NBER WORKING PAPER SERIES

DO HOT HANDS EXIST AMONG HEDGE FUND MANAGERS? AN EMPIRICAL EVALUATION

Ravi Jagannathan
Alexey Malakhov
Dmitry Novikov

Working Paper 12015
http://www.nber.org/papers/w12015

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2006

We would like to thank Dmitri Alexeev, Torben Andersen, Duke Bristow, Art Bushonville, Bhagwan Chowdhry, Randolph Cohen, Kent Daniel, Harry DeAngelo, Dobrislav Dobrev, Shingo Goto, David Hsieh, Ken Judd, Robert Korajczyk, Ken Kroner, Ananth Madhavan, Rosa Matzkin, Narayan Naik, Todd Pulvino, Che-Lin Su, Melvyn Teo, Sergey Tsyplakov, Motohiro Yogo, participants of the 2006 NBER Asset Pricing Program Meeting, 2006 Western Finance Association Meeting, 2007 American Finance Association Meeting and seminar participants at Barclays Global Investors, China Europe International Business School, East Carolina University, Indian School of Business, National University of Singapore, Northwestern University, Singapore Management University, University of Arkansas, University of Illinois at Urbana-Champaign, University of North Carolina at Chapel Hill, University of South Carolina, University of Southern California, York University for their helpful discussions and suggestions. We are grateful to HFR Asset Management for providing us with the data. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Do Hot Hands Exist Among Hedge Fund Managers? An Empirical Evaluation
Ravi Jagannathan, Alexey Malakhov, and Dmitry Novikov
NBER Working Paper No. 12015
February 2006, Revised July 2007
JEL No. I1

ABSTRACT

We examine whether hot hands exist among hedge fund managers. In measuring performance, we use hedge fund style benchmarks. This allows us to control for optionlike features inherent in returns from hedge fund strategies. We take into account the possibility that reported asset values may be based on stale prices. We develop a statistical model that relates a hedge fund?s performance to its decision to liquidate or close in order to infer the performance of a hedge fund that left the database. While we find significant performance persistence among superior funds we find little evidence of persistence among inferior funds.

Ravi Jagannathan
J.L. Kellogg Graduate School of Management
2001 Sheridan Road
Leverone/Anderson Complex
Evanston, IL 60208-2001
and NBER
rjaganna@northwestern.edu

Alexey Malakhov
College of Business
University of Arkansas
malakhov@walton.uark.edu

Dmitry Novikov
Goldman, Sachs & Co.
dmitry.novikov@gs.com

1 Introduction

"Investors have made a trillion-dollar bet that hedge funds will bring them rich returns" claims an article in The Economist. Indeed, the hedge fund industry grew at an astounding pace from 610 funds controlling $39 billion in 1990 to more than 9,000 funds with $1.3 trillion in 2006.[1] While it appears that investors have enthusiastically embraced hedge funds as an investment vehicle, and are especially eager to invest in hedge funds that have exhibited outstanding past returns, there is little consensus in the empirical finance literature on performance evaluation, and whether there is performance persistence among hedge funds. *"If you are thinking about investing in a hedge fund, you probably want to figure out how it stacks up against the competition. Good luck"* - was a recent comment in The Wall Street Journal on the subject.[2] In part that is due to the fact that any rigorous research about hedge fund performance has to overcome numerous biases and irregularities in the available data. These biases arise due to the unregulated nature of the hedge fund industry. There are no legal requirements for hedge funds to report performance numbers, although there are several different databases, to which hedge funds provide information about themselves on a voluntarily basis.[3] Ackerman, McEnally, and Ravenscraft (1999), Liang (2000), Fung and Hsieh (2000) and Fung and Hsieh (2002) discuss the issues that arise when using data from these sources.

In this paper we study performance persistence among hedge fund managers, while correcting for measurement errors as well as for the backfill, serial correlation, and look-ahead biases in the data. We introduce a relative performance measure, alpha, for hedge fund managers. It reflects the performance of a hedge fund manger relative to the market and the group of "peers", i.e. hedge funds pursuing similar strategies. We find relative performance persistence over a three year horizon, i.e. that managers with higher estimated alphas in one three year period tend to have higher estimated alphas in the following three year period.

An important feature of a hedge fund database is backfill bias - the case when hedge funds bring their history with them when they join a database. Since only funds with relatively superior historical performance enter a database, when possible backfilling of data is ignored, it results in a bias toward mistakenly assigning superior ability to managers of funds in their earlier years. Since our HFR data contains the information on when funds actually joined the database, we are able to eliminate the backfill bias by deleting all the

[1] See "The New Money Men," The Economist, Feb 17, 2005, and "Rolling in It," The Economist, Nov 16, 2006.

[2] See "Race to Rate Hedge Funds Begins in Heavy Fog" by Scott Patterson, The Wall Street Journal, Sep 28, 2005

[3] Among them are MAR, TASS and HFR (we use the HFR database in the paper).

backfill observations in our data set. Moreover, our data is survivorship bias free, since the HFR database retains all hedge funds, including those that ceased to exist.

Another issue with hedge fund analysis is that hedge fund returns exhibit substantial serial correlation, a feature that is extensively investigated in Getmansky, Lo, and Makarov (2004) and Okunev and White (2003). They showed that the presence of illiquid assets in hedge fund portfolios are the primary source for the serial correlation. If serial correlation is not accounted for properly, the manager's performance measure will be biased. Notice that when hedge fund returns exhibit serial correlation due to the presence of illiquid assets in the portfolio, benchmark style index factor returns will also exhibit such serial correlation. We assume that unobserved "true" returns on assets are serially uncorrelated, and identify them using the MA2 approach suggested by Getmansky, Lo, and Makarov (2004). We measure performance relative to a carefully chosen portfolio of fund specific style index benchmarks and a broad stock market index, i.e., we use alpha relative to peers. To the extent peers within each hedge fund style take similar risks, we are able to control for option-like features in returns.

We evaluate hedge fund performance persistence by comparing the alphas over consecutive nonoverlapping three year intervals. This is a fairly long time period relative to the time periods examined in the literature reviewed in the following section. Considering a three-year period allows us to accurately capture relative alphas for individual funds, and also provides us with a better sense of investor returns accounting for lockup, notice, and redemption periods. For example, an investor in a fund with a two year lockup period can realistically expect to receive her money from two years and three months to two years and six months later. Lockup periods vary among different funds, but periods exceeding two years are not uncommon, and they have gotten more prevalent in recent years.[4] Following Hsieh,[5] we employ a method of weighted least squares in order to minimize the downward bias in persistence caused by measurement errors in alphas. We assign more weight to more precisely measured alphas in our sample. We further apply this approach to study persistence among the best performing and the worst performing funds separately.

Finally, some hedge funds stop reporting to the database before the end of the sample period used in the study.[6] That may lead to a biased estimate of alpha-persistence when

[4] For example, in 1996, LTCM allowed to withdraw one third of investor's capital in years 2, 3, and 4 (Perold (1999)). The adoption of a new SEC rule in December 2004 provided further incentives for hedge funds to adopt lockup periods in excess of two years (the rule was struck down by the US Court of Appeals in June 2006).

[5] Mimeo, private communication.

[6] Notice that the fact of nonreporting to a database does not mean fund liquidation. For example, a fund may stop reporting after it has been closed for *new* investors. Such a hedge fund will continue to manage funds of current investors.

the likelihood of a fund leaving the database is related to its past and expected future performance. Therefore, estimating performance persistence by regressing future alpha on past alpha without addressing conditional nature of the observed distribution of alphas may produce a biased estimate of alpha persistence. We follow the terminology of Baquero, Ter Horst, and Verbeek (2005), and refer to it as a look-ahead bias. We simultaneously address measurement errors and the look-ahead bias by building a statistical model that assumes that hedge funds that are liquidated are more likely to be ones with low past performance and those that are closed are more likely to be ones with high past performance. Our statistical model provides additional information about the unobserved performance of funds thereby reducing the measurement error in estimated alphas, provided the model is right. We assume that hedge funds that stop reporting but do not give a reason are drawn from the same distribution as funds that continue to report or stop reporting but tell us why. With these assumptions, which we empirically show are reasonable, we develop a GMM estimation method that estimates all parameters in the model and produces an estimate of performance persistence. Our approach is also consistent with the observation in Brown, Goetzmann, and Park (2001) and Liang (2000) that hedge funds with low past performance are primary candidates for liquidation. Overall, both weighted least squares and GMM approaches produce similar estimates of performance persistence.

The unobserved performance of a hedge fund after it stopped reporting to the database can result in a biased persistence estimate. For example, a fund that has a large positive alpha during the first three year period may perform poorly during the second three year period and liquidate; a fund that has a large negative alpha during the first three year period may perform extremely well during the second three year period and close; and both funds will stop reporting their performance data. That could cause a positive bias in measured persistence in the alphas of funds that survived during both three year periods. While it is a possibility, we provide diagnostics indicating that it is not a likely scenario.

We find that relative performance tends to persist among hedge fund managers. The average of performance persistence parameter estimates is 23% from the weighted least squares approach,[7] and 26% from the GMM procedure.[8] This implies that a hedge fund that outperformed its benchmark by 100 basis points in the past will on average continue to outperform its benchmark by more than 20 basis points in the future. In comparison, a simple regression of future alphas on past alphas gives a downward biased average estimate of only 14% for alpha persistence. The weighted least squares approach also provides strong evidence of performance persistence among the top hedge funds with the average persistence

[7] Individual cross-section estimates for the weighted least squares approach vary from 9% to 38%.

[8] Individual cross-section estimates from the GMM procedure vary from 0.3% to 54%.

estimate of 45% among the top 33% of the funds, and of 72% among the top 10% of the funds. In contrast, there is no evidence of persistence among the bottom funds. Our findings are consistent with Berk and Green (2004) who show, using a rational model of active portfolio management, that in equilibrium more money will flow to managers with superior skills leading to an erosion of performance over time and equalization of after fee returns available to investors from managers with different levels of skills when there are diminishing returns to scale; however only part of the superior performance erodes.

The rest of this paper is organized as follows. The next section provides a connection to the existing hedge fund performance persistence literature. Section 3 describes the methodology for empirical testing. The model of hedge fund performance is introduced, factor selection, return smoothing and look-ahead bias issues are discussed there. Tests for performance persistence are also explained. Section 4 contains data description, along with estimation of hedge fund performance persistence. Section 5 concludes.

2 Related Literature

There are several papers in the literature that examine hedge fund managers' performance persistence. Brown, Goetzmann, and Ibbotson (1999) estimated the offshore hedge fund performance using raw returns, risk adjusted returns using the CAPM, and excess returns over self reported style benchmarks. They found little persistence in relative performance across managers. On the contrary, Agarwal and Naik (2000a) and Agarwal and Naik (2000b) when using both offshore and onshore hedge funds found significant quarterly persistence - that is hedge funds with relatively high returns in the current quarter tend to earn relatively high returns in the next quarter. They used the return on a hedge fund in excess of the average return earned by all funds that follow the same strategy as a measure of performance.[9] They used both parametric and nonparametric tests for performance persistence. In their case the persistence was driven mostly by "losers". Edwards and Caglayan (2001) considered an eight-factor model to evaluate hedge fund performance. They found the evidence of performance persistence over one and two year horizons. They also showed that the persistence holds among both "winners" and "losers".

More recently, Bares, Gibson, and Gyger (2003) applied a non-parametric approach to individual funds, as well as an eight-factor APT model to fund portfolios with a conclusion of performance persistence only over one to three month horizons. Capocci and Hübner (2004) followed the methodology of Carhart (1997), discovering no evidence of performance per-

[9] They also examined the standardized measure of performance, i.e., the excess return dividend by its standard deviation.

sistence for best and worst performing funds, but providing limited evidence of persistence for middle decile funds. Boyson and Cooper (2004) have found no evidence of performance persistence if only common risk and style factors are used in estimation, but discovered quarterly persistence when manager tenure was taken into consideration. Baquero, Ter Horst, and Verbeek (2005) concentrated on accounting for the look-ahead bias in evaluating hedge fund performance. Comparing raw and style-adjusted performance of performance-ranked portfolios they found evidence of positive persistence at the quarterly level. Kosowski, Naik, and Teo (2007) used a seven-factor model and applied a bootstrap procedure, as well as Bayesian measures to estimate hedge fund performance. Considering performance-ranked portfolios they found evidence of performance persistence over a one year horizon. Finally, Fung, Hsieh, Naik, and Ramadorai (2007), using data for fund of hedge funds, show that it is possible to identify fund of funds that deliver superior alphas. However, they find that new money flows faster into such funds leading to a deterioration of their performance over time.

This paper contributes to the above literature in three ways. First, control for the measurement errors in alphas using weighted least squares and GMM procedure. The latter deals with measurement errors and the look-ahead bias simultaneously. Second, to our knowledge, this paper is first to account for all three major biases in hedge fund data, i.e. backfill, serial correlation, and look-ahead biases. Third, we present evidence of hedge fund managers' performance persistence over longer (three year) horizons, especially among the top performing funds.

3 Econometric Methodology

In this section we describe the estimation of hedge fund performance and then we propose a method to check for performance persistence.

3.1 Modeling the Relative Performance of a Hedge Fund

Hedge fund returns have several distinctive features. This can make the analysis of hedge funds' performance different from the analysis of performance of other assets like stocks and mutual funds.

First, hedge funds are not required to reveal their financial information including their returns.[10] This raises a question about the selectivity of returns in hedge fund databases. We

[10] According to SEC regulation 13F institutional investors with assets under management more than $100M are supposed to reveal their long position holdings on quarterly basis.

should take into account possible reasons for a hedge fund to reveal its performance information. One possible explanation is that some hedge funds need to raise funds. Reporting their returns could be a way to advertise themselves. This implies that we will probably not find the most and the least successful hedge funds in the database. The most successful funds most likely have enough clients without any additional promotions. The least successful funds probably would not reveal their information to a broad set of investors.

Second, hedge fund strategies produce returns that cannot be well explained by standard factors,[11] and they also exhibit option-like features.[12] The usual way to estimate the performance in such a case is to include options on factors in addition to these factors, following the suggestion made by Glosten and Jagannathan (1994).

Third, hedge funds often hold illiquid securities in their portfolios. Usually, it is difficult to obtain current prices for such securities. In this case, managers use past prices to estimate the current value of assets. Therefore, we may observe serial correlation in returns. If we completely ignore this issue, then we will get inconsistent estimates of hedge fund performance. Scholes and Williams (1977) proposed a simple way to account for stale prices. They used lags of factors along with factors in estimating the asset performance. These lags control for the serial correlation in returns. Asness, Krail, and Liew (2001) using this technique showed that the performance of indices[13] may not be as attractive as it appears from a regular regression without including any lags. Lo (2002) showed that annualized Sharpe ratios can be significantly overstated if the serial correlation in returns is not taken into account. Getmansky, Lo, and Makarov (2004) and Okunev and White (2003) introduced models for hedge fund returns, taking into account stale prices and return smoothing practices among hedge funds. Getmansky, Lo, and Makarov (2004) also estimated smoothing patterns for individual hedge funds and indices.

Fourth, the history of hedge funds is relatively short. Even for long-livers the reliable data in most cases does not exceed ten years. This creates a problem in analyzing hedge fund risks. The hedge fund return history may simply be too short for a high risk (low probability) event to happen. Weisman (2002) explains several simple strategies[14] that can be successful for a relatively long period of time (several years), but finally lead to bankruptcy. Those strategies will not be correlated with systematic factors. Pastor and Stambaugh (2002b),

[11] See Fung and Hsieh (1997).

[12] See for example, Fung and Hsieh (1997), Fung and Hsieh (2001), Mitchell and Pulvino (2001), Okunev and White (2003), Agarwal and Naik (2004), and Bondarenko (2004) for the discussion of the issues that option-like features in managed portfolio returns create when measuring performance.

[13] In the case of Hedge Fund Research style indices.

[14] Consider for example a strategy from St. Petersburg Paradox. You place one dollar on a coin to be tossed heads. If you lose, then you double your bets (if you do not have your own capital then you have to borrow). If you play long enough, then with probability one you will face a borrowing constraint.

Pastor and Stambaugh (2002a), and Ben Dor, Jagannathan, and Meier (2003) developed techniques for dealing with short histories. Ben Dor, Jagannathan, and Meier (2003) used two stage regressions; Pastor and Stambaugh (2002b) and Pastor and Stambaugh (2002a) used Bayesian analysis. Kosowski, Naik, and Teo (2007) applied Bayesian technique to the hedge fund performance analysis.

Finally, the life of hedge funds can be pretty short. Hedge funds can be liquidated or closed for new investments. Even if a database is survivorship bias free (that is, it stores all the liquidated and closed funds), there is the issue of how these hedge funds should be taken into account when analyzing performance persistence.

While analyzing the performance of hedge funds and performance persistence, we will try to control for the above features of hedge fund returns. We follow Getmansky, Lo, and Makarov (2004) in designing an appropriate model for the estimation of hedge fund performance.

Let the true equilibrium (unobserved) returns follow:

$$R_{i,t}^{un} = \alpha_i + X_t\beta_i + \varepsilon_{i,t} \tag{1}$$

where X_t is the vector of returns on factor portfolios $(T \times l)$, ε_{it} are i.i.d. We define α_i as the performance of a hedge fund. We assume that the observed returns (as reported by the hedge fund managers) are smoothed. Hence we observe the following returns

$$R_{i,t} = \theta_0^i R_{i,t}^{un} + ... + \theta_s^i R_{i,t-s}^{un}$$

Note that s may be different for different hedge funds. For identification purposes we will use the following normalization on the parameters:

$$\theta_0^i + ... + \theta_s^i = 1 \text{ for any } i$$

Combining with equation (1) we can write the observed returns as follows:

$$R_{i,t} = \alpha_i + X_t\theta_0^i\beta_i + ... + X_{t-s}\theta_s^i\beta_i + u_{i,t} \tag{2}$$

where

$$u_{i,t} = \theta_0^i\varepsilon_{i,t} + ... + \theta_s^i\varepsilon_{i,t-s} \tag{3}$$

As we see from (3), the error term $u_{i,t}$ follows an $MA(s)$ process. The next step is to choose appropriate factors for the model given by (2) and (3).

3.2 Factor Selection

In selecting factors we use the following criteria:

1) The number of factors should be relatively small as we do not have a long time series of observations on hedge fund returns. This also avoids overparametrization.

2) Factors should reflect the non-linear (option-like) strategies used by hedge funds.

Given this, we choose the following three factors.

Variable	Description
R_t^{mkt}	Return on the market portfolio (CRSP)
$I_t^{J,self}$	Self reported style index J from HFR
$I_t^{K,aux}$	Additional style index K from HFR

Therefore, $X_t' = [R_t^{mkt}, I_t^{J,self}, I_t^{K,aux}]$. The first factor is the CRSP market portfolio, and the other two factors are HFR style indices. Style indices are defined as an equally weighted average of returns for all hedge funds with the same strategy. The hedge funds themselves provide information about strategies they use. The list of strategies[15] defined in the database can be found in table 1.

Style indices are good proxies for non-linear strategies of hedge funds, however there are problems with self reported styles. For all hedge funds in the database we can find the styles that were reported by the hedge funds themselves. However, hedge funds may change their styles over time, and this may not be reflected in the database. We observe only one style per hedge fund and we do not know if a hedge fund has been using this style lately or some time ago (it may depend on the willingness of a hedge fund to report any changes in its style). To account for this "unpleasant" feature, we are going to add one more style index[16] in addition to the self reported index to try to capture changes in hedge fund styles. This additional style index is chosen by the Schwarz's Bayesian criterion (SBC) (details are provided in the next subsection).

The second problem is with style indices as factors. We know that the reported hedge fund returns are smoothed. By definition, a style index is the (equally weighted) average of returns for all hedge funds with the same self-reported strategy. Therefore, we should expect style indices to display serial correlations (or be "smoothed") as well. To deal with this problem, we consider the following model of "smoothed" indices (again we follow here

[15] For the official definition of self reported index, please refer to the web page of Hedge Fund Research at http://www.hedgefundresearch.com/pdf/HFR_Strategy_Definitions.pdf.

[16] We also found little evidence that adding more than one additional style index improves the fit of the model.

Getmansky, Lo, and Makarov (2004)):

$$I_t^J = \gamma_0^J \eta_t^J + ... + \gamma_l^J \eta_{t-l}^J \tag{4}$$

where η_t^J represents the unobservable "true" factor J at time t. Let us assume that $\eta_t^J \sim N\left(\mu_J, \sigma_J^2\right)$. Equation (4) is a moving average process of order l. To identify this process, as before we assume $\gamma_0^J + ... + \gamma_l^J = 1$. From equation (4) we see that I_t^J follow an $MA(l)$. Hence, the true factors η_t^J can be estimated from (4) by maximum likelihood. For this estimation we set $l = 2$ (i.e. we assume that indices are smoothed up to two lags[17]). We will use η_t^J as factors in (2).

The autocorrelations of orders from 1 to 12 for the original database indices I_t^J are presented in figure 1. We can see that several indices have significant[18] first and second order autocorrelation. The examples of such strategies are "convertible arbitrage", "distressed securities", "emerging markets", etc. These strategies involve heavy trading in illiquid securities. Figure 2 displays the autocorrelations of orders from 1 to 12 for unsmoothed indices η_t^J. None of the unsmoothed indices η_t^J have statistically significant autocorrelations, and their autocorrelations are substantially smaller than corresponding autocorrelations in figure 1.

3.3 Estimation procedure

In order to check for performance persistence we have to have at least two periods with performance estimates, see figure 3. For every period, we run the following regression based on the model given by (2) and (3):

$$R_{i,t} = \alpha_{zi} + X_t \delta_{0,i} + ... + X_{t-s} \delta_{s,i} + u_{i,t} \tag{5}$$

$$u_{i,t} = \theta_0^i \varepsilon_{i,t} + ... + \theta_s^i \varepsilon_{i,t-s} \tag{6}$$

where z is either 0 or 1, depending on if $T \leq t < T + k$ or $T + k \leq t < T + 2k$; X_t is the vector of factors described in the previous subsection.

We estimate the alphas by Maximum Likelihood. We also take into account the fact that the error term $u_{i,t}$ follows moving average process of order s. As a result of the maximum likelihood estimation procedure, we obtain consistent and asymptotically efficient estimators.

For every hedge fund we have to determine how many lags s to include and which additional indices are to be used in (5). We use Schwarz's Bayesian Criterion (Schwarz

[17] Getmansky, Lo, and Makarov (2004) use two lags to estimate the smooth model of hedge fund returns.

[18] At the a 5% significance level.

(1978)) to select the best model:

$$SBC = -2\log(L) + l \times \log(n)$$

where L is the likelihood function, l is the number of parameters and n is the number of observations. Given a hedge fund, we estimate several models like (5) that will be different in the number of lags and additional style indices. We then pick the model with the highest value of the Schwarz's Bayesian Criterion. For different hedge funds we may have different number of lags[19] in regression (5) and different additional indices.[20]

We use primary and additional style indices as factors in estimation of hedge fund performance. Therefore, we look at the *relative* performance of hedge funds with respect to hedge funds that follow similar investment strategies.

3.4 Testing Hedge Fund Performance Persistence

Here we provide an econometric framework for testing a hypothesis of performance persistence.

3.4.1 Simple (Naive) Regressions

Suppose we have obtained the hedge fund alphas for two periods α_{0i} and α_{1i}. Then we can run a simple regression

$$\alpha_{1i} = a + b\alpha_{0i} + \varepsilon_i \tag{7}$$

The persistence would mean that the slope coefficient b is statistically different from zero. However, a statistically insignificant slope coefficient would not necessarily mean the absence of persistence. That is because the slope estimate can be biased toward zero due to measurement errors. We discuss the nature of this bias in the next subsection.

3.4.2 Measurement Errors and Estimation Bias

If the true alphas were known, then the regression (7) would have given us an unbiased estimate of performance persistence. However, in reality there is always a measurement error present in our estimates of alphas. Assume that we observe

$$\begin{aligned} \alpha_{0i} &= \alpha_{0i}^* + u_i \\ \alpha_{1i} &= \alpha_{1i}^* + v_i \end{aligned}$$

[19] We consider up to three lags for each hedge fund.

[20] We also consider a model without an additional style index.

where α^*_{0i} and α^*_{1i} are "true"measures of relative performance, and noise components u_i, v_i are independent from the "true" alphas and from each other.

The OLS slope estimator from the regression (7) is equal to

$$\hat{b}_{OLS} = \frac{cov\left(\alpha_{1i}, \alpha_{0i}\right)}{Var\left(\alpha_{0i}\right)} = \frac{cov\left(\alpha^*_{1i}, \alpha^*_{0i}\right)}{Var\left(\alpha^*_{0i}\right) + Var(u_i)} \tag{8}$$

It is easy to see from (8) that the error in measuring α_0 creates the downward bias in the naive OLS estimate $\hat{b}_{OLS}$ compared to the "true" persistence estimate $\hat{b}^*$, since

$$\left|\hat{b}_{OLS}\right| = \left|\frac{cov\left(\alpha^*_{1i}, \alpha^*_{0i}\right)}{Var\left(\alpha^*_{0i}\right) + Var(u_i)}\right| < \left|\frac{cov\left(\alpha^*_{1i}, \alpha^*_{0i}\right)}{Var\left(\alpha^*_{0i}\right)}\right| = \left|\hat{b}^*\right|$$

Further, notice that the error in measuring α_1 does not result in a biased estimate of persistence, and thus we assume without loss of generality that $\alpha_{1i} = \alpha^*_{1i}$ throughout the rest of the paper.

3.4.3 Weighted Least Squares Approach

We employ a method of weighted least squares in order to minimize the downward bias in persistence caused by measurement errors in alphas. Performing regression (7) in terms of the t-statistic of alpha would result in a more accurate estimate of persistence, since more accurately measured alphas would have higher absolute t-statistic values, while less accurately measured alphas would have lower absolute t-statistic values. Unfortunately, such regression results could be difficult to interpret as a measure of performance persistence, since the weights would be different across the evaluation and prediction periods.

We employ a stylized t-statistic of alpha that is obtained by dividing all alphas by their standard deviations during the evaluation period, i.e we consider

$$t^*_{\alpha_{1i}} = a + bt_{\alpha_{0i}} + \varepsilon_i, \tag{9}$$

where

$$t_{\alpha_{0i}} = \frac{\alpha_{0i}}{\sigma_{\alpha_0}}, \quad t^*_{\alpha_{1i}} = \frac{\alpha_{1i}}{\sigma_{\alpha_0}}.$$

This results in assigning more weight to more precisely measured alphas in our sample, and it also allows us to interpret the regression result as a measure of performance persistence. We further apply this approach to see whether performance persists among the best performing or the worst performing funds by running regression (9) for the upper and the lower terciles according to their alpha t-statistic during the evaluation period.

3.4.4 Selective Reporting Model

In this section we address the errors in variables problem and potential look-ahead bias by modeling the nature of the dependence of the closing/liquidation decision of a fund on its true "alpha". We estimate the model parameters using the generalized method of moments. While estimating alphas in the prediction period, one can notice that some hedge funds, which were available in the evaluation period, disappeared from the database. A hedge fund can be liquidated or closed.[21] Closed funds typically stop reporting to the database, since they do not need to attract any additional investments. In the HFR database, hedge funds that opt out of the database may indicate the reason (liquidated fund or closed for new investments fund). For some hedge funds this information is missing.

We build the following model. Suppose that the hedge fund performance is measured by alphas: α_{0i} - alpha in the evaluation period and α_{1i} - alpha in the prediction period. We can observe α_{0i} for all funds in our sample during the evaluation period, but we can only observe α_{1i} for funds that were not liquidated or closed during the prediction period. We can also observe whether a hedge fund was liquidated or closed for new investments. We model the above pattern in hedge funds' performance and reporting as follows:

$$\alpha_{1i}^* = a + b\alpha_{0i}^* + \varepsilon_i \tag{M}$$

$$\alpha_{0i} = \alpha_{0i}^* + u_i$$

$$\alpha_{1i} = \begin{cases} \textit{liquidated}, & \text{with probability } p_0\left(\alpha_{0i}^*\right) \\ \alpha_{1i}^*, & \text{with probability } p_2\left(\alpha_{0i}^*\right) \\ \textit{closed}, & \text{with probability } p_1\left(\alpha_{0i}^*\right) \end{cases}$$

where $p_0\left(\alpha_{0i}^*\right) + p_1\left(\alpha_{0i}^*\right) + p_2\left(\alpha_{0i}^*\right) = 1$.

This model implies that we observe noisy[22] variables of hedge fund performance, however the decision on hedge fund liquidation, or closing is based on the "true" α_{0i}^* measure of performance.

The noise in this model follows

$$\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$$
$$u_i \sim N(0, \sigma_u^2)$$

[21] A hedge fund is called closed if it is closed for new investors. It continues to manage capital of its current investors.

[22] The measurment error can be attributed for example to the incomplete set of factors in the performance estimation regression.

and these random variables are independent.

We assume that hedge fund alphas are normally distributed as well.

$$\alpha_{0i}^{*} \sim N(\mu_{\alpha}, \sigma_{\alpha^*}^2)$$

and

$$\alpha_{0i} \sim N(\mu_{\alpha}, \sigma_{\alpha}^2)$$

One can easily establish the relationship between the variance of α_{0i}^{*} and α_{0i} :

$$\sigma_{\alpha}^2 = \sigma_u^2 + \sigma_{\alpha^*}^2 \tag{10}$$

For notational convenience, we consider σ_{α^*} as an unknown parameter, which is to be estimated (instead of σ_u), then σ_u can be easily found from (10).

3.4.5 GMM Estimation

Consider the following specification for probabilities of liquidation and closure:

$$\begin{aligned} p_0\left(\alpha_{0i}^{*}\right) &= \begin{cases} \max\{\min\{g_0(\mu - \alpha_{0i}^{*}) + c_0, 1 - c_1\}, 0\}, \text{ if } \alpha_{0i}^{*} \leq \mu_{\alpha^*} \\ c_0, \text{ if } \alpha_{0i}^{*} > \mu_{\alpha^*} \end{cases} \\ p_1\left(\alpha_{0i}^{*}\right) &= \begin{cases} c_1, \text{ if } \alpha_{0i}^{*} \leq \mu_{\alpha^*} \\ \max\{\min\{g_1(\alpha_{0i}^{*} - \mu) + c_1, 1 - c_0\}, 0\}, \text{ if } \alpha_{0i}^{*} > \mu_{\alpha^*} \end{cases} \end{aligned} \tag{P}$$

where μ_{α^*} is the mean value of α_0^*. Then model (M) with specification (P) has nine parameters: a, b, c_0, c_1, g_0, g_1, σ_{ε}, σ_{α^*}, and μ_{α^*}. Of these parameters, μ_{α^*} is obviously identified, and it is estimated by the sample mean of α_0. The remaining eight parameters $P = (a, b, c_0, c_1, g_0, g_1, \sigma_{\varepsilon}, \sigma_{\alpha^*})$ in model (M) with specification (P) are identified and can be estimated via GMM using the following moment conditions:

1) Conditional probability of liquidation, given $\alpha_{0i} \leq \mu_{\alpha^*}$:

$$\Pr(liquidation|\alpha_{0i} \leq \mu_{\alpha^*}) = \Pr(liquidation|\tilde{\alpha}_{0i} \leq \mu_{\alpha^*}) \tag{11}$$

2) Conditional probability of liquidation, given $\alpha_{0i} > \mu_{\alpha^*}$:

$$\Pr(liquidation|\alpha_{0i} > \mu_{\alpha^*}) = \Pr(liquidation|\tilde{\alpha}_{0i} > \mu_{\alpha^*}) \tag{12}$$

3) Conditional probability of closure, given $\alpha_{0i} \leq \mu_{\alpha^*}$:

$$\Pr(closure|\alpha_{0i} \leq \mu_{\alpha^*}) = \Pr(closure|\tilde{\alpha}_{0i} \leq \mu_{\alpha^*}) \tag{13}$$

4) Conditional probability of closure, given $\alpha_{0i} > \mu_{\alpha^*}$:

$$\Pr(closure|\alpha_{0i} > \mu_{\alpha^*}) = \Pr(closure|\tilde{\alpha}_{0i} > \mu_{\alpha^*}) \tag{14}$$

5) Expected value of alpha α_0 for liquidated funds::

$$E(\alpha_0|liquidation) = E(\tilde{\alpha}_{0i}|liquidation) \tag{15}$$

In the above equations (11) - (15), $\tilde{\alpha}_{0i}$ belongs to a simulated distribution F of α_0 according to the model specification with free parameters g_0, g_1, c_0, c_1, $\sigma^2_{\alpha^*}$. Further denote F^* to be a simulated distribution of α_0^* for observable funds that is derived from the model specification with parameters g_0, g_1, c_0, c_1, $\sigma^2_{\alpha^*}$. Then

6) Expected value of α_{1i}

$$\begin{aligned} E\left(\alpha_{1i}|\alpha_{1i} \text{ is observable}\right) &= E\left(\alpha_{1i}^*|\alpha_{0i}^* \sim F^*\right) \\ &= E\left(a + b\alpha_{0i}^* + \varepsilon_i|\alpha_{0i}^* \sim F^*\right) \\ &= a + bE\left(\alpha_{0i}^*|\alpha_{0i}^* \sim F^*\right) \end{aligned} \tag{16}$$

7) Variance of α_{1i}

$$\begin{aligned} Var\left(\alpha_{1i}|\alpha_{1i} \text{ is observable}\right) &= Var\left(a + b\alpha_{0i}^* + \varepsilon_i|\alpha_{0i}^*|\alpha_{0i}^* \sim F^*\right) \\ &= \sigma_\varepsilon^2 + b^2 Var\left(\alpha_{0i}^*|\alpha_{0i}^* \sim F^*\right) \end{aligned} \tag{17}$$

8) Covariance between α_{1i} and α_{0i}

$$\begin{aligned} & cov\left(\alpha_{1i}, \alpha_{0i}|\alpha_{1i} \text{ is observable}\right) \\ = \; & cov\left(a + b\alpha_{0i}^* + \varepsilon_i, \alpha_{0i}^* + u_i|\alpha_{0i}^* \sim F^*\right) \\ = \; & bVar\left(\alpha_{0i}^*|\alpha_{0i}^* \sim F^*\right) \end{aligned} \tag{18}$$

Notice that estimates for parameters g_0, g_1, c_0, c_1, $\sigma^2_{\alpha^*}$ can be obtained by solving the system of equations (11), (12), (13), (14), (15). The estimate for the slope b can be found from (18), the intercept a estimate can be computed from (16), and the variance σ_ε^2 estimate can be obtained from (17). This proves that the above eight moment conditions (11) - (18) specify the exactly identified case for estimating the set of parameters $P = (a, b, g_0, g_1, c_0, c_1, \sigma_\varepsilon, \sigma_{\alpha^*})$. We estimate the parameters and standard errors via the two step GMM procedure described in Hansen (1982) and Hansen and Singleton (1982) by numerically solving[23] the system of equations (11) - (18) for numerically simulated distributions F and F^*.

3.4.6 Biases in Simple vs. GMM Models

The OLS slope estimate from the naive regression (7) is equal to

$$\hat{b}_{OLS} = \frac{cov\left(\alpha_{1i}, \alpha_{0i}\right)}{Var\left(\alpha_{0i}\right)}, \tag{19}$$

and the consistent GMM estimator can be obtained from (18) as

$$\hat{b}_{GMM} = \frac{cov\left(\alpha_{1i}, \alpha_{0i}\right)}{Var\left(\alpha_{0i}^* | \alpha_{0i}^* \sim F^*\right)}. \tag{20}$$

In order to compare $\hat{b}_{OLS}$ and $\hat{b}_{GMM}$ estimators we have to account for the two types of estimation bias:

1) Measurement bias: $Var\left(\alpha_{0i}\right) > Var\left(\alpha_{0i}^*\right)$,

2) Look-ahead bias: $Var\left(\alpha_{0i}^*\right) > Var\left(\alpha_{0i}^* | \alpha_{0i}^* \sim F^*\right)$.

The combined effect of the above biases is that $Var\left(\alpha_{0i}\right) > Var\left(\alpha_{0i}^* | \alpha_{0i}^* \sim F^*\right)$, which results in

$$\left|\hat{b}_{OLS}\right| < \left|\hat{b}_{GMM}\right|.$$

This means that the naive regression OLS slope estimate (19) is biased toward zero compared to the GMM slope estimate (20).

4 Estimation Results

In this section we present the data and the results of the estimation of all the models proposed in the last section.

4.1 Data Description

The data for this research was generously provided by Hedge Fund Research. The database contains the history of monthly hedge fund returns beginning in 1990.[24] However, the

[23] We wold like to thank Ken Judd and Che-Lin Su for suggesting SNOPT software that we used in our algorithm. We also confirmed our approach by conducting Monte Carlo tests.

[24] For some funds, the history goes back to 1980s.

information about when a fund actually joined the database is only available since May 1996. Hence, we consider the time period from May 1996 until April 2005. We consider only hedge funds with dollar returns (both offshore and onshore), which report their returns as net of all fees. The yearly summary statistics is presented in table 2.

When a hedge fund joins the HFR database, it is given an option to select one strategy from the HFR list. These strategies are used in computation of monthly self reported style indices.[25] The indices are computed as returns on equally weighted portfolios of all funds using the same strategy.

4.2 Data Biases, Model Selection and Distribution of Alphas

In this section we demonstrate empirically how the distribution of hedge fund alphas is affected by different biases. In particular, we estimate three different models, eliminating one by one the problems related to the hedge fund data and then observe the differences in the distributions of alphas. Stale prices and changes in hedge fund strategies are considered. We run the following three regressions.

1. Stale prices are not taken into account:

$$R_{i,t} = \alpha_i + \beta_i R_t^{mkt} + \gamma_i \eta_t^{J,self} + \varepsilon_{i,t} \tag{21}$$

   We assume that residuals ($\varepsilon_{i,t}$) are i.i.d., so that the data is exposed to stale prices. To estimate hedge fund performance we use a market index, and a self declared style as benchmarks.

2. Now we take into account the stale prices. To do this we run a different regression:

$$\begin{aligned} R_{i,t} &= \alpha_i + \beta_{0,i} R_t^{mkt} + ... + \beta_{s,i} R_{t-s}^{mkt} \\ &\quad + \beta_{0,i}^{self} \eta_t^{J,self} + ... + \beta_{s,i}^{self} \eta_{t-s}^{J,self} + u_{i,t} \\ u_{i,t} &= \theta_0^i \varepsilon_{i,t} + ... + \theta_s^i \varepsilon_{i,t-s} \end{aligned} \tag{22}$$

   In this regression we include lags of the benchmarks, and assume that the error term ($u_{i,t}$) follows MA(s) process, ($\varepsilon_{i,t}$) are i.i.d. The number of lags is selected by SBC (Schwartz - Bayesian Criterion). For the details of the regression estimations see subsection 3.3.

[25] Only hedge funds with dollar returns reported on monthly basis, net of all fees are used in the computation of self reported indices. These indices are also free of the backfill bias, since backfill observations are excluded from index calculations.

3. To account for hedge funds changing their strategies overtime, we add an additional index into the regression (22). The additional index and the number of lags are selected by SBC. The regression equation is as follows:

$$\begin{aligned} R_{i,t} &= \alpha_i + \beta_{0,i} R_t^{mkt} + ... + \beta_{s,i} R_{t-s}^{mkt} \\ &\quad + \beta_{0,i}^{self} \eta_t^{J,self} + ... + \beta_{s,i}^{self} \eta_{t-s}^{J,self} \\ &\quad + \beta_{0,i}^{aux} \eta_t^{K,aux} + ... + \beta_{s,i}^{aux} \eta_{t-s}^{K,aux} + u_{i,t} \\ u_{i,t} &= \theta_0^i \varepsilon_{i,t} + ... + \theta_s^i \varepsilon_{i,t-s} \end{aligned} \tag{23}$$

In our estimation of the above regressions, we only consider hedge funds that had at least three years of observations. This leaves us with 1755 hedge funds. The brief summary statistics of alphas for the above three models are presented in table 3, and the final distribution of alphas from the third model is presented in figure 4. Since we use HFR indices in our regressions, and these indices are equally weighted averages of returns for all hedge funds within the same strategies, we might expect the mean and the median of all alphas to be approximately equal to zero and the number of positive alphas to be about fifty percent. This would have been true in the absence of an "inverse survivorship bias" that results from the fact that a successful fund, which was active over the length our study, would contribute only one positive alpha to the final sample, while a few unsuccessful funds that functioned over different times in our study would bring in several negative alphas to the same sample. Notice that the magnitude of the described "inverse survivorship bias" increases with the length of the time period considered. Taking this bias into consideration, we expect the mean and the median of all alphas in our sample to be less than zero and the number of positive alphas to be below fifty percent.

We can clearly see from table 3 that our alpha estimations in model 1 suffer from a severe positive bias. When we take into account stale prices, the percentage of positive alphas decreases from 59.2% to 48.38%. When we take into consideration stale prices along with an additional style index, the percentage of positive alphas goes further down to 46.27%. These results provide us with a preliminary indication of the effectiveness of our approach to estimating relative alphas.

4.3 Estimation of Hedge Fund Alphas

As described in the econometrics methodology section, in order to test for the persistence in hedge fund returns, we first estimate alphas α_{0i} in the evaluation period, then estimate alphas α_{1i} in the prediction period for the same hedge funds (if available) and proceed with a cross-section of hedge fund alphas (future and past alphas) which is tested for persistence. We form four overlapping cross-sections with three year evaluation and prediction periods using the nine years of available backfill bias free data. Table 4 shows the timeline for the estimation of alphas.

Notice that we cannot compute alphas α_{1i} for hedge funds that disappear from the database by the end of the evaluation period. We further winsorized the data at 1% in our subsequent analysis.

4.4 Performance Persistence

4.4.1 Simple (Naive) Regression

The first approach to check for persistence is to run the naive regression (7):

$$\alpha_{1i} = a + b\alpha_{0i} + \varepsilon_i.$$

The results of the naive regression are presented in table 5 and the scatter plots are presented in figure 5. The slope coefficient b is significant in two out of four cross-sections. The average estimate of performance persistence across all cross-sections is 14%. However, the persistence estimate, b, suffers from the downward bias due to measurement errors, and it also does not account for the fact that some hedge funds disappeared from the database due to different reasons. We address these biases in subsections that follow.

4.4.2 Weighted Least Squares Regression

Here we employ a method of weighted least squares in order to minimize the downward bias in persistence caused by measurement errors in alphas. We estimate the regression (9), i.e.

$$t^*_{\alpha_{1i}} = a + bt_{\alpha_{0i}} + \varepsilon_i,$$

where

$$t_{\alpha_{0i}} = \frac{\alpha_{0i}}{\sigma_{\alpha_0}}, \quad t^*_{\alpha_{1i}} = \frac{\alpha_{1i}}{\sigma_{\alpha_0}}.$$

The results of the weighted least squares regression are presented in table 6, and the scatter plots are presented in figure 6. The slope coefficient b is statistically significant[26] in all cross-sections, and the average estimate of performance persistence across all cross-sections is 23%. However, the magnitude of the persistence estimate, b, is noticeably smaller in the third cross-section. That cross-section has the closest breaking point to the worst

[26] At the 1% significance level.

overall performance year for the hedge fund industry.[27] This suggests that the superior skill that is reflected in our measure of the relative performance persistence may not be as valuable to an investor during periods of adverse economic conditions for the hedge fund industry as a whole. We conjecture that when there are fewer opportunities in the economy for hedge fund managers as a group, there will be less cross sectional dispersion in managers' alphas, i.e., the performance differential between the more talented and the less talented managers is likely to be less pronounced. We leave modeling this dependence of relative performance on market conditions to future research.

It is also important to note that an investor can only benefit from our approach by investing in hedge funds run by talented managers, and staying away from the ones that have not demonstrated persistent skill. Hence it may be of little value to an investor to find evidence of negative performance persistence, since an investor cannot take a short position in a hedge fund. On the other hand, evidence of positive performance persistence could be extremely valuable, since it would allow an investor to achieve superior returns by taking long positions in hedge funds run by talented managers.

We study whether positive or negative performance persists by running regression (9) separately for funds in the upper and the lower terciles according to their alpha t-statistic during the evaluation period. Remarkably, we find strong evidence of performance persistence among top hedge funds, while we find no evidence of persistence among bottom funds. These results are presented in table 7. We further extend our analysis of the role of superior managerial talent by studying performance persistence separately among the top and bottom 10% of the funds. The results are presented in table 8. We find even stronger evidence of performance persistence among the top 10% of the funds, while the results for the bottom 10% of the funds fail to provide a consistent picture of persistence. This is consistent with the interpretation of superior performance persistence as a result of superior managerial talent, which is also reflected in superior prior performance. Our findings also support the view that managers of superior skills restrict inflow of new money in order to maintain their performance.

4.4.3 GMM Estimation

During the prediction period, a hedge fund can either remain or drop out from the database. Funds may disappear from the database due to liquidation, closing, or stop reporting for unknown reasons. Summary statistics of hedge funds according to this decomposition are presented in tables 9 and 10. If probabilities of liquidation and closure are influenced

[27] Measured by the HFR total index.

by fund's "true" prior performance, α_0^*, it will result in biased persistence estimates, which is also known as a look-ahead bias. Considering histograms of distributions of liquidated and closed funds by deciles of α_0 (figure 7) and conditional probabilities of liquidation and closure conditional on α_0 being in top and bottom parts of its distribution (table 11), we conclude that there is a relationship between fund's prior performance and probabilities of fund's liquidation and closure. We model this relationship by specifying different patterns of liquidation and closure for the top and bottom parts of the true alpha distribution through model (M) with specification (P). This approach allows us take into account measurement errors along with the look-ahead bias, and it is estimated via the GMM procedure. Estimates from the GMM procedure are provided tables 12 and 13. The estimates of the persistence coefficient b are roughly consistent with the weighted least squares estimates from subsection 4.4.2, and the average GMM estimate of performance persistence across all cross-sections is 26% compared to the weighted least squares average of 23%. GMM estimated conditional probabilities of liquidation and closure (figure 8) are also consistent with observed probabilities in table 11 and figure 7.

Notice that in the first two cross-sections liquidated funds tend to have low alphas, while closed funds tend to have high alphas (see table 9). This is consistent with our statistical model (M) and the specification (P), but it is the only consequence of the model. In fact, the specification (P) is flexible to allow decreasing probabilities of closure with increasing α_0, as demonstrated by negative g_1 parameter estimates in third and fourth cross-sections. These estimates are consistent with descriptive statistics in the last two cross-sections, as closed funds do not outperform liquidated and observable funds (see table 10).

However, it is worth pointing out that the underlying fundamentals of the decision to close a fund to new investors might have changed after 2001. In order to test this conjecture we performed probit tests of the decision to liquidate vs. close among the funds that were either closed or liquidated in our data. The estimates of the probability of liquidation are provided in table 14. The results indicate that while α_0 was significant in the liquidation vs. closure decision in the first two cross-sections, it is not significant in the last two cross-sections, while the ratio of last flows to assets becomes significant in the last two cross-sections. This supports our conjecture that the role of the relative performance measure, α_0, in the decision to liquidate or close a fund has diminished since 2001.

4.4.4 Non-Reporting Funds

The non-reporting funds[28] comprise on average 15.55% of the data among all cross-sections. Can we use these funds in our further performance analysis? The answer to this question lies in the distribution of observable characteristics of the non-reporting funds during the evaluation period. We may attempt to classify the non-reporting funds as closed or liquidated on the basis of their evaluation period performance α_0. Such classification would be consistent with assumptions of the model (M) and the specification (P), but only if the distribution of the relative performance measure α_0 for non-reporting funds resembles the distributions of α_0 for funds that stopped reporting, but indicated a reason for doing so (i.e. liquidated and closed funds). Unfortunately, Kolmorogov-Smirnov test for distribution closeness does not indicate consistently close fit between the distribution of non-reporting funds and the distribution of liquidated and closed funds. In fact, best matches between the distribution of non-reporting funds and the distribution of liquidated and closed funds come from the first and the fourth cross-sections, while in the second and the third cross-sections the non-reporting funds distribution is closest to the distribution of all reporting funds.[29]

Hence we conclude that classifying non-reporting funds as either closed or liquidated would result in model (M) misspecification, and that treating non-reporting funds as missing data would be the most consistent approach.

4.4.5 Potential Biases

Here we consider a possibility of a scenario when funds with large positive alphas during the first three year period perform poorly during the second three year period and liquidate, and funds with large negative alphas during the first three year period perform extremely well during the second three year period and close. Such a pattern could contribute to findings positive measured persistence in alphas of funds that survived during both three year periods.

However, as seen in figure 7 and tables 9, 10, and 11 funds with lower performance during the first period were more likely to be liquidated. This indicates that the scenario of performance reversal for liquidated funds between the two periods is unlikely.

In case of closed funds, figure 7 along with tables 9 and 11 indicate that in the first two cross-sections funds with higher prior performance were more likely to be closed. This does not suggest performance reversal in the first two cross-sections. On the other hand, in the last two cross-sections (see figure 7 and tables 10 and 11) closed funds with lower first period

[28] The non-reporting funds are those that dropped out of the HFR dataset without reporting a reason.

[29] See table 15 for Kolmorogov-Smirnov test results.

performance were more likely to be closed, which could be an indication of performance reversal among a subset of underperforming funds in the first period. If that was the case, we would have been more likely to find an indication of stronger performance persistence among the lower performing hedge funds. Nevertheless, our weighted least squares analysis produced no evidence of performance persistence among the lower performing hedge funds, hence we conclude that it is unlikely that there could be a performance reversal pattern strong enough to significantly bias our previous findings of performance persistence.

While the above observations allow us to suggest that our finding of performance persistence is not a spurious phenomenon, a completely definitive answer on the matter could only be obtained by completely eliminating the bias caused by funds dropping out of the database by tracking down the performance of all the funds that dropped out without being completely liquidated.

4.5 Portfolio Performance Interpretation

Here we attempt to interpret the significance of the main results about performance persistence. We construct portfolios of hedge funds based on their past performance in the evaluation period, and then track their performance during the prediction period. However, the fact that some hedge funds disappear in the prediction period makes it impossible to track portfolio performances exactly. Hence, it is impossible to make an unbiased portfolio performance comparison during the prediction period. Here we attempt to estimate the persistence magnitude using all the hedge funds that are available in the evaluation period.

We sort all the hedge funds by their evaluation period performance measured by the t-statistic of alpha. We compose an inferior portfolio of all hedge funds in the bottom 10% of all funds, a superior portfolio of all funds in the top 10%, and a neutral portfolio of all the remaining funds. We then invest one dollar to every portfolio in the beginning of the prediction period. One dollar is equally split among all the hedge funds in a given portfolio.

We consider three scenarios. Under a pessimistic scenario we assume that the money invested into disappeared hedge funds cannot be recovered at all. That is, if a hedge fund disappears from our database, we lose all the money invested there, regardless of the reason the hedge fund disappeared. The pessimistic scenario is modeled by assigning -100% return to a fund during the month after its disappearance from the database, and zero returns thereafter. Under a neutral scenario, we assume that we can take all the money from a disappeared hedge fund and invest it into the HFR index of the strategy that the fund was following. This implies a zero alpha strategy after the fund's disappearance, and it is modeled as follows. We estimate α_1 using available observations, and then take a weighted average of the resulting alpha and zero with weights representing the number of observations available and the number of observations missing during the evaluation period. Under a realistic scenario we assume that we reinvest all the money from a disappeared fund into the surviving funds within the group. We assume a fund's α_1 to be a weighted average of the available α_1 and the average α_1 of the surviving funds within the group.

For each scenario we calculate each portfolio performance as an equally weighted average of individual fund alphas. This is summarized in figure 9. Under the assumptions of our model, liquidated hedge funds performed poorly, and closed funds performed well, relative to other hedge funds in the evaluation period. In the pessimistic scenario we may significantly underestimate the performance of every portfolio, since, in reality, some money can be recovered even from liquidated funds. In the neutral scenario, the relationship of the estimated performance to the actual performance is more ambiguous. The performance of the inferior portfolio is probably overestimated, as liquidation would be the main reason for fund disappearance. The performance of the superior portfolio is most likely underestimated, as the main reason for a fund to disappear is to close for new investors. One can expect that such hedge funds will perform better than average funds in the prediction period. It is difficult to make any preliminary conclusions for the neutral portfolio. Arguably, the realistic scenario has the closest resemblance with the actual portfolio performances. These results are summarized in table 16.

We report the performance of the three portfolios in the evaluation and prediction periods in table 17. The performance for the evaluation period is presented in column 'Past Alpha'. We also report the performance of all portfolios during the evaluation period, as well as the statistics about surviving funds in each portfolio.

As we see from table 17, portfolio performances in the neutral scenario are in line with predictions from table 16. The inferior portfolio's performance in realistic scenario falls between performance estimates in pessimistic and neutral scenarios. The superior portfolio's performance in realistic scenario is higher than performance estimates in both the neutral and pessimistic scenarios.

In all cross-sections the superior portfolio outperforms the inferior portfolio under all three scenarios, and the difference is statistically significant in three out of four cross-sections. These results indicate that in reality we should expect the superior portfolio to outperform the inferior portfolio.

5 Conclusion

Hedge fund managers are given much more flexibility regarding where and how to invest compared to mutual fund managers. The growth of hedge funds, with 1.3 trillion dollars invested in assets by 2006, may well reflect the need for giving talented managers who know where superior opportunities exist at a given point in time the necessary flexibility to exploit that talent. A natural question that arises is whether it is possible to identify those hedge fund managers who are able to exploit the flexibility given to them better than others.

While the flexibility given to hedge fund managers may help in generating superior returns, it also makes performance evaluation more difficult. Hedge fund returns are unlike returns from standard asset classes, and exhibit option-like features that have to be taken into account when evaluating performance. Further, since hedge funds invest in illiquid assets, care has to be exercised in measuring their systematic risk. In this paper we develop a method for evaluating the performance of a hedge fund manager relative to a suitably constructed peer group. Our method takes into account option-like features in hedge fund strategies and serial correlation in hedge fund returns caused possibly by investments in illiquid assets. We also take into account the backfill bias in our data set and the look-ahead bias (i.e. the fact that a hedge fund may be liquidated or closed and exit the data set). We employ a method of weighted least squares in order to minimize the downward bias in persistence caused by measurement errors in alphas.

We find evidence of persistence in the performance of funds relative to their style benchmarks. It appears that on average more than 20% of the abnormal performance during a three year interval will spill over into the following three year interval. We provider further support for the interpretation of performance persistence as evidence of superior managerial talent by finding strong evidence of performance persistence among top hedge funds, while finding little evidence of persistence among bottom funds.

Our analysis highlights difficulties that arise in predicting how a hedge fund manager will perform in the future relative to his peer group. While the assumptions we had to make in order to answer the question of performance persistence among hedge fund managers appear reasonable, we need a better understanding of what happened to funds that vanished from publicly available databases to provide a quantitative answer to that question with utmost confidence. We hope that our findings will stimulate research examining how funds that discontinue reporting their performance do subsequently.

References

Ackerman, C., R. McEnally, and D. Ravenscraft, 1999, "The Performance of Hedge Funds: Risk, Return, and Incetives," *Journal of Finance*, 54, 833–874.

Agarwal, V., and N. Y. Naik, 2000a, "Multi-Period Performance Persistance Analysis of Hedge Funds," *Journal of Financial and Quantitative Analysis*, 35, 327–342.

Agarwal, V., and N. Y. Naik, 2000b, "On Taking the Alternative Route: Risks, Rewards, and Performance Persistence of Hedge Funds," *Journal of Alternative Investments*, 2, 6–23.

Agarwal, V., and N. Y. Naik, 2004, "Risk and Portfolio Decisions Involving Hedge Funds," *Review of Financial Studies*, 17, 63–98.

Asness, C., R. Krail, and J. Liew, 2001, "Do Hedge Funds Hedge?," *Journal of Portfolio Management*, pp. 6–19.

Baquero, G., J. Ter Horst, and M. Verbeek, 2005, "Survival, Look-Ahead Bias and the Persistence in Hedge Fund Performance," *Journal of Financial and Quantitative Analysis*, 40, 493–517.

Bares, P.-A., R. Gibson, and S. Gyger, 2003, "Performance in the Hedge Fund Industry: An Analysis of Short and Long-Term Persistence," *Journal of Alternative Investments*, 6, 25–41.

Ben Dor, A., R. Jagannathan, and I. Meier, 2003, "Understanding Mutual Fund and Hedge Fund Styles Using Return-Based Style Analysis," *Journal of Investment Management*, 1, 94–134.

Berk, J. B., and R. C. Green, 2004, "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, 112, 1269–1295.

Bondarenko, O., 2004, "Market Price of Variance Risk and Performance of Hedge Funds," *Working Paper*.

Boyson, N. M., and M. J. Cooper, 2004, "Do Hedge Funds Exhibit Performance Persistence? A New Approach," *Working Paper*.

Brown, S. J., W. N. Goetzmann, and R. G. Ibbotson, 1999, "Offshore Hedge Funds: Survival and Performance 1989-1995," *Journal of Business*, 72, 91–118.

Brown, S. J., W. N. Goetzmann, and J. Park, 2001, "Careers and Survival: Competition and Risk in the Hedge Fund and CTA Industry," *Journal of Finance*, 56, 1869–1886.

Capocci, D., and G. Hübner, 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.

Edwards, F. R., and M. O. Caglayan, 2001, "Hedge Fund Performance and Manager Skill," *Journal of Futures Markets*, 21, 1003–1028.

Fung, W., and D. A. Hsieh, 1997, "Empirical Charachteristics of Dynamic Trading Strategies: The Case of Hedge Funds," *Review of Financial Studies*, 10, 275–302.

Fung, W., and D. A. Hsieh, 2000, "Performance Characteristics of Hedge Funds and Commodity Funds: Natural Vs. Spurios Biases," *Journal of Financial and Quantitative Analysis*, 35, 291–307.

Fung, W., and D. A. Hsieh, 2001, "The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers," *Review of Financial Studies*, 14, 313–341.

Fung, W., and D. A. Hsieh, 2002, "Hedge-Fund Benchmarks: Information Content and Biases," *Financial Analysts Journal*, 58, 22–34.

Fung, W., D. A. Hsieh, N. Y. Naik, and T. Ramadorai, 2007, "Hedge Funds: Performance, Risk and Capital Formation," *Journal of Finance*, forthcoming.

Getmansky, M., A. W. Lo, and I. Makarov, 2004, "An Econometric Model of Serial Correlation and Illiquidity in Hedge Fund Returns," *Journal of Financial Economics*, 74, 529–609.

Glosten, L., and R. Jagannathan, 1994, "A Contingent Claims Framework for Analyzing the Performance of Portfolio Managers," *Journal of Empirical Finance*, 1, 133–160.

Hansen, L. P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Hansen, L. P., and K. J. Singleton, 1982, "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.

Kosowski, R., N. Y. Naik, and M. Teo, 2007, "Do Hedge Funds Deliver Alpha? A Bayesian and Bootstrap Analysis," *Journal of Financial Economics*, 84, 229–264.

Liang, B., 2000, "Hedge Funds: The Living and the Dead," *Journal of Financial and Quantitative Analysis*, 35, 309–326.

Lo, A., 2002, "The Statistics of Sharpe Ratios," *Financial Analysts Journal*, 58, 36–50.

Mitchell, M., and T. Pulvino, 2001, "Characteristics of Risk and Return in Risk Arbitrage," *Journal of Finance*, 56, 2135–2175.

Okunev, J., and D. White, 2003, "Hedge Fund Risk Factors and Value at Risk of Credit Trading Strategies," *Working Paper*.

Pastor, L., and R. F. Stambaugh, 2002a, "Investing in Equity Mutual Funds," *Journal of Financial Economics*, 63, 351–380.

Pastor, L., and R. F. Stambaugh, 2002b, "Mutual Fund Performance and Seemingly Unrelated Assets," *Journal of Financial Economics*, 63, 315–349.

Perold, A., 1999, "Long Term Capital Management, LP (A)," *HBS Case Study No. 9-200-007*.

Scholes, M., and J. Williams, 1977, "Estimating Betas from Nonsynchronous Data," *Journal of Financial Economics*, 5, 309–327.

Schwarz, G., 1978, "Estimating the Dimension of a Model," *Annals of Statistics*, 6, 461–464.

Weisman, A. B., 2002, "Informationless Investing and Hedge Fund Performance Measurement Bias," *Journal of Portfolio Management*, pp. 80–91.

#	HFR Strategy Style Index	#	HFR Strategy Style Index
1	Convertible Arbitrage	17	Fund of Funds: Conservative
2	Distressed Securities	18	Fund of Funds: Diversified
3	Emerging Markets: Asia	19	Fund of Funds: Market Defensive
4	Emerging Markets: E. Europe/CIS	20	Fund of Funds: Strategic
5	Emerging Markets: Global	21	Macro
6	Emerging Markets: Latin America	22	Market Timing
7	Equity Hedge	23	Merger Arbitrage
8	Equity Market Neutral	24	Regulation D
9	Equity Non-Hedge	25	Relative Value Arbitrage
10	Event-Driven	26	Sector: Energy
11	Fixed Income: Arbitrage	27	Sector: Financial
12	Fixed Income: Convertible Bonds	28	Sector: Health Care/Biotechnology
13	Fixed Income: Diversified	29	Sector: Miscellaneous
14	Fixed Income: High Yield	30	Sector: Real Estate
15	Fixed Income: Mortgage-Backed	31	Sector: Technology
16	Fund of Funds (Total)	32	Short Selling

Table 1: Style indices in Hedge Fund Research database.

year	total	entered	left	attrition	mean return	median return	std. dev.
1996	1123	1123	91	8.10%	0.57%	0.61%	5.10%
1997	1326	294	163	12.29%	1.14%	0.86%	5.31%
1998	1436	273	206	14.35%	-0.19%	0.23%	7.98%
1999	1479	249	199	13.46%	1.50%	0.67%	7.97%
2000	1546	266	251	16.24%	-0.40%	0.12%	7.12%
2001	1851	556	204	11.02%	0.12%	0.24%	4.64%
2002	2183	536	277	12.69%	-0.09%	0.13%	4.34%
2003	2744	838	281	10.24%	1.11%	0.76%	3.31%
2004	3274	811	364	11.12%	0.23%	0.20%	2.86%

Table 2: Yearly distribution of hedge funds. The table presents the total number of funds that reported during a year, the number of funds that entered and left the database, and mean, median, and standard deviation of monthly excess returns. A year represents the time period from May of that year until April of the next year.

model	**description**	**mean**	**median**	**percent of positive alphas**
1	stale prices	.03818	.12477	59.20%
2	no stale prices	-.14263	-.03366	48.38%
3	multiple indices	-.18719	-.06130	46.27%

Table 3: Summary statistics for alpha. Statistics are provided for three models designed to correct different data biases. Alphas are measured as monthly percentage returns. There are 1755 funds with at least a three-year history available, after excluding backfill observations.

	Evaluation Period		**Prediction Period**	
Cross-section	**Begins**	**Ends**	**Begins**	**Ends**
1	May 1996	April 1999	May 1999	April 2002
2	May 1997	April 2000	May 2000	April 2003
3	May 1998	April 2001	May 2001	April 2004
4	May 1999	April 2002	May 2002	April 2005

Table 4: Timeline for evaluation and prediction periods.

	1996-1999 - 1999-2002			1997-2000 - 2000-2003		
Parameter	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
a	-0.1135	-1.26	0.2101	-0.0457	-0.93	0.3503
b	**0.3276**	**3.69**	**0.0003**	**0.0775**	**1.78**	**0.0754**
	1998-2001 - 2001-2004			1999-2002 - 2002-2005		
Parameter	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
a	-0.0869	-2.37	0.0184	-0.2456	-6.59	<.0001
b	**0.0030**	**0.09**	**0.9319**	**0.1537**	**4.69**	**<.0001**

Table 5: Naive regression results. Persistence is captured by the slope coefficient b, which is statistically significant in two out of four cross-sections.

	1996-1999 - 1999-2002			1997-2000 - 2000-2003		
Parameter	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
a	-0.0602	-0.39	0.6933	0.0353	0.38	0.7012
b	**0.3835**	**6.02**	**<.0001**	**0.2758**	**6.45**	**<.0001**
	1998-2001 - 2001-2004			1999-2002 - 2002-2005		
Parameter	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
a	-0.1510	-1.91	0.0561	-0.5709	-7.27	<.0001
b	**0.0935**	**2.71**	**0.0070**	**0.1822**	**5.33**	**<.0001**

Table 6: Weighted least squares regression results. Persistence is captured by the slope coefficient b, which is statistically significant in all cross-sections.

		Top 33%			Bottom 33%		
Cross-section	Parameter	Estimate	t-stat	p-value	Estimate	t-stat	p-value
1996-1999 - 1999-2002	b	0.7425	5.21	<.0001	0.1842	0.68	0.4958
1997-2000 - 2000-2003	b	0.5706	6.20	<.0001	-0.0929	-0.44	0.6603
1998-2001 - 2001-2004	b	0.2072	2.24	0.0262	-0.0032	-0.03	0.9767
1999-2002 - 2002-2005	b	0.2769	2.92	0.0040	0.1034	0.83	0.4055

Table 7: Weighted least squares regression results. Persistence is estimated separately for top and bottom terciles of the t_{α_0} ranking.

		Top 10%			Bottom 10%		
Cross-section	Parameter	Estimate	t-stat	p-value	Estimate	t-stat	p-value
1996-1999 - 1999-2002	b	1.0850	4.25	0.0002	-2.1255	-2.14	0.0415
1997-2000 - 2000-2003	b	0.8040	4.50	<.0001	-0.6653	-1.54	0.1303
1998-2001 - 2001-2004	b	0.4892	2.13	0.0374	0.1726	0.54	0.5858
1999-2002 - 2002-2005	b	0.4846	2.20	0.0322	-0.1224	-0.36	0.7210

Table 8: Weighted least squares regression results. Persistence is estimated separately for top and bottom 10 percent of the t_{α_0} ranking.

1996-1999 - 1999-2002	Observable	Liquidated	Closed	Non-Reporting	Total
number of hedge funds	319	61	25	88	493
percent	64.71%	12.37%	5.07%	17.85%	100%
α_0 mean	-0.0804	-0.2200	0.1974	-0.0068	-0.0704
α_0 median	0.0371	-0.2332	0.2552	0.0558	0.0363
α_0 std. dev.	1.0177	1.2528	1.1194	1.1927	1.0866
assets ($M) mean	235.29	21.47	58.52	95.09	174.70
assets ($M) median	54.73	6.35	31.25	20.00	39.00
assets ($M) std. dev.	652.36	34.37	71.04	180.83	536.23
1997-2000 - 2000-2003	Observable	Liquidated	Closed	Non-Reporting	Total
number of hedge funds	456	73	32	112	673
percent	67.76%	10.85%	4.75%	16.64%	100%
α_0 mean	-0.0434	-0.2525	0.2257	-0.1444	-0.0701
α_0 median	0.1078	-0.0190	0.6111	0.1375	0.1401
α_0 std. dev.	1.1250	1.3671	1.2597	1.4254	1.2146
assets ($M) mean	226.94	31.99	55.73	70.60	170.94
assets ($M) median	57.38	7.21	10.55	18.40	37.98
assets ($M) std. dev.	622.89	89.13	93.92	141.05	521.93

Table 9: Distribution of hedge funds in the prediction period from first and second cross-sections. Alphas are measured as monthly percentage returns.

1998-2001 - 2001-2004	Observable	Liquidated	Closed	Non-Reporting	Total
number of hedge funds	507	83	37	96	723
percent	70.12%	11.48%	5.12%	13.28%	100%
α_0 mean	-0.0027	-0.1819	-0.1560	-0.0761	-0.0408
α_0 median	0.1433	-0.0649	-0.0557	0.0588	0.1010
α_0 std. dev.	1.0648	1.2713	1.4846	1.1531	1.1257
assets ($M) mean	301.38	58.99	59.24	77.32	230.62
assets ($M) median	68.00	10.49	11.18	17.76	44.68
assets ($M) std. dev.	723.13	263.84	105.38	155.66	623.04
1999-2002 - 2002-2005	Observable	Liquidated	Closed	Non-Reporting	Total
number of hedge funds	519	103	30	110	762
percent	68.11%	13.52%	3.94%	14.44%	100%
α_0 mean	0.0569	-0.0918	-0.7322	-0.1780	-0.0281
α_0 median	0.1719	-0.0260	-0.2865	0.0039	0.1063
α_0 std. dev.	1.1370	1.0738	1.6988	0.9966	1.1467
assets ($M) mean	328.51	33.08	38.69	179.04	256.73
assets ($M) median	80.00	9.98	10.10	11.46	41.30
assets ($M) std. dev.	688.56	100.58	72.13	921.87	678.70

Table 10: Distribution of hedge funds in the prediction period from third and fourth cross-sections. Alphas are measured as monthly percentage returns.

Cross-section	Pr(L|$\alpha_0 \leq \mu_{\alpha_0}$)	Pr(L|$\alpha_0 > \mu_{\alpha_0}$)	Pr(C|$\alpha_0 \leq \mu_{\alpha_0}$)	Pr(C|$\alpha_0 > \mu_{\alpha_0}$)
1996-1999 - 1999-2002	0.1889	0.1200	0.0389	0.0800
1997-2000 - 2000-2003	0.1535	0.1141	0.0439	0.0661
1998-2001 - 2001-2004	0.1581	0.1127	0.0699	0.0507
1999-2002 - 2002-2005	0.1789	0.1417	0.0596	0.0354

Table 11: Observed probabilities of liquidation and closure conditional on observed values of α_0.

	1996-1999 - 1999-2002			1997-2000 - 2000-2003		
Parameter	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
a	-0.0951	-1.4711	0.1419	-0.0385	-0.9626	0.3361
b	**0.5404**	**2.5292**	**0.0117**	**0.2761**	**1.8557**	**0.0639**
g_0	0.0775	0.5944	0.5525	0.1611	0.2168	0.8284
g_1	0.0753	0.4077	0.6837	0.0894	0.3218	0.7477
c_0	0.0772	2.1987	0.0284	0.0889	0.9973	0.3190
c_1	0.0372	1.2166	0.2243	0.0261	0.8775	0.3805
σ_ε	1.5850	11.1591	<.0001	1.0328	19.6883	<.0001
σ_{α^*}	0.8157	0.9766	0.3293	0.6190	0.5438	0.5868

Table 12: Results of the GMM procedure for the first two cross-sections.

	1998-2001 - 2001-2004			1999-2002 - 2002-2005		
Parameter	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
a	-0.0868	-2.9313	0.0035	-0.2508	-7.9928	<.0001
b	**0.0032**	**0.0866**	**0.9310**	**0.2388**	**4.2732**	**<.0001**
g_0	0.0430	0.9182	0.3588	0.0796	0.8879	0.3749
g_1	-0.0235	-0.7482	0.4546	-0.0396	-0.9694	0.3327
c_0	0.1169	7.3773	<.0001	0.1473	11.4650	<.0001
c_1	0.0717	7.4678	<.0001	0.0690	7.4634	<.0001
σ_ε	0.8248	26.9162	<.0001	0.8367	17.2526	<.0001
σ_{α^*}	1.0265	1.4153	0.1406	0.9241	1.9986	0.0460

Table 13: Results of the GMM procedure for the last two cross-sections.

	1996-1999 - 1999-2002			1997-2000 - 2000-2003		
Parameter	Estimate	ChiSq	Pr>ChiSq	Estimate	ChiSq	Pr>ChiSq
Intercept	0.4860	8.62	0.0033	0.3375	4.94	0.0262
α_0	-0.2957	3.77	0.0521	-0.1736	2.65	0.1039
last_returns	-0.1404	5.84	0.0156	-0.1013	2.55	0.1103
last_flows_to_assets	-0.0110	0.09	0.7646	-0.0422	1.03	0.3091
	1998-2001 - 2001-2004			1999-2002 - 2002-2005		
Parameter	Estimate	ChiSq	Pr>ChiSq	Estimate	ChiSq	Pr>ChiSq
Intercept	0.3831	7.30	0.0069	0.6782	20.44	<.0001
α_0	-0.0551	0.34	0.5613	0.0146	0.01	0.9046
last_returns	0.0285	0.33	0.5636	0.0607	0.82	0.3653
last_flows_to_assets	-0.1970	3.04	0.0813	-0.3962	4.43	0.0353

Table 14: Probit estimates of the probability of liquidation. α_0 is estimated over the evaluation period. *last returns* is the cumulative fund's return over the last year of a fund's presence in the database. *last flows to assets* is a ratio of cumulative cash flows over a fund's last assets over the last year of a fund's presence in the database.

	1996-1999 - 1999-2002		1997-2000 - 2000-2003	
Distributions	KSa statistic	p-value	KSa statistic	p-value
Observable funds	1.2692	0.0798	0.5763	0.8940
Liquidated and closed funds	0.7146	0.6869	0.7756	0.5843
All reporting funds	1.0511	0.2192	0.5551	0.9176
	1998-2001 - 2001-2004		1999-2002 - 2002-2005	
Distributions	KSa statistic	p-value	KSa statistic	p-value
Observable funds	0.7011	0.7095	1.0822	0.1920
Liquidated and closed funds	0.6694	0.7613	0.6974	0.7155
All reporting funds	0.6608	0.7752	0.9401	0.3398

Table 15: Kolmogorov-Smirnov tests for closeness of alpha 0 distributions. KSa statistic denotes the asymptotic Kolmogorov-Smirnov statistic, and the p-value is provided for the test of the null hypothesis that there is no difference between the two distributions. The non-reporting funds distribution is compared to the observable funds distribution, liquidated and closed funds distribution, and to the all reporting funds (i.e. observable, liquidated, and closed funds) distribution.

	Pessimistic Scenario	$\gtreqless$	Actual Portfolio	$\gtreqless$	Neutral Scenario
Inferior	$\bullet$	$<$	$\bullet$	$<$	$\bullet$
Neutral	$\bullet$	$<$	$\bullet$	?	$\bullet$
Superior	$\bullet$	$<$	$\bullet$	$>$	$\bullet$

Table 16: Performance comparison under pessimistic and neutral scenarios to the actual case

1996-1999 - 1999-2002					
Portfolios \ Performance	Past Alpha	Survivors Alpha	Pessimistic Scenario	Neutral Scenario	Realistic Scenario
Inferior (lowest performance in the past)	-2.1111 N = 50	-1.0552*** N = 29	-1.4392	-0.6352***	-0.9324***
Neutral	0.0122 N = 393	-0.1189 N = 163	-1.3337	-0.2040	-0.2270
Superior (highest performance in the past)	1.2733 N = 50	0.3942*** N = 30	-0.6294	0.3253***	0.4291***

1997-2000 - 2000-2003					
Portfolios \ Performance	Past Alpha	Survivors Alpha	Pessimistic Scenario	Neutral Scenario	Realistic Scenario
Inferior (lowest performance in the past)	-2.1552 N = 69	-0.0562* N = 45	-1.2254**	-0.1303**	-0.1445***
Neutral	0.0194 N = 535	-0.1407 N = 359	-1.1024	-0.1864	-0.2177
Superior (highest performance in the past)	1.1120 N = 69	0.2374* N = 52	-0.1359**	0.1974**	0.2395***

1998-2001 - 2001-2004					
Portfolios \ Performance	Past Alpha	Survivors Alpha	Pessimistic Scenario	Neutral Scenario	Realistic Scenario
Inferior (lowest performance in the past)	-2.4095 N = 74	-0.0751 N = 43	-1.6713**	-0.0996	-0.1221
Neutral	0.0874 N = 575	-0.1027 N = 406	-1.1880	-0.1574	-0.1765
Superior (highest performance in the past)	1.1059 N = 74	0.0240 N = 57	-0.6665**	0.0042	0.0073

1999-2002 - 2002-2005					
Portfolios \ Performance	Past Alpha	Survivors Alpha	Pessimistic Scenario	Neutral Scenario	Realistic Scenario
Inferior (lowest performance in the past)	-2.2106 N = 78	-0.5550*** N = 50	-1.0685	-0.4441***	-0.5591***
Neutral	0.0709 N = 606	-0.2715 N = 411	-0.9630	-0.2448	-0.2940
Superior (highest performance in the past)	1.1960 N = 78	0.0131*** N = 59	-1.0206	0.0154***	0.0180***

Table 17: Out of sample performance of three portfolios. Portfolios are formed and ranked according to the previous relative t-alpha performance in the evaluation period. Then portfolio alphas in the prediction period are caclulated under pessimistic and neutral scenarios. Differences between superior and inferior portfloio performances marked with ***, **, and * are statistically significant at 1, 5, and 10 percent respectively.

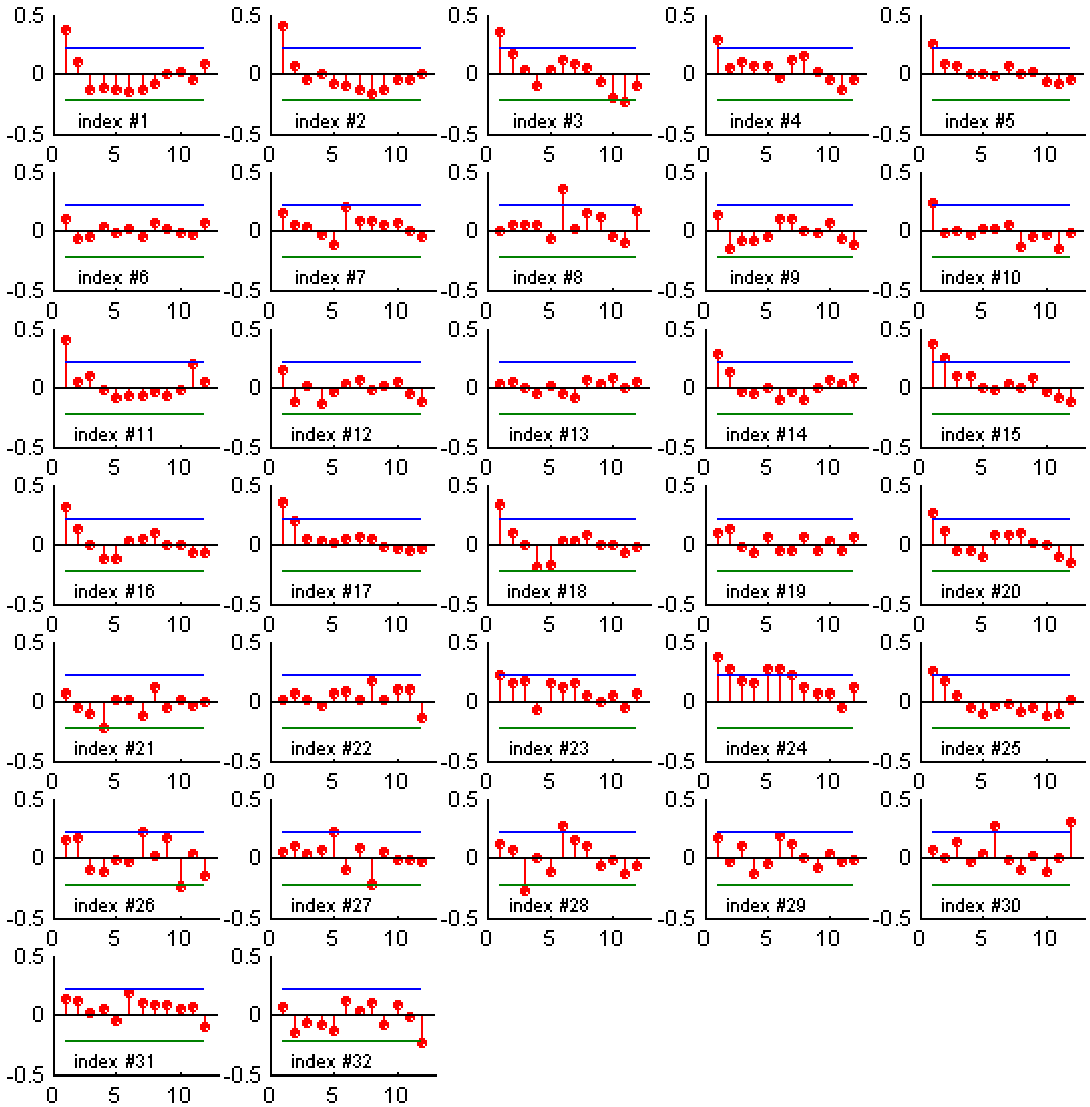

Figure 1: The autocorrelation functions for style indices are presented in this figure. The style indices used are *before* the adjustment for smoothing (i.e. as they were presented in the original database). The autocorrelations were computed for lags from 1 to 12. The thin horizontal lines around the horizontal axes represent 95% confidence intervals. Style index names can be retrived from table 1. For example, index #1 stands for Convertible Arbitrage index.

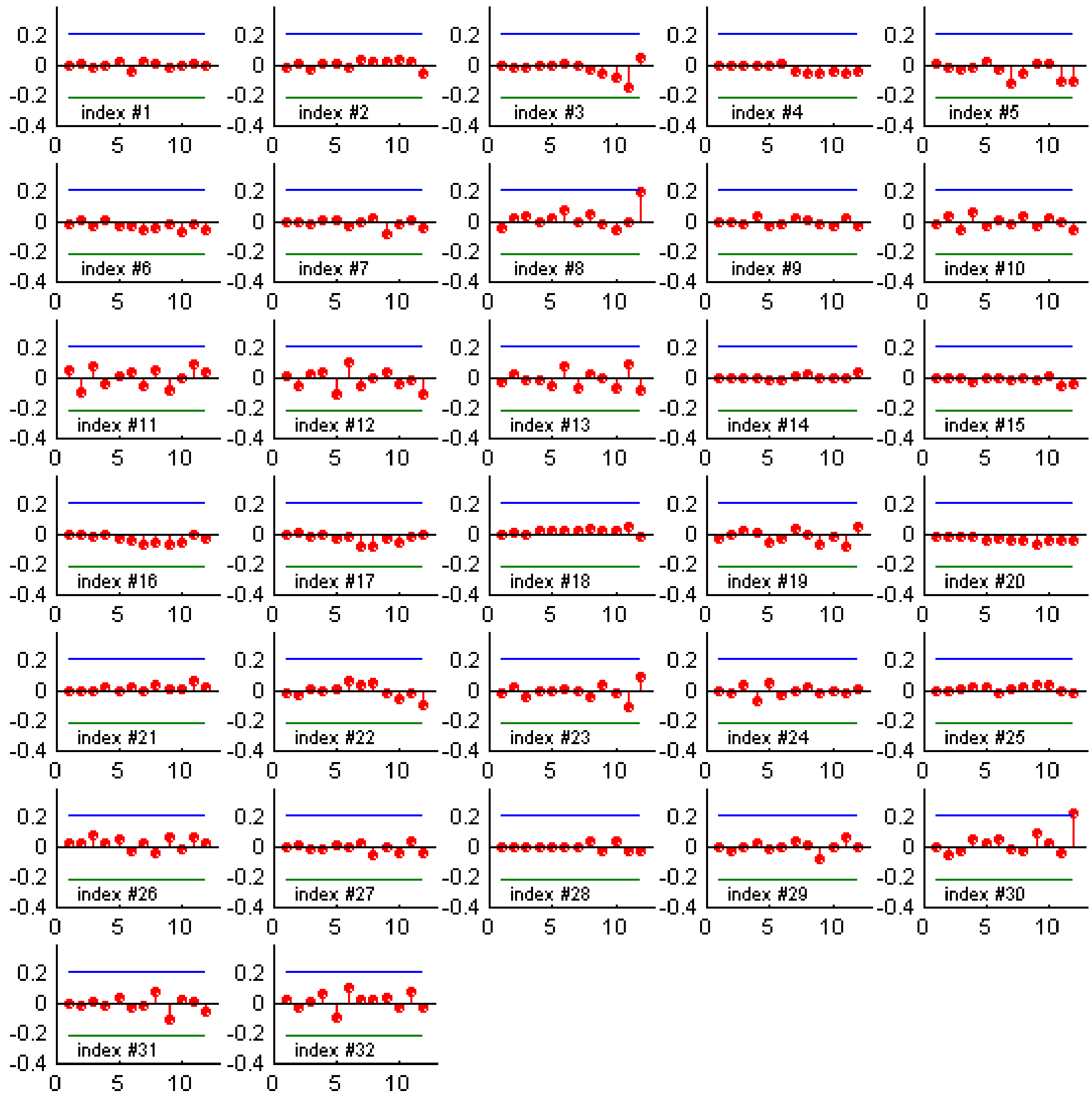

Figure 2: The autocorrelation functions for style indices are presented in this figure. The style indices used are *after* the adjustment for smoothing (η_t^J from (4)). The autocorrelations were computed for lags from 1 to 12. The thin horizontal lines around the horizontal axes represent 95% confidence intervals. Style index names can be retrived from table 1. For example, index #1 stands for Convertible Arbitrage index.

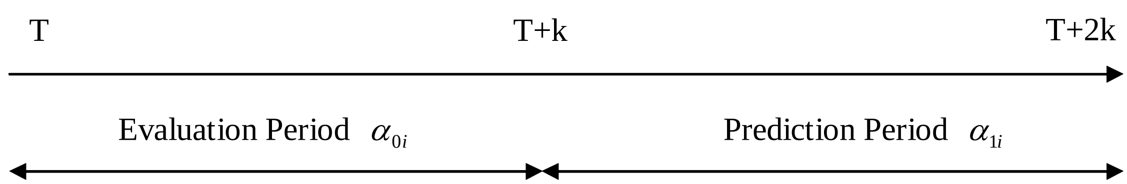

Figure 3: This diagram shows the timeline for the estimation of hedge fund alphas. In this paper k is equal to 36 months. That is evaluation and prediction periods are 3 years. The hypotheses is tested if alphas from the evaluation period can explain alphas from the prediction period.

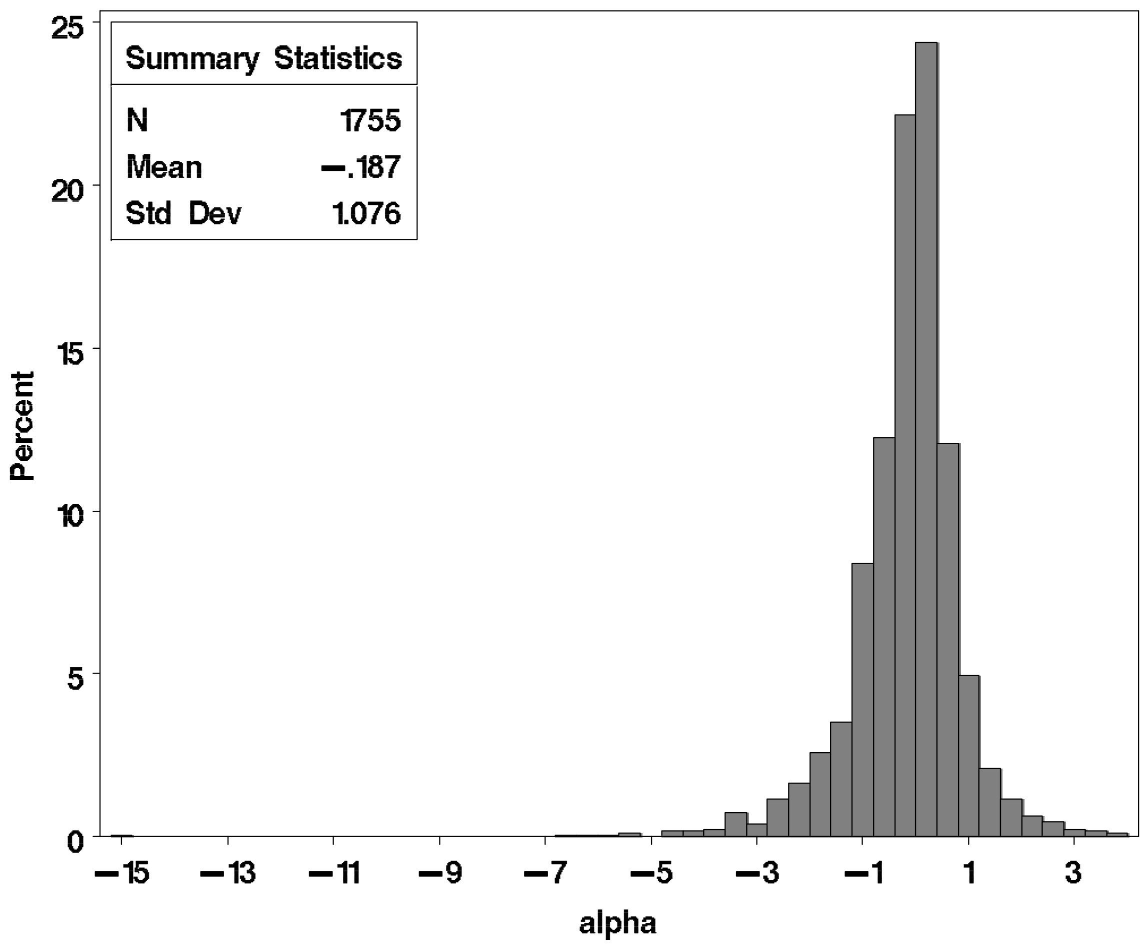

Figure 4: The distribution of α in model 3. Alphas are measured as monthly percentage returns and they are estimated over the period from May 1996 until April 2005.

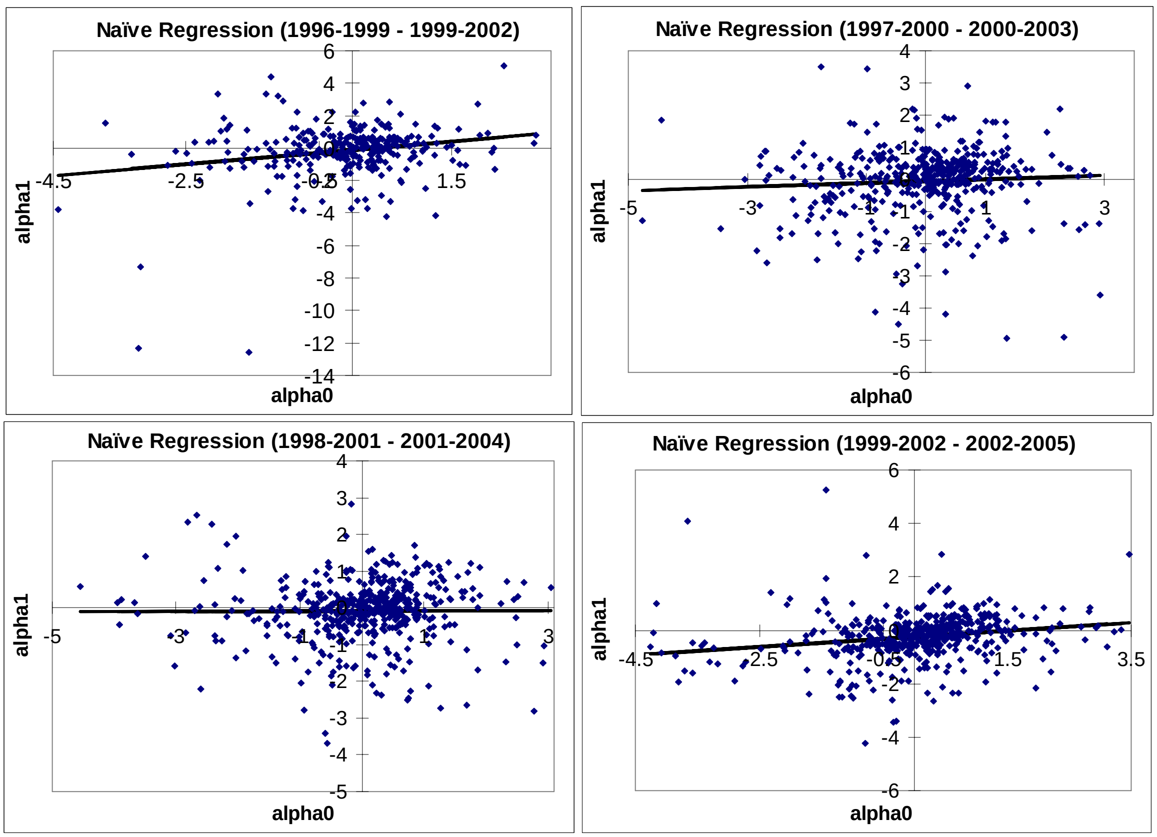

Figure 5: Scatter plots from the naive regression.

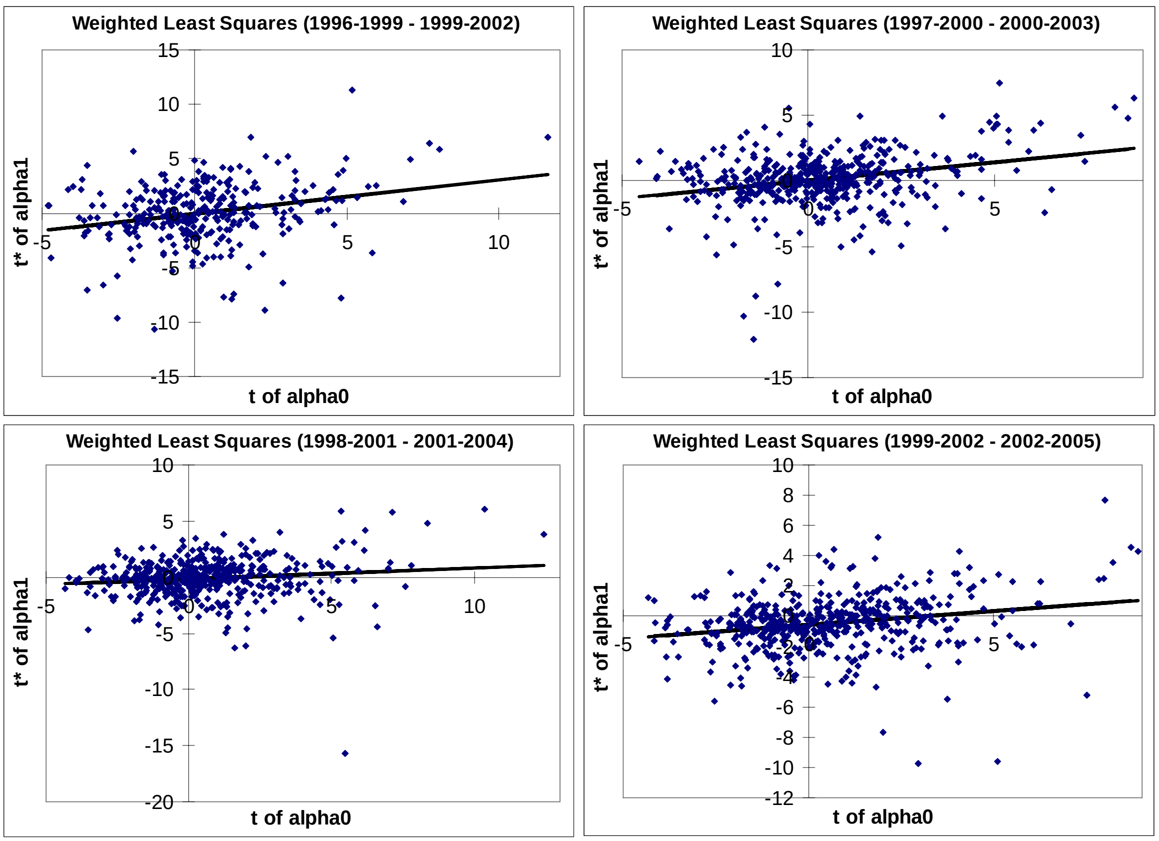

Figure 6: Scatter plots from the weighted least squares approach.

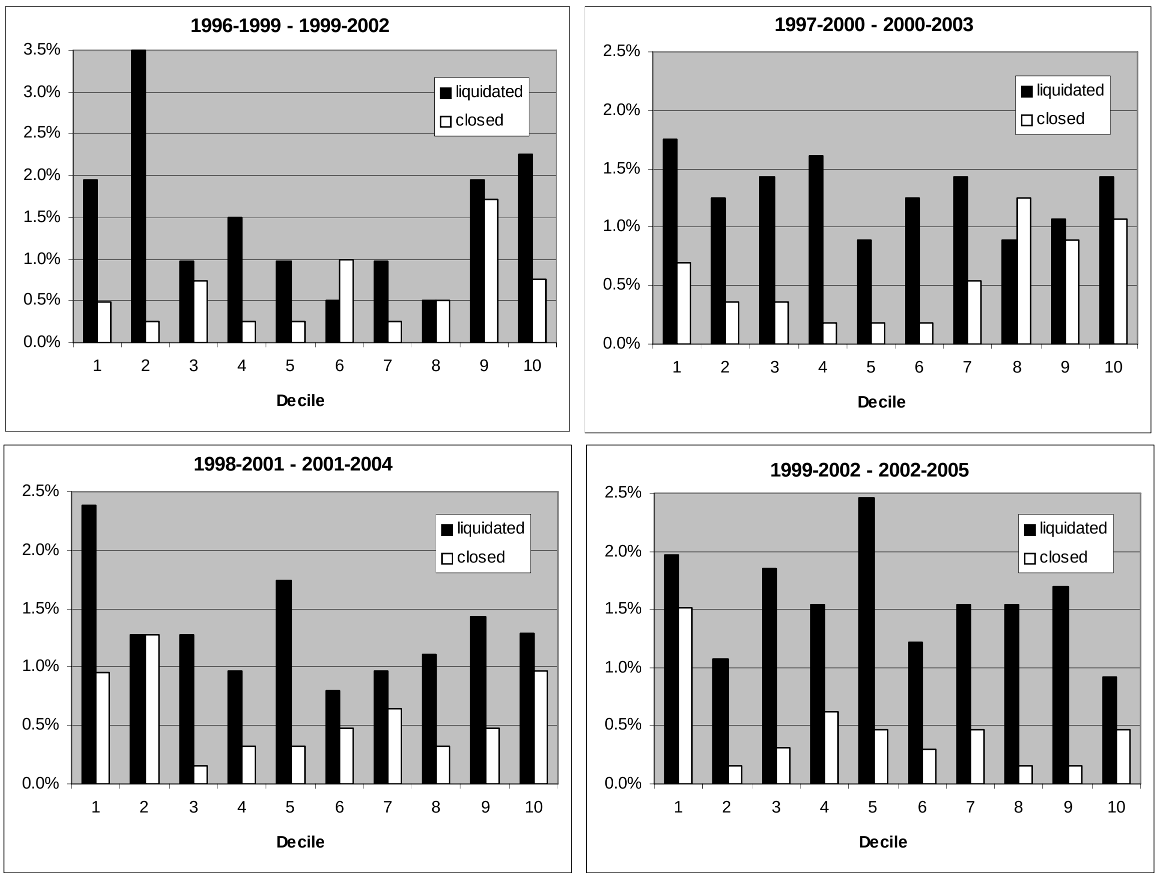

Figure 7: Histograms of distributions of liquidated and closed funds by deciles of α_0.

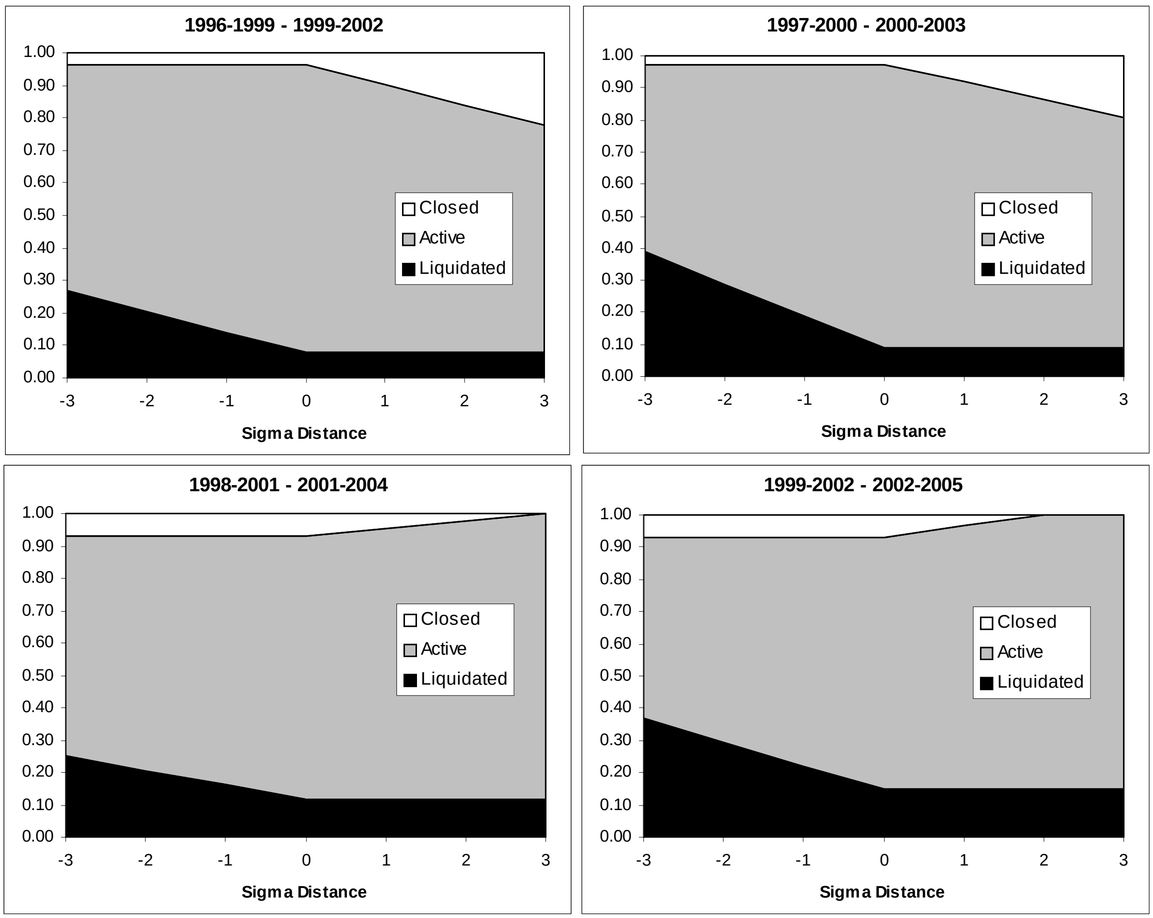

Figure 8: GMM estimated conditional probabilities of liquidation and closure with respect to α_0^*.

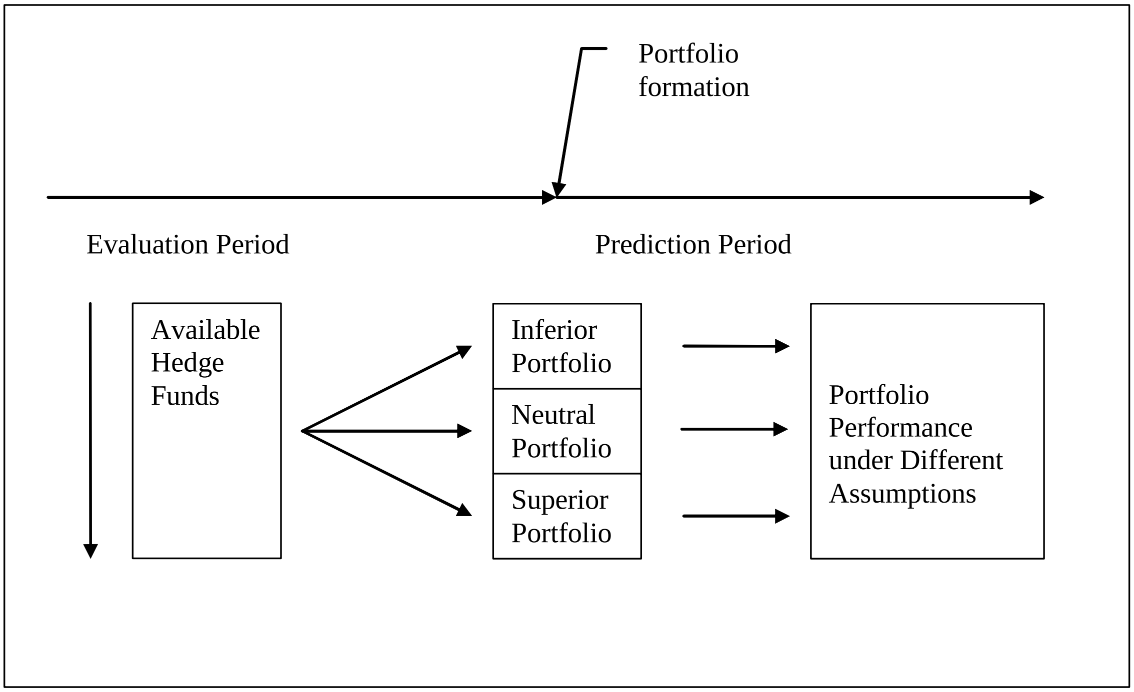

Figure 9: Test portfolio formation process.