1372	0.1344	0.0074
	(0.0252)	(0.0242)	(0.0270)	(0.0054)
6	0.1342	0.1288	0.1290	0.0052
	(0.0318)	(0.0369)	(0.0361)	(0.0089)
7	0.1403	0.1340	0.1341	0.0090
	(0.0381)	(0.0396)	(0.0400)	(0.0095)
8	0.1565	0.1556	0.1581	−0.0016
	(0.0325)	(0.0337)	(0.0339)	(0.0087)
9	0.1900	0.1835	0.1806	0.0093
	(0.0468)	(0.0479)	(0.0471)	(0.0097)
10 (winners)	0.2029	0.1844	0.1824	0.0204
	(0.1016)	(0.1006)	(0.0995)	(0.0131)
Winners − losers	0.0428	0.0922	0.0998	−0.0570
	(0.0693)	(0.0728)	(0.0720)	(0.0188)

databases. First, funds may liquidate or close down owing to their poor performance, and, second, hedge fund managers may voluntary stop reporting to a database vendor (self-selection). In this paper we empirically investigate the determinants of fund liquidation and self-selection, and analyze the impact of these two process upon persistence measures of hedge fund performance.

Using information from the TASS database, covering the period 1994–2000, we find that both liquidation and self-selection are more likely for hedge funds that have a poor return history. While the relationship is somewhat stronger for the liquidation process, this implies that look-ahead bias due to self-selection affects persistence measures in the same direction as does look-ahead bias due to fund liquidation. Consequently, double correcting persistence tables leads to a stronger persistence pattern than obtained in Baquero et al. (2005), where look-ahead bias due to liquidation is the focus of interest. At the annual horizon, the expected excess return on a winner minus loser portfolio, based upon previous year returns, is close to 10% when both biases are taken into account, while it is only 4.3% if no correction is employed. These biases are almost entirely located in the bottom decile, where expected returns may be significantly overestimated by almost 8% per year.

Appendix

This appendix contains the estimation results for the binary liquidation and self-selection models, when fund size is excluded from the conditioning set. These models, estimated over a larger sample, are used to correct for liquidation and self-selection bias in cases where fund NAV is unobserved.

Table IX. Estimation results liquidation model

The table reports the estimation results of the liquidation model, excluding net asset value (size); 21,297 fund/period observations.

Parameters	Estimate	Std.error	Parameters	Estimate	Std. Error
Intercept	3.986	0.719	Offshore	−0.110	0.049
$r(-1)$	1.250	0.201	Incentive Fees	−0.008	0.003
$r(-2)$	1.065	0.216	Mng. Fees	−0.018	0.023
$r(-3)$	1.282	0.217	Underwater	−0.340	0.063
$r(-4)$	0.499	0.219	Emerging markets	−0.017	0.078
$r(-5)$	0.252	0.178	Equity market neutral	−0.162	0.088
$r(-6)$	0.012	0.162	Event driven	0.155	0.102
			Fixed income arbitrage	−0.067	0.200
StDev	0.522	0.352	Global macro	0.019	0.174
ln(Age)	−0.343	0.363	Long/short equity	−0.133	0.073
ln(Age)^2	0.061	0.048	Managed futures	−0.285	0.067

Loglikelihood: −1817.495 | Chi-squared test: 502.38 ($DF = 41$)
pseudo R^2: 0.1214 | ($p = 0.000$)

Table X. Estimation results self-selection model

Estimation results self-selection model, excluding net asset value (size); 20,876 fund/period observations.

Parameters	Estimate	Std.error	Parameters	Estimate	Std. Error
Intercept	4.107	1.085	Offshore	0.104	0.074
$r(-1)$	1.575	0.343	Incentive Fees	−0.013	0.005
$r(-2)$	1.303	0.347	Mng. Fees	−0.049	0.034
$r(-3)$	0.668	0.325	Underwater	0.031	0.112
$r(-4)$	0.608	0.340	Emerging markets	0.025	0.128
$r(-5)$	0.030	0.253	Equity market neutral	0.061	0.145
$r(-6)$	0.914	0.382	Event driven	0.190	0.148
			Fixed income arbitrage	0.105	0.363
StDev	0.652	0.566	Global macro	0.287	0.345
ln(Age)	−0.557	0.570	Long/short equity	−0.117	0.104
$\ln(\text{Age})^2$	0.088	0.075	Managed futures	−0.043	0.105
Loglikelihood: −693.909 pseudo R^2: 0.0861			Chi-squared test: 130.71 ($DF = 41$) ($p = 0.000$)		

References

Ackermann, C., McEnally, R. and Ravenscraft, D. (1999) The performance of hedge funds: risk, return and incentives, *Journal of Finance* **54**, 833–874.

Agarwal, V., Daniel, N. D. and Naik, N. Y. (2004) Flows, Performance, and Managerial Incentives in the Hedge Fund Industry, EFA 2003 Glasgow Meetings, available at SSRN: http://ssrn.com/abstract=424369.

Agarwal, V. and Naik, N. Y. (2000) Multi-period performance persistence analysis of hedge funds, *Journal of Financial and Quantitative Analysis* **35**, 327–342.

Baquero, G., ter Horst, J. R. and Verbeek, M. (2005) Survival, look ahead bias and the persistence in hedge fund performance, *Journal of Financial and Quantitative Analysis* **40**, 493–518.

Baquero, G. and Verbeek, M. (2006) A Portrait of Hedge Fund Investors: Flows, Performance and Smart Money, EFA 2006 Zurich Meetings, available at SSRN: http://ssrn.com/abstract=773384.

Bares, P. A., Gibson, R. and Gyger, S. (2003) Performance in the hedge fund industry: an analysis of short and long term persistence, *Journal of Alternative Investments* **6**, 25–41.

Berk, J. B. and Green, R. C. (2004) Mutual fund flows and performance in rational markets, *Journal of Political Economy* **112**, 1269–1295.

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

Boyson, N. M. and Cooper, M. J. (2004) Do Hedge Funds Show Performance Persistence? A New Approach, working paper, Northeastern University, available at SSRN: http://ssrn.com/abstract=686572.

Brown, S. J. and Goetzmann, W. N. (2003) Hedge funds with style, *Journal of Portfolio Management* **29**, 101–112.

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

Brown, S. J., Goetzmann, W. N. and Ibbotson, R. G. (1999) Offshore hedge funds: survival and performance 1989–1995, *Journal of Business* **72**, 91–117.

Brown, S. J., Goetzmann, W. N. and Park, J. (2001) Careers and survival: competition and risk in the hedge fund and CTA industry, *Journal of Finance* **56**, 1869–1886.

Capocci, D. and Hübner, G. (2004) Analysis of hedge fund performance, *Journal of Empirical Finance* **11**, 55–89.

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

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

Elton, E. J., Gruber, M. J. and Blake, C. R. (1996) Survivorship bias and mutual fund performance, *Review of Financial Studies* **9**, 1097–1120.

Fitzgerald, F., Gottschalk, P. and 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.

Fung, W. and Hsieh, D. A. (1997) Empirical characteristics of dynamic trading strategies: the case of hedge funds, *Review of Financial Studies* **10**, 275–302.

Liang, B. (2000) Hedge funds: The living and the dead, *Journal of Financial and Quantitative Analysis*, **35**, 309–326.

Getmanksy, M., Lo, A. W. and Makarov, I. (2004) An econometric model of serial correlation and illiquidity in hedge fund returns, *Journal of Financial Economics* **74**, 529–609.

Goetzmann, W.N., Ingersoll, J. and Ross, S.A. (2003) High-water marks and hedge fund management contracts, *Journal of Finance* **58**, 1685–1718.

Hendricks, D., Patel, J. and Zeckhauser, R. (1997) The J-shape of performance persistence given survivorship bias, *The Review of Economics and Statistics* **79**, 161–170.

ter Horst, J. R., Nijman, T. and Verbeek, M. (2001) Eliminating look-ahead bias in evaluating persistence in mutual fund performance, *Journal of Empirical Finance* **8**, 345–373.

ter Horst, J. R. and Verbeek, M. (2000) Estimating short-run persistence in mutual fund performance, *The Review of Economics and Statistics* **82**, 646–655.

Huij, J. and Verbeek, M. (2007) Cross-sectional learning and short-run persistence in mutual fund performance, *Journal of Banking and Finance*, forthcoming; available at SSRN: http://ssrn.com/abstract=567701.

Jagannathan, R., Malakhov, A. and Novikov, D. (2006) Do Hot Hands Persist Among Hedge Fund Managers? An Empirical Evaluation, NBER Working Paper W12015, available at SSRN: http://ssrn.com/abstract=881249.

Kosowski, R., Naik, N. Y. and Teo, M. (2007) Do hedge funds deliver alpha? A bayesian and bootstrap analysis, *Journal of Financial Economics*, forthcoming.

Malkiel, B. (1995) Returns from investing in equity funds 1971 to 1991, *Journal of Finance* **50**, 549–572.

Malkiel, B. G. and Saha, A. (2005) Hedge funds: risk and return, *Financial Analysts Journal* **61**, 80–88.

Posthuma, N. and van der Sluis, P. J. (2003) A Reality Check on Hedge Funds Returns, working paper, available at SSRN: http://ssrn.com/abstract=438840.

Sirri, E. R. and Tufano, P. (1998) Costly search and mutual fund flows, *Journal of Finance* **53**, 1589–1622.
Wermers, R. (2003) Is Money Really “Smart”? New Evidence on the Relation between Mutual Fund Flows, Manager Behavior, and Performance Persistence, working paper, University of Maryland, available at SSRN: http://ssrn.com/abstract=414420.