World Review of Business Research
Vol. 4. No. 1. March 2014 Issue. Pp. 76 – 89

Does Conditioning Bias Exist in Asian Hedge Funds' Returns?

Lan Thi Phuong Nguyen*, Ming Yu Cheng**, Sayed Hossain***
and Malick Ousmane Sy****

The availability of various databases raises an important question on the reliability of hedge fund data. Researchers using different databases may derive different conclusions on the performance of hedge funds, either overstating or understating the funds' performance. The differences can be explained by a phenomenon called "data conditioning bias" as defined by Ackermann, McEnally, and Ravenscraft (1999). There are six forms of biases, namely survivorship, termination, self-selection, liquidation, backfilling, and multi-period sampling biases. In this paper, we focus on three biases that might exist in a chosen sample of Asian hedge funds. These biases are survivorship bias, multi-period sampling bias, and omission bias as defined in the paper. We choose a crisis-free period, i.e. 1st January 2000 and 15th June 2008 for the study. We select the largest possible sample size (456 funds) with at least 4-year-monthly data. Our results showed that the studied sample has passed through the test with no significant survivorship bias, no significant multi-period sampling as well as omission bias. Overall, we could conclude that the sample is fit to be used for further analysis in Asian hedge funds.

1. Introduction

Although hedge fund industry has existed for more than 50 years since its first introduction in 1940, information about hedge funds is not as widely available as compared to other types of traditional investment. This is mainly due to the specific characteristic of hedge funds, which is subjected to modest regulation by Security Commissions (SECs). Basically, hedge fund managers are not allowed to advertise their investment publicly, thus they provide information about their funds to certain data vendors, on voluntary basis. Hedge fund managers often choose to report their funds' performance to a data vendor for two main purposes: (1) to inform the existing investors of their funds' current performance and (2) to attract new ones to increase their funds' assets. Thus, hedge fund data are only available to the "qualifying public" (Cappocci & Hübner, 2004). There are only a handful of data vendors supplying universal hedge fund data, i.e. Paradigm LDC, TASS Management Limited (TASS), Hedge Fund Research Inc. (HFR), Managed Account Reports Inc. (MAR), Asia Hedge and Eurekahedge Inc.. Most of the researches on hedge funds are done based on the data provided by the above databases.

* Dr. Lan Thi Phuong Nguyen, Faculty of Management, Multimedia University, Malaysia, Email: nguyen.thi.phuong.lan@mmu.edu.my
**Prof. Dr. Ming Yu Cheng, Faculty of Accountancy and Management, Universiti Tunku Abdul Rahman, Malaysia, Email: chengmy@utar.edu.my
***Dr. Sayed Hossain, Faculty Associate, Colline College, the United States of America, Email: sydhos@gmail.com
****Prof. Dr. Malick Ousmane Sy, School of Economics, Finance, and Marketing, College of Business, RMIT University, Australia, Email: malick.sy@rmit.edu.au

The last two data vendors are specializing in providing Asian Hedge Fund (AHF) data. The availability of different databases creates an issue on the reliability of hedge fund data as researchers who use different hedge fund databases may end up with different conclusions on the performance of hedge funds, either overstating or understating their performances. The differences in findings can be explained by the phenomenon called “data conditioning bias”, which often exists in hedge fund and mutual fund returns.

According to Ackermann, McEnally, and Ravenscraft (1999), data conditioning bias can be classified into six forms, namely survivorship, termination, self-selection, liquidation, backfilling, and multi-period sampling. It is stressed by Ball and Watts (1979) that a frequent problem in the time series analysis of firms’ financial data is the need to impose survivorship criteria in order to obtain sufficient observations for estimation purposes.

Survivorship bias is defined by Ackermann et al. (1999) as the bias of considering only the performance of funds that are alive and present in the database at the end of the sample period. Termination and self-selection biases are two subsets of survivorship bias. As information on hedge funds are reported to the data vendors on voluntary basis, the supply of fund information may stop for two reasons. Firstly, if a fund performs poorly, the fund may die and exit from the database, this is called discontinuation. Secondly, when a fund has raised enough capital and does not need to attract more investors, it will stop reporting its status to the data vendor and this is referred to as self-selection. Excluding funds that performed poorly and died will cause an upward bias in hedge funds’ average return. Excluding well-performing funds that do not want to increase their capital anymore may cause a downward bias in hedge funds’ average return. Funds that experience liquidation often take sometimes to liquidate their funds after declaring as dead funds. Thus their post-liquidation performance, which is often not good, might be reported in the database. If the performance of funds with and without the post-reporting returns appears to be significantly different, then liquidation bias exists.

When a new fund is added, the database providers typically request for the full performance history of the fund, thus only funds that survived the backfilling period are included. If the performance of funds with and without backfilling period is significantly different, then backfilling bias exists. Researchers often choose to study samples of funds that exist for a certain period of time. Multi-period sampling bias exists when funds that failed to survive for the whole period and new funds that emerged during that period are excluded. Conditioning on survival over multiple years may impart an upward performance bias.

Acknowledging the possibility of providing misleading information to investors as well as fund researchers, a number of hedge funds providers have begun to keep data of funds that stop reporting to them. Examples are HFR who started to do this in December 1992 and MAR in December 1993. In the case of the Eurekahedge database, it has kept all dead funds’ details since its establishment. For some special cases, if funds are requested to be removed from the database due to some reasons, their historical data will not be reported in the database. However, there are a few funds that do so.

In this study, we attempt to test whether conditioning biases exists in AHF data provided by Eurekahedge Inc. for the period of 1st January 2000 - 15th June 2008.

Results of the study are expected to provide a justification on the accuracy of research findings produced for AHFs during the above-mentioned study period. AHFs that are selected for this study are required to have at least 4 years of full monthly data, which means funds that were introduced after July 2004 are not included in our sample. Given the practical implication of the study, we examine two common conditioning biases, i.e. survivorship bias and multi-period sampling bias for our sample of AHFs. In addition, due to the unique feature of the Eurekahedge database Inc., we attempt to test a new type of data conditioning bias, named by us as "omission bias". We define the omission bias as an upward bias that could occur when dead funds exist from the database as requested by respective managers. This study can be considered as the first study that examines the three types of data bias as mentioned on AHFs with a large sample size (456 funds) over a long study period (1st January 2000 – 15th June 2008). The overall results suggested that the selected sample of AHFs was bias-free; thus, fit to be used for further tests.

The rest of the paper is organized as follows: related studies are reviewed in Part 2, followed by Part 3 which discusses hypothesis and testing methods. Main findings are presented in Part 4, and key conclusions are given in Part 5.

2. Literature Review

2.1 Survivorship Bias

It is often believed that competition is much fiercer in the hedge fund industry; thus, more funds are expected to disappear as compared to those in the mutual fund industry. However, many studies show that this may not be the case. In Fung and Hsieh (1997), the disappearing rate (4.3% per year) found for hedge funds is somewhat the same with what is found for mutual funds in Grinblatt and Titman (1989) and similar to the rate (4.8% per year) found in Brown, Goetzmann, Ibbotson, and Ross (1992). Liang (1999) also finds a low attrition rate of 10% for several years by 1996 for hedge funds. However, the disappearance of hedge funds certainly causes a survivorship bias in a hedge fund database. According to Liang (1999), survivorship bias for hedge funds should be between 0.5% and 3.4% per year. This is because the author takes into account the lowest possible bias found in mutual fund studies, i.e. 0.5% per year, and the survivorship bias found for CTA in Fung and Hsieh (1997), i.e. 3.4% per year.

Results on survivorship bias for hedge funds found for the period prior to 1994 are somewhat interesting. Firstly, when a surviving condition is not imposed on a sample of hedge funds, there is hardly any significant evidence found for survivorship bias in the set of hedge fund data (Ackermann et al., 1999). Ackermann et al. (1999) measure survivorship bias between the performance of surviving and disappearing funds (the segregated approach) during the period of 1988 and 1995 by employing a combined sample (1,272 funds) selected from two databases, i.e. MAR and HFR. The authors name all funds that survive until December 1995 as extant funds. Disappearing funds are weighted with the number of months they were in existence in order to scale down the impact of short history funds on the overall results. Results show that survivorship bias for hedge funds during the study period is about 0.16% per year, which is below the estimated range (0.5% – 1.4% per year) for mutual funds. Although disappearing funds are found to be below-average performers, the evidence for this is weak and insignificant. There is also no strong or significant evidence that disappearing funds are highly volatile funds. In examining termination bias, the authors find some significant evidence that these extant funds outperform the set of terminating funds, i.e. in three out of five cases. When testing the self-selection bias, although evidence shows that the set of self-selected funds outperforms the set of extant funds in five out of six cases, those differences are insignificant. Overall, results support the fact that funds with either superior or inferior performances tend to voluntarily stop reporting to a database.

In another study by Brown, Goetzmann, and Ibbotson (1999), survivorship bias is found in the sample of U.S. offshore hedge funds for the period of 1989 to 1995. In this study, a sample of 399 offshore hedge funds from the *U.S. Offshore Fund Directory* is used. The *U.S. Offshore Fund Directory* is known as one of the few databases that keep a record of defunct funds. Two samples with different surviving conditions are employed in the study. In the first sample, the selected surviving funds are required to survive for seven years until the end of the study period, which is 1995. For the second sample, all funds are required to survive until the end of the sample period, which is 1995. To obtain a more accurate estimate of the performance of hedge funds, the authors use three different indices: (1) an equally-weighted index constructed from the full sample, (2) an equally-weighted index constructed from the set of extant funds, and (3) an equally-weighted index constructed from the set of funds that survived throughout the entire seven-year period. Their results reveal a bias of about 2.75% per year in raw returns. Given the fact that onshore and offshore U.S. hedge funds often employ almost the same strategies, their findings may represent both onshore and offshore hedge funds (Fung and Hsieh, 2000). The study implies that using survival conditioning may lead to a significant difference in performance between ex post observed returns and ex ante expected future returns. Thus, if one uses the past performance of a sample of hedge funds to make an investment decision, one may expect volatility in its future performance. The study also suggests that poor performance is likely to be the cause of a fund's closure. This is because survival bias in ex post observed data is likely a result of high-watermark provisions.

Despite the findings mentioned above, it is important to note that data of disappearing funds was not collected by most hedge fund databases until 1994. Therefore, the attrition rate for hedge funds prior to 1994 should be near zero. In this context, results of survivorship bias for hedge funds prior to 1994 produced in past studies are somewhat under-represented. In other words, a more reasonable estimation of survivorship bias for hedge funds should be expected from studies employing data beyond 1994 (Liang, 2000).

Liang's study (1999) is considered the earliest study that tests survivorship bias for hedge funds after 1994. Using Malkiel (1995)'s method, in which a full sample of hedge funds (1,162) is compared with a sample of surviving funds (1,054), Liang (1999) examines survivorship bias for HFR data as of July 1997 and finds that survivorship bias is estimated to be around 0.84% per year. This low estimate of survivorship bias for hedge funds results from a lack of information concerning disappearing funds that were in existence before HFR started its data collection in 1993. For this reason, this study is expanded in Liang (2000), in which one additional database - TASS - is employed. By July 1998, TASS consisted of 1,627 hedge funds, among which 1,201 were surviving funds. As compared to HFR, TASS covers more defunct funds. In addition, TASS also has more funds that reveal their assets, fees, and minimum investment as compared to HFR. Results show that survivorship bias for hedge funds collected from HFR is relatively low (0.6% per year) for the period of 1994-1997. However, when using the TASS database, this bias is found to be much higher (2.24% per year), which is somewhat consistent with the results (3% per year) found in Brown et al. (1999) for offshore hedge funds. The lower bias estimate found in Liang (2000) reflects the lower risk of their sample consisting of both offshore and onshore funds as compared to that of the sample with only offshore hedge funds employed in Brown et al. (1999). The reason for the higher bias estimated for TASS data is that TASS has more comprehensive data of disappearing funds as compared to HFR. This implies that the accuracy of estimation for survivorship bias depends heavily on how well a database collects its data.

In addition to the extensive findings on the survivorship bias for hedge funds from two databases (HFR and TASS), Liang (2000) seems to be the first study that contributes to the area of survivorship bias for different hedge fund styles. Using data for the sample period of 1994-1997, Liang (2000) studies survivorship bias for hedge fund styles throughout their entire existence. Results show insignificant survivorship bias (0.31% per year) for the 17 hedge fund styles defined by HFR. However, there is more significant evidence (10 out of 15) collected for the 15 hedge fund styles defined by TASS on survivorship bias (1.49% per year) for hedge funds. Results found for survivorship bias according to styles are slightly different from those previously obtained for all funds because of the difference in the sample periods used. Again, TASS proves itself as a better source of data for hedge fund studies.

In an extended performance study for hedge funds, Liang (2001) examines the survivorship bias for a larger sample consisting of 1,921 hedge funds collected from the TASS database as of July 1999. Results show that the annual survivorship bias for the sample of hedge funds is 2.43% per year, which is very close to what was found in Liang (2000). The significant positive survivorship bias found in the study explains the fact that hedge funds disappear from the industry mostly because of their poor performance.

By adding more years of data to that used in Fung and Hsieh (1997), Fung and Hsieh (2000) examine survivorship for both individual hedge funds and funds of hedge funds (FOFs) between 1994 and 1998. Using Malkiel (1995)'s method, in which survivorship bias is measured as the difference between a set of surviving funds and a set of all funds, Fung and Hsieh (2000) analyse survivorship bias for individual hedge funds as well as for FOFs selected from the TASS hedge fund database for the period of 1994-1998. Their sample contains 1,722 funds, of which 1,120 are surviving funds. Due to the lack of information on assets under management (AUM) for individual hedge funds as well as FOFs, an equally- weighted portfolio is used as a proxy for performance measurement instead of a value-weighted portfolio for individual hedge funds and FOFs present in TASS. The test results show that the difference in the average returns between a full sample of individual hedge funds (10.2%) and a sample of only surviving hedge funds (13.2%) is estimated to be about 3% per year. For FOFs, survivorship bias seems to be lower than half (1.4%) of that of individual hedge funds per year. When measuring the selection bias for FOFs, Fung and Hsieh (2000) find that selection bias is relatively small for their sample of FOFs, implying that a FOF is more willing to disclose its track record in order to attract more capital. This may be because a FOF normally invests in a diversified portfolio of hedge funds, and thus does not have much constraint imposed on it. In short, survivorship bias as well as selection bias is found to be smaller for FOFs than for individual hedge funds in the study. Fung and Hsieh (2002) recommend a solution that could mitigate survivorship and self-selection biases through a construction of hedge fund benchmarks that is funds of funds (FOFs). Their empirical results confirm that both biases can be avoided in FOFs.

Using the same database (MAR) employed in Ackermann et al. (1999), Edwards and Caglayan (2001) attempt to estimate survivorship bias in a sample of 1,665 hedge funds as of August 1998. By comparing alphas (excess returns) between a sample of surviving funds (1,169) and a sample of all funds, Edwards and Caglayan (2001) find that survivorship bias for hedge funds in their sample is 1.85% per year. This finding is quite similar to what is found in Liang (2000) and Fung and Hsieh (2000).

Unlike other studies that use a definition of survivorship bias made either by Ackermann et al. (1999) or Liang (2000 and 2001), Capocci and Hübner (2004) choose to use both in order to examine survivorship bias for hedge funds within the period of 1984 - 2000 using the MAR and HFR databases. The authors divide the study period into two sub-periods: 1984-1993 and 1994-2000. According the definition of survivorship bias made by Ackermann et al. (1999), results show that survivorship bias for the full sample period is 0.36% per month (4.45% per year). However, a lower survivorship bias (0.07% per month or 0.9% per year) is found for the same period if Liang (2000)'s definition of survivorship bias is used. When analysing results produced for the two sub-periods, a lower survivorship bias is found for the period prior to 1994 compared to that found for the period after 1994. This evidence may reflect the fact that MAR and HFR only started collecting information on disappearing funds after 1994 and 1993 respectively. Overall, results found in the study show a higher bias for two periods, i.e. 1984-2000 and 1994-2000, and a similar bias for the period of 1984-1993, as compared to those obtained in Ackermann et al. (1999). However, when Liang (2000)'s definition is used, survivorship bias for the period of 1994-2000 is found to be lower than that (0.3% per monthly and 3% per year) found in Fung and Hsieh (2000) and Liang (2001) respectively. These differences in results may be due to the differences in the sample periods used in those studies.

For the TASS database, the issue of survivorship bias is further tested in Malkiel and Saha (2005) with a study period ending as of April 2004. When comparing the average annual performances of two samples, i.e. the sample of surviving funds and the sample of defunct funds, during the period of 1996 - 2003, a significant positive bias of 8.35% per year is found at the 5% level. The authors suggest that the performance of a sample of all hedge funds, including both defunct and live funds, is the best representation of the performance of the hedge fund industry. However, different results are presented in Ammann and Moerth (2008), in which a suvivorship bias of only 3.54% per year is found for the the TASS database for the period of January 1994 - June 2005. The difference in the above results may be due to the difference in the sample periods chosen for the two studies.

Capocci and Hübner (2004)'s method for testing survivorship bias was adopted in another study conducted on Italian hedge funds by Steri, Giorgino, and Viviani (2009). The estimated bias for Italian hedge funds is negative (-0.64%). This interesting finding implies that non-surviving Italian hedge funds have a higher average return compared to that of surviving funds. In other words, poor performance should not be a reason for Italian hedge funds to exit a database. However, results may be influenced by the small number of non-surviving Italian hedge funds in the sample. Furthermore, due to unobtainable information on self-selected funds, the authors assume that self-selection bias is zero or does not exist in the study.

Survivorship bias results similar to those found in Ackermann et al (1999) and Liang (2000) are also found in Eling (2009) for a sample of 2,936 hedge funds collected from Center for International Securities and Derivatives markets (CISDM) during the period of January 1996 - December 2005. Specifically, Eling (2009) finds that the estimates of suvivorship bias for hedge funds in the sample are small, i.e. 0.08% per month. Similar findings are also present in Capocci (2009) for the period of 1995 - 2002, from which survivorship biases of 1.49% per year are found for the sample of 3,060 hedge funds from both MAR/CISDM and Barclay databases.

The issue of survivorship bias is also addressed and examined in Aggarwal and Jorion (2010) for the TASS database for two study periods, i.e. (1) from 1994 to 2001, and (2) from 2002 to 2008. Using equally-weighted monthly returns which are then compounded to produce annual returns, a survivorship bias of 5.43% per year is found to be statistically significant for the sample consisting of surviving funds during the first study period. However, there is no evidence that survivorship bias exists during the second study period. Moreover, in the most recent study concerning this issue, Ibbotson, Cheng, and Zhu (2011) updates the survivorship bias that was estimated in Brown et al. (1999) using an extended sample period from 1995 untill December 2009 and a larger sample of TASS hedge funds consisting of both dead (3,917) and surviving (2,252) funds. As the authors compare the performance of the equally weighted porfolio of surviving funds and that of both dead and surviving funds, a survivorship bias of 3.16% is found. This result is somewhat similar to the previous finding in Brown et al. (1999), in which survivorship bias was 3% per year. However, when excluding the backfilled data for both above-mentioned portfolios, the survivorship bias becomes larger, i.e. 5.13% per year, which is consistent with the finding found in Aggarwal and Jorion (2010).

In short, the issue of survivorship bias has become more challenging for researchers as databases have grown in terms of quantity and quality as mentioned in Fung and Hsieh (2009). According to Fung and Hsieh (2009), based on the number of funds collected from four different hedge fund databases, i.e. Barclays, CISDM, HFR, and TASS, it is observed as of 2007 that 49% of unique funds came from a single database, while 31% of these funds appeared in two databases, 13% appeared in three databases, and 7% appeared in four databases. As hedge funds may appear in more than one database simultaneously, challenges arise in situations when funds may be reported as dead funds in one database, but as surviving funds in another database. Given the fact that it is impossible for researchers to access all hedge fund databases available in the world due to their financial constraints, capturing the exact survivorship bias may still be unrealistic. Therefore, the best way to bring the estimated survivorship bias of a database closer to its real value if possible, is to choose a database that comprehensively reports a specific population of hedge funds, if possible.

2.2 Multi-period Sampling Bias

Many studies impose the condition that funds selected for a study sample must have full monthly returns for a number of periods, i.e. 12 months, 24 months, or 36 months. This required condition is to ensure the minimal calculation needed to have meaningful measures (Eling, 2009). However, only a few studies examine the issue of multi-period sampling bias, namely Akermann et al. (1999), Fung and Hsieh (2000), and Edwards and Caglayan (2001).

In Ackermann et al. (1999), a fund that does not possess a full 24 months' worth of monthly data for the period of 1994 - 1995 may have started reporting only after 1994, or may not have survived until 1995. Multi-period sampling bias is measured as the difference in performance between two samples of funds, of which one is with less than and the other is with the full data set during 1994 - 1995. The bias is estimated to be about 12% per month, implying that a fund with less than two years of history performs better than a fund with a longer history. In other words, using a sample of funds with a full data set of 24 months will not exaggerate the overall performance of hedge funds.

Being stricter than the previous studies mentioned above, Fung and Hsieh (2000) impose a minimum history of 36 months for a hedge fund which is to be included in their sample. The estimation of multi-period sampling bias for hedge funds during the period of 1994 - 1998 is only 0.6% per month, reflecting a very small amount of bias. This suggests that the sample of Fung and Hsieh (2000) is a good representation of the hedge fund industry during the study period.

In a different approach, Edwards and Caglayan (2001) believe that requiring hedge funds to have monthly returns for a long period of time, i.e. 36 months (Fung and Hsieh, 2000) may lead to an excessive exclusion of non-surviving funds from a sample, thus resulting in a survivorship bias. The authors impose a requirement of only 24 monthly returns for all funds in their study sample. This was done after excluding the first 12 months of returns for all hedge funds in the sample to avoid potential backfilling bias. The study reveals an annual multi-period sampling bias of 0.32% for the selected sample of hedge funds. The authors also find a similar multi-period sampling bias of 0.29% when a requirement of only 12 monthly returns is imposed on all funds in the sample. The study suggests that the multi-period sampling bias is not significantly different when a requirement of either 24 or 12 monthly returns is imposed.

The past studies clearly revealed the importance to address the conditioning bias in a sample of hedge funds before proceeding to any further analysis. Due to the different ways of keeping data by various researchers, the differences in databases may also generate different type of data biases which need to be explored by researchers. In most of the studies conducted on AHFs, for instance, Kok, Koh, and Teo (2003), Do, Faff, and Wickramanayake (2005), Hakamada, Takahashi, and Yamamoto (2007), Fang, Phoon, and Xiang (2008), and Wong, Phoon, and Lean (2008), the issue of data bias has not been investigated. Thus, this study attempts to fill in this gap.

3. Research Methodology

We now evaluate data reliability of AHFs through the investigation on types of conditioning bias that might exist in our sample collected from the Eurekahedge database, which in turns might have an impact on the performance of AHFs. The total number of funds available in the Eurekahedge database by 15th June 2008 is 1428, including surviving (1032) and dissolved (396) funds.

We form four different samples (A, B, C, and D) in order to test the conditioning bias that might exist in the Eurekahedge database as well as in our sample of AHFs for the period of 1st January 2000 to 15th June 2008 (see *Figure 1*). Sample A represents all AHFs in the industry, Sample B represents AHFs that were present in the Eurekahedge database, Sample C represents AHFs that survive throughout the sample period, and Sample D represents AHFs that have had at least 4 years of performance as of 15th June 2008. Descriptive statistics for the four constructed samples (A, B, C, and D) described in *Table 1* below.

Figure 1: Samples for Testing Conditioning Biases

In this chart, four different constructed samples used to test conditioning biases are shown. Sample A includes all surviving funds, disappearing funds, and omitted funds. Sample B consists of all surviving funds and disappearing funds. Sample C includes all surviving funds. Sample D includes only surviving funds that have at least 4 years of full monthly returns.

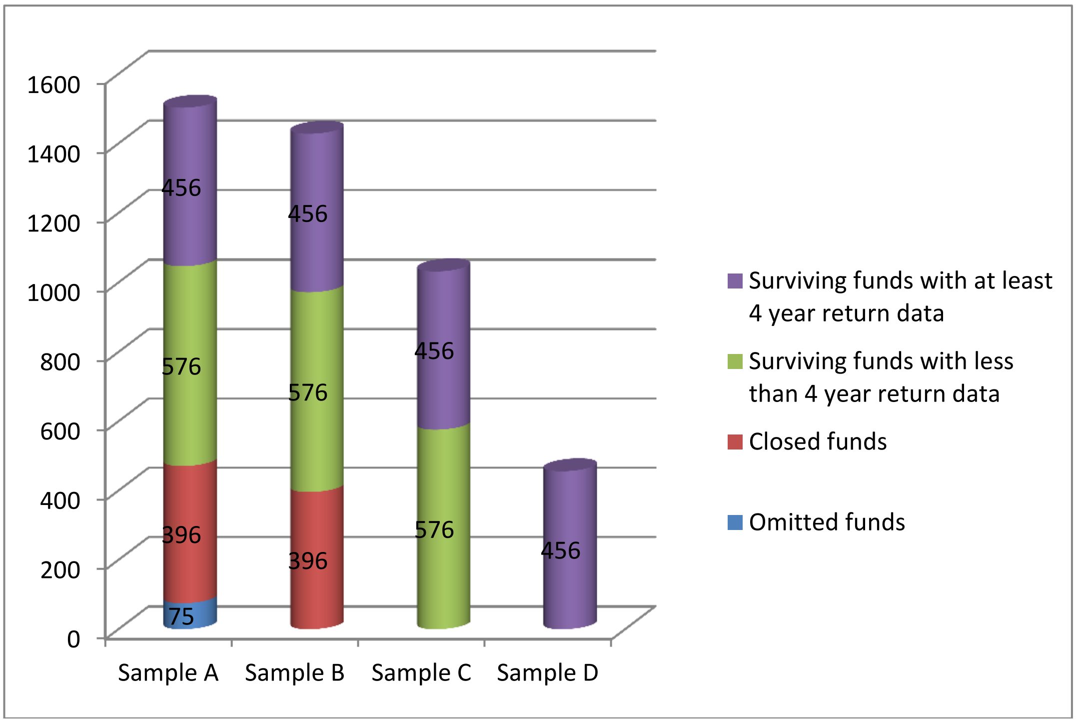

Table 1: Descriptive Statistics of the Four Constructed Samples of AHFs

In this table, average monthly returns, standard deviations, Sharpe ratios for the four constructed samples of AHFs are reported.

Samples	Average Monthly Returns	Standard Deviations	Sharpe Ratios
Sample B	0.83	4.36	-0.79
Sample A	0.83	4.33	-0.79
Sample C	0.91	4.42	-0.76
Sample D	0.89	4.37	-0.78

Eurekahedge Inc. normally keeps all historical information of dead funds in its database unless fund managers of those dead funds made a request to remove them totally from the database. We identified 75 funds that were omitted from the database for this reason between 1st January 2000 and 15th June 2008. We then compare the performance of samples with (sample A) and without (sample B) these omitted funds. The omission of closed funds may cause upward bias in the overall performance of AHFs during the study period. To our best knowledge, this kind of bias has not been studied before, which may be due to the non-reporting on these cases by the data vendors. Thus, we give a new name for this kind of bias as “omission bias”.

To test the difference between two population means, we form the following null and alternative hypotheses for a two-tailed test.

Ho: $\mu_1 - \mu_2 = 0$
Ha: $\mu_1 - \mu_2$ $\mu \neq 0$

Where: μ_1 and μ_2 denote the means for populations 1 and 2, respectively.
Our test statistics is as follows (Ott & Mendenhall, 1990)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$\bar{x}_1 \; and \; \bar{x}_2$ are two samples’ means. S_p is the common population standard deviation and is calculated as

$$s_p = \sqrt{\frac{(n_1 - 1){s_1}^2 + (n_2 - 1){s_2}^2}{n_1 + n_2 - 2}}$$

n_1 and n_2 are the two sample sizes. s_1 and s_2 are the two samples’ standard deviations. The degree of freedom for the above t-test is

$$df = n_1 + n_2 - 2$$

If t-statistics does not fall in the rejection region, we conclude that there is insufficient evidence to indicate a difference in the means of two samples. We exclude outliers from the samples before carrying out our tests.

4. Findings

When the average returns for samples A and B are compared, there is no apparent difference found between the two. This is confirmed by the t-statistics (-0.986) shown in *Table 2*, which suggests that there is no statistically significant omission bias in the Eurekahedge database at the 5% level of significance. In other words, the set of AHFs presented in the Eurekahedge database as of 15th June 2008 can be used to represent the population of the AHF industry. Thus, sample B can be used to measure two data bias for AHFs, i.e. survivorship bias and multi-period sampling bias.

Based on the average returns for Sample B and Sample C as shown in *Table 1*, the estimated survivorship bias is (-0.08%) per month or (-0.96%) per year. The higher average return of Sample C as compared to that of Sample B implies an upward bias in the sample of surviving AHFs, i.e. 0.08% per month or 0.96% per year. The survivorship bias found for the sample of AHFs in this study is therefore similar to what is found in Eling (2009) for hedge fund data collected from CISDM during the period of 1996 to December 2005. However, the two-tailed t-test results shown in *Table 2* suggest that this upward bias is statistically insignificant. In other words, survivorship bias in Sample C is not significant, thus there is no significant survivorship bias presented in the selected sample of study. This finding may confirm the claim in Koh et al. (2003) that the survivorship bias for a sample of 417 AHFs in their study is mitigated.

As shown in *Table 1*, the average return for Sample D is 0.89% per month, while the average return for Sample B is 0.83% per month. Thus, multi-period sampling bias for Sample D is about 0.06% per month or 0.72% per year. This finding may imply that AHFs that survive for at least 4 years is because of their good performances. However, the two-tailed t-statistic (0.038) shown in *Table 2* suggests that the multi-period sampling bias is not statistically significant at the 5% level.

When the average returns of Sample C (0.91%) and Sample D (0.89%) are compared, there is a slight downward bias in the performance of Sample D. This bias implies that a fund which survives for a longer period of time (4 years) does not perform better than a fund with a shorter history (less than 4 years). This result is directly opposed to the finding obtained in Brown, Goetzmann, and Park (2001) which states that funds with shorter histories are likely to perform poorly. However, the p-value for t-statistic shown in *Table 2* confirms that this difference is not statistically significant at the 5% level. In short, Sample D which consists of 456 AHFs is a good representative of the overall AHFs and can be used for further analysis.

Table 2: Results of Data Bias for the Four Selected Samples

In the following tables, we report data bias in monthly returns between the following paired samples: (1) sample of all funds including omitted funds by the data vendor (Sample A) and sample of all funds excluding omitted funds (Sample B), (2) sample that contains both surviving and disappearing funds (Sample B) and sample of all surviving funds (Sample C), (3) sample that contains both surviving and disappearing funds (Sample B) and sample of surviving funds with at least 4 year monthly return data (Sample D), (4) sample of all surviving funds (Sample C) and sample of surviving funds with at least 4 year monthly return data (Sample D).

No.	Samples	Difference	t-statistics	p-value
1	Sample A & Sample B	-0.0077	-0.9860	0.3260
2	Sample B & Sample C	-0.0284	-0.1140	0.9090
3	Sample B & Sample D	-0.0095	0.0380	0.9700
4	Sample C & Sample D	0.0200	-1.6900	0.0900

5. Conclusions

The study on the data bias of AHFs is important in ensuring that the sample chosen for this study can represent the entire AHF industry during the sample period. The two-sample t-test was employed to test three selected types of data bias, i.e. omission bias, survivorship bias, and multi-period sampling bias. Despite the fact that some of the AHFs were omitted from the Eurekahedge database as requested by their fund managers, results showed no significant evidence of omission bias present in the selected sample of 456 AHFs. Although survivorship bias was found to be around 0.0285 per month for the sample of surviving AHFs, it was not significant at the 5% level. The findings for multi-period sampling bias in the sample of AHFs were contrary to the common belief that funds which stay in the industry for a long period of time because of their good performance often perform better than funds with short histories of performance (Brown, Goetzmann, and Park, 2001). In fact, there was a downward multi-period sampling bias found in the sample of AHFs with four years of full monthly return data, although it was not statistically significant at the 5% level.

In short, the overall findings suggest the sample of 456 AHFs was free from the three types of above-mentioned bias. In other words, the sample can be used for further tests, and results obtained from this sample can be generalized for the AHF industry as of 15th June 2008, the end period of this study. For the first time, the three above-mentioned types of data bias were tested for a large sample of AHFs over a long study period of 1st January 2000 – 15th June 2008. Thus, the confirmation on the data-bias-free sample of AHFs in this study is believed to be helpful to researchers who work with AHF data as of 15th June 2008.

We acknowledge that this study has a limitation. The main AHFs database used for this study is the Eurekahedge database. Thus, if it is possible, the use of more than one database may allow a larger population of AHFs to be captured and studied so that stronger evidence may be produced.

References

Ackermann, C, McEnally, R, & Revenscraft, D 1999, 'The Performance of Hedge Funds: Risk, Return, and Incentives',*The Journal of Finance,*vol. LIV, no.3, pp.833-874.

Aggarwal, RK, & Jorion, P 2010, 'Hidden Survivorship in Hedge Fund Returns', *Financial Analysts Journal, vol.66, no.2*, pp.1-6 (Ahead of Print Version).

Ammann, M, & Moerth, P 2008b, 'Performance of Funds of Hedge Funds', *The Journal of Wealth Management, Summer* , pp.46-63.

Ball, R, & Watts, R 1979, 'Some Additional Evidence on Survival Biases', *The Journal of Finance* , vol. XXXIV, no.1, pp.197-206.

Baquero, G, Horst, JT, & Verbeek, M 2005, 'Survival, Look-Ahead Bias, and Persistence in Hedge Fund Performance', *Journal of Financial and Quantitative Analysis* , vol.40, no.3, pp.493-517.

Blake, C, Elton, E, & Gruber, M 1993, 'The Performance of Bond Mutual Funds', *Journal of Business, vol. 66,* pp.371- 403. .

Brown, SJ, & Goetzmann, WN 1995, 'Performance Persistence', *Journal of Finance, vol. 50, no.2* , pp.679-698.

Brown, SJ, Goetzmann, WN, & Ibbotson, RG 1999, 'Offshore Hedge Funds: Survival and Performance 1989-1995', *Journal of Business, vol.72,* , pp.91-118 .

Brown, SJ, Goetzmann, WN, & Park, J 2001, 'Careers and Survival: Competition and Risk in the Hedge Fund and CTA Industry', *The Journal of Finance, vol.LVI, no.5* , pp.1869-1886.

Brown, SJ, Goetzmann, WN, Ibbotson, RG, & Ross, SA 1992, 'Survivorship Bias in Performance Studies', *The Review of Financial Studies, vol. 5, no.4* , pp.553-580. .

Bu, Q, & Lacey, N 2007, 'Exposing Survivorship Bias in Mutual Fund Data', *Journal of Business and Economics Studies, vol.13, no.1*, pp. 22-37.
Capocci, DP 2007, 'The Sustainability in Hedge Fund Performance: new insights', *Working paper, HEC-ULG Management School* , pp.1-21.
Cappocci, DP & Hübner, G 2004, 'Analysis of hedge fund performance', *Journal of Empirical Finance* , vol.11, no.1, pp.55-89.
Capocci, DP 2009, 'Persistence in Hedge Fund Performance: Extended Analysis', *International Journal of Finance & Economics, vol.14* , pp.233-255.
Carhart, MM 1997, 'On Persistence in Mutual Fund Performance', *Journal of Finance, vol. 52, no.1* , pp.57-82.
Carpenter, JN & Lynch, AW 1999, 'Survivorship Bias and Attribtion Effects in Measures of Performance Persistence', *Journal of Financial Economics, vol. 54* , pp.337-374.
Do, V, Faff, R, & Wickramanayake, J 2005, 'An empirical analysis of hedge fund performance: The case of Australian hedge funds industry', *Journal of Multinational Financial Management* , vol.*15*, pp.377-393.
Edwards, FR & Caglayan, MO 2001, 'Hedge Fund Performance and Manager Skill', *Journal of Future Markets, vol.21, no.11* , pp.1003-1028.
Eling, M 2009, 'Does Hedge Fund Performance Persist? Overview and New Empirical Evidence', *European Fianncial Management Journal, vol.15, no.2*, pp.362-401.
Elton EJ & Brown MJ 1996, 'Survivorship bias and Mutual Fund Performance', *The Review of Financial Studies, vol.9, no. 4* , pp.1097-1120.
Fang, V, Phoon, K F, & Xiang, VY 2008, 'A Heuristic Approach to Asian Hedge Fund Allocation', *The Journal of Wealth Management, vol.10, no.4* , pp.42-52.
Fung, W & Hsieh, DA 1997, 'Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds', *The Review of Financial Studies, vol.10, no.2,* , pp.275 - 302.
Fung, W & Hsieh, DA 2000, 'Performance Characteristics of Hedge funds and Commodity Funds: Natural vs. Spurious Biases', *Journal of Financial and Quantitative Analysis, vol.35, no.3* , pp.291-307.
Fung, W & Hsieh, DA 2002, 'Hedge-Fund Benchmarks: Information Content and Biases', *Financial Analysts Journal, vol.58, no.1* , pp.22-34.
Grinblatt, MA 1989, 'Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings', *Journal of Business, vol.62, no.3* , pp.393-416.
Hakamada, T, Takahashi, A, & Yamamoto, K 2007, 'Selection and Performance Analsysis of Asia-Pacific Hedge Funds', *The Journal of Alternative Investments, Winter* , pp.7-29.
Ibbotson, RG, Cheng, P, CFA, & Zhu, KX 2011, 'The ABCs of Hedge Funds: Alphas, Betas, and Costs', *Financial Analysts Journal, vo.67, no.1* , pp.15-25.
Kok, F, Koh, WTH, &Teo, M 2003, '*Asian Hedge Funds: Return Persistence, Style, and Fund Characteristics', Working Paper.* Singapore: Singapore Management University.
Liang, B 1999, 'On the Performance of Hedge Funds', *Financial Analyst Journal, vol.4* , pp.72-85.
Liang, B 2000, 'Hedge Funds: The Living and the Dead', *Journal of Financial and Quantitative Analysis* , vol.35, no.3, pp.309-326.
Liang, B 2001, 'Hedge Fund Performance: 1990-1999', *Financial Analyst Journal, January/February* , pp.11-18.
Lyman 1999, '*Understanding Statistics',* Boston, USA: PWS-KENT Publishing Company.

Malkiel, BG 1995, 'Returns from investing in equity mutual funds 1971 to 1991', *Journal of Finance, vol.50* , pp.549-572.

Malkiel, BG, & Saha, A 2005, 'Hedge Funds: Risk and Return, *Financial Analysts', Journal, vol.61, no.6*, pp.80-88.

Ott, L, & Mendenhall, W 1990, '*Understanding Statistics' (5th ed.),* Boston, USA: PWS-KENT Publishing Company.

Sheridan, J 2007, '*SPSS 14.0 for Window - Analysis Without Anguish',* Australia: John Wiley and Sons Australia Ltd.

Steri, R, Giorgino, M, & Viviani, D 2009, 'The Italian Hedge Fund Indusry: An Empirical Analysis of Performance and Persistence', *Journal of Multinational Financial Management* , vol.19, pp.75-91.

Wong, WK, Phoon, KF, & Lean, HH 2008, 'Stochastic Dominance Analysis of Asian Hedge Funds', *Pacific-Basin Finance Journal, vol.16*