CTA Performance Persistence: 1994-2010

Marat Molyboga
Co-Chief Risk Officer
Efficient Capital Management

www.efficientcapital.com

630 657 6800 T
630 657 6686 F

4355 Weaver Parkway
Warrenville, IL 60555

Abstract

This paper reports results of tests of the performance persistence hypothesis for Commodity Trading Advisors (CTAs). Using Fama-MacBeth regression and quintile analysis, we find that ranking CTAs using the t-statistic of alpha with respect to a CTA benchmark is predictive of future unleveraged returns. Sorting on the t-statistic of alpha yields around 4.6% annual spread of unleveraged returns between equally-weighted portfolios of the top and bottom quintiles. This finding is robust to the choice of CTA benchmark and model parameters. We examined the impact of incubation and backfill bias on the above results by repeating the analysis after excluding the first 12 and 24 months of data for each fund. We find that while on average there is no impact on the relationship between previous rankings and future unleveraged returns, and on persistence of worst performing funds, the identified strong persistence of the best performing funds is potentially solely driven by the incubation and backfill biases. We use Chi-square and Fisher tests to confirm that the worst performing funds have a significantly higher probability of liquidation than those of the other quintiles, and the top performing funds have a higher conditional probability of staying top performers versus becoming worst performers than that of the worst performing funds.

I am grateful to Sol Waksman for providing CTA data and professor Eugene Fama for his helpful comments. I also thank Keli Han and Zhongjin Yang for their valuable suggestions.

I. Introduction

There has been an increasing interest in study on the performance of Commodity Trading Advisors ("CTAs") since the late 1980s. Billingsley and Chance (1996), and Edwards and Park (1996) showed that CTA funds can add value to traditional portfolios of stocks and bonds in mean-variance framework. Schneeweis, Savanayana and McCarthy (1991), Schneeweis (1996) showed that managed futures add more value to traditional portfolios of stocks and bonds than hedge funds do. Kat (2004) showed that besides improving risk-adjusted returns better than hedge funds, managed futures also enhance the parameters of a portfolio's skewness and kurtosis. Fung and Hsieh (1997) constructed a CTA style factor index that persistently had a positive return when the S&P 500 index had a negative return. Worthington (2001) identified that between 1990 and 1998 the correlation of managed futures to the S&P 500 was 0.33 during the best 30 months and - 0.25 during the worst 30 months of the S&P 500 index. Fung and Hsieh (2002) showed that CTAs' impact on a traditional portfolio is similar to that of a lookback call and lookback put. Edwards and Caglayan (2001) found that CTAs provide better downside protection and higher returns along with negative correlation during bear markets than hedge funds. Schneeweis and Georgiev (2002) showed that CTAs add value to traditional portfolios especially during bear markets. However, during bull markets CTAs' performance is typically inferior to hedge funds, as reported by Georgiev (2001).

Using mean returns and modified Sharpe ratio to rank funds, Elton, Gruber, and Rentzler (1990) reported that performance persistence was statistically insignificant. Schwager (1996) ranked funds based on return/risk for funds with positive return and and returns for funds with negative return. He concluded that there is little evidence that top performing funds can be predicted. McCarthy, Schneeweis and Spurgin (1997) analyzed CTAs' returns and found that there is some persistence in performance if CTA returns are adjusted for market risk.

Brorsen and Townsend (2002) found that although persistence is stronger if return/risk measures of performance and long time series of data are used, it is still weak relative to the noise in the data. Capocci (2004) adopted Carhart's decile methodology to test CTA persistence. He ranked all funds based on their performance during in-sample period, divided them into deciles, created weighted index portfolios for each decile and tracked their out-of-sample performance. Capocci detected significant persistence in badly performing CTAs and weak persistence in well performing CTAs.

Schneeweis, Spurgin, and McCarthy (1996) examined survivorship bias in CTA returns. Schneeweis and Spurgin (1999) further analyzed performance of dissolved CTAs and concluded that dissolved CTAs begin underperforming 18-24 months before dissolution. Diz (1999) concludes that ignoring survival issues in the selection of managed futures programs results in significant reduction of performance in the range of 4.2-4.7 percent per year. Capocci (2004) reported that dead funds significantly underperform existing ones, and dissolution frequencies can reach 60 percent. Gregoriou, Hubner, Papageorgiou, and Rouah (2005) discovered that CTA survivorship is highly contingent on the fund strategy and that low assets under management, poor returns and high-risk exposure hasten CTA mortality.

In this study we test the performance persistence hypothesis of Commodity Trading Advisors. First, we discuss the CTAs' data we used and the cleaning procedures we performed in Section II. Then we describe the methodology that includes Fama-MacBeth regression and quintile methodology in Section III. We present empirical results in Section IV. We test whether the t-statistic of alpha is predictive of future performance. Since previous studies have reported persistence in performance of worst funds and lack of persistence in best funds, we additionally test hypothesis of persistence in performance of top and bottom funds. We also test whether relative performance is predictive of future attrition rates. Finally, we conclude our paper with the summary of our findings in Section V.

II. Data

There are six commonly used CTA databases: Barclay, CISDM (formely the MAR database), TASS, ITR, Stark, and Autumn Gold. The CISDM database was one of the first databases that began tracking CTA data in 1979. It currently includes data for over 500 active CTAs. International Traders Research (ITR) has been providing CTA data since 1996; it currently includes over 500 active programs and approximately 900 defunct funds. Autumn Gold currently includes 428 active programs. The Stark CTA database contains around 500 CTA programs. TASS database reports data for 628 active CTAs and 1842 defunct funds. Barclay Trading Group includes the largest list of active and defunct traders. From a statistical point of view, the larger the database, the more accurate the results that can be obtained. The current study uses the Barclay database containing 3912 funds, including 1126 active CTAs and 2786 defunct funds.

In order to get accurate results on persistence in performance of CTAs, the database had to be examined for data errors and biases. First, we excluded all data prior to 1994 as the number of available CTAs was too small. Then we excluded all multi-advisors and benchmarks as the scope of this study is limited to individual funds. We excluded all funds that only reported gross returns to preserve comparability. We excluded all funds with assets under management below US $1 million as they would be too small for institutional investors and their returns seem the most noisy. Furthermore, we excluded all funds with abnormal monthly returns in excess of 100% and removed zero returns at the end of defunct fund return streams.

Additional data corrections were made for attrition rates research. As we know, managers voluntary report to a database because they are actively marketing their funds and looking to attract new investors. Consequently, CTAs often stop reporting when they are not looking to attract new investors. There are three categories of funds that stop reporting. The first category ('Liquidated funds') includes funds with returns that are insufficient to cover operational expenses because of either poor performance or inability to raise assets. Second category ('Closed funds') includes successful funds that are not interested in attracting more investors, either because they have already reached their capacity level, or because they have a good client network sufficient to reach that level. The third category ('Unknown') includes funds closed for reasons not related to performance (for example, owners decide to close the fund and retire, etc). Only traders from the first category should be considered defunct in the attrition rate research. Since the second and third categories contain successful traders, including them in the dissolution analysis as defunct funds artificially increases the attrition rate of the well-performing funds. In his study Capocci (2004) defined dissolved CTAs as all funds that stop reporting and, therefore, considers traders from all three categories as defunct CTAs. Consequently, he reported that the 10th decile funds (containing top CTAs) have a higher dissolution probability than CTAs from the 6th, 7th, 8th and 9th deciles. In order to identify traders from the second and third categories, we examined each CTA's returns, assets under management (AUM), and fund status. Appendix A contains the full description of this step. The filtered dataset contains 2595 funds, including 835 active CTAs and 1760 defunct funds (1417 liquidated, 134 closed and 209 with unknown status).

It is well known that CTA databases contain incubation and backfill biases due to the voluntary nature of self-reporting. To mitigate these biases, we perform our analysis excluding the first 12 months of the data for each fund as suggested in Kosowski et al (2005). To investigate potential impact of the incubation and backfill biases on persistence result, we repeat our analysis by excluding the first 24 months of the data and then once again by not excluding any data. Inclusion of defunct funds mitigates the survivorship bias.

We use three CTA benchmarks commonly used in the literature: TASS, Barclay and CISDM. Table 1 contains information about performance of the benchmarks over 1994-2010 period.

Table 1 • Performance of TASS, Barclay and CISDM CTA indices in 1994-2010

	TASS	Barclay	CISDM
Annualized Return	3.77%	2.86%	5.56%
Annualized Std Dev	9.66%	7.61%	8.66%
Sharpe Ratio	0.39	0.38	0.64

This table displays annualized excess returns, annualized standard deviation of excess returns, and the Sharpe ratio calculated for each benchmark over 1994-2010 period.

Schneeweis and Spurgin (1996), Schneeweis et al (2007) report results of comparative analysis of various CTA benchmarks along with their descriptions. We repeat our analysis for each benchmark, to ensure robustness of the results to the choice of a benchmark. Three month t-bill rate is used for calculation of excess returns.

III. Methodology

This study uses two methodologies: Fama-MacBeth regression and quintile analysis. Both techniques use rolling CTA ranking based on the t-statistic of alpha. In order to calculate it at time t, we regress the last k net-of-fee excess returns of a CTA r_τ^i on the corresponding excess returns of a CTA benchmark I_τ.

$$r_\tau^i = \alpha_t^i(k) + \beta_t^i(k) \cdot I_\tau + \varepsilon_\tau^i$$

Then we estimate the standard error of alpha $\sigma(\alpha)_t^i(k)$ and define standard t-statistic of alpha as $T_t^i(k) = \alpha_t^i(k) \big/ \sigma(\alpha)_t^i(k)$.

A. Fama-MacBeth Regression

Fama-MacBeth regression was introduced in Fama and MacBeth (1973). In this study it is used because of its superior benefits when working with panel data. At each point in time t, t-statistic of alpha $T_t^i(k)$ is calculated for each fund that doesn't have missing data during that period and meets minimum AUM requirements. Then future unleveraged returns over the next l months (as defined in appendix B) $R_{t+l}^i(k)$, are regressed against the corresponding values of the t-statistic of alpha:

$$R_{t+l}^i(k) = \alpha_t + \beta_t \cdot T_t^i(k) + \varepsilon_t^i$$

Values of α_t and β_t are recorded. Then the data window is shifted by l months and the estimation procedure is repeated. Finally, the slope coefficient $\hat{\beta}$ is estimated as the average of all slope coefficients

$$(1) \quad \hat{\beta} = \frac{1}{s}\sum_{m=1}^{s}\beta_m$$

along with the corresponding standard error $s(\beta)$ and its t-statistic:

$$(2) \quad t(\hat{\beta}) = \frac{\hat{\beta}}{s(\beta)/\sqrt{s}}$$

The value of $\hat{\beta}$ represents the average impact of the past t-statistic of alpha of a fund on its future unleveraged returns. The corresponding t-statistic $t(\hat{\beta})$ shows the statistical significance of that relationship.

While Fama-MacBeth regression indicates the average relationship between past rankings of funds with their future unleveraged returns, potentially it can be driven by a relatively small group of funds. Therefore, we complement our study with quintile analysis that explicitly focuses on persistence of funds within quintiles.

B. Quintile Analysis

While in this study Fama-MacBeth regression is calculated for a wide range of parameters *k* and *l* for all three CTA benchmarks, the scope of the quintile analysis in this study is limited to using the Barclay CTA Index for ranking funds and applying one set of parameters commonly used in the industry. The window length *k*, used for ranking funds, is assumed equal to 24 months and the frequency of rebalancing *l* is assumed equal to 12 months.

The quintile analysis is performed similarly to the octile methodology of Hendricks et al (1993) and decile methodology of Carhart (1997). On December of each year, t-statistic of alpha $T_t^i(k)$ is calculated for each fund that doesn't have missing data during the most recent 24 months and meets minimum AUM requirements. Quintiles are formed based on the ranking and their equally weighted unleveraged portfolios are built and tracked for the next 12 months. At that point the process of re-ranking funds, forming portfolios and tracking their performance is repeated. If a fund stops reporting during that period, its allocation is assumed to be re-invested in the risk free asset with zero excess return until the end of the year. Performance of each portfolio and transition probabilities are recorded.

Since database reporting is voluntary, most likely a liquidated fund doesn't record its last losing month. Therefore, the above assumption of re-investment in the risk-free asset results in the upward bias in performance of portfolios for each quintile. However, the difference of returns between top and bottom quintiles would be understated if the bottom portfolio contains more liquidated funds than the top portfolio.

IV. Empirical Results

A. Fama-MacBeth Regression

Results of Fama-MacBeth regression are strikingly similar regardless of the choice of the benchmark. Table 2 presents values of the slope coefficients and their t-statistics as defined in (1) and (2), calculated using our standard parameter set with the window length k, used for ranking funds, equal to 24 months and the frequency of rebalancing l equal to 12 months.

Table 2 • Betas for Fama-MacBeth regression with k = 24, l = 12

	TASS	Barclay	CISDM
Beta	1.27%	1.28%	1.17%
	(3.23)	(3.39)	(3.17)

This table presents values of betas calculated in the Fama-MacBeth regression using TASS, Barclay and CISDM CTA indices for CTA ranking. The t-statistics are in parentheses.

To get a sense of the economic significance of the result, let's consider two hypothetical funds with the t-statistics of alpha, calculated with respect to TASS CTA Index over the most recent 24 months, equal to +2 and -2. The difference in their next year's expected unleveraged returns would be equal to $5.08\% = (2-(-2)) \cdot 1.27\%$, which is substantial given the expected future volatility of 15%.

We repeated analysis for the range of the ranking window k between 12 and 60 months as well as the range of the rebalancing frequency l between 1 and 12 months. Our results are summarized in Table 3 for the case of Barclay CTA Index used as the benchmark for calculation of the t-statistic of alpha.

Table 3 • Values of Betas for Fama-MacBeth regression using Barclay CTA Index for CTA ranking

	1	3	6	12
12	1.52	1.40	1.31	1.32
	(4.53)	(3.96)	(5.23)	(4.03)
18	1.76	1.56	1.46	1.20
	(5.99)	(5.56)	(5.46)	(4.89)
24	1.40	1.20	1.22	1.27
	(4.66)	(4.02)	(3.48)	(3.23)
30	1.35	1.05	1.12	1.09
	(4.19)	(3.40)	(3.34)	(3.58)
36	1.17	1.03	1.04	1.04
	(3.49)	(3.23)	(2.84)	(2.77)
42	1.02	0.88	0.88	0.83
	(3.02)	(2.69)	(2.40)	(2.03)
48	0.94	0.73	0.90	0.88
	(2.73)	(2.16)	(2.27)	(2.40)
54	0.95	0.75	0.83	0.86
	(2.77)	(2.30)	(1.92)	(1.96)
60	0.83	0.64	0.71	0.73
	(2.33)	(1.69)	(1.27)	(1.48)

This table displays values of betas calculated for various values of the ranking window k, presented in the left column, as well as a range of values of the rebalancing frequency l, presented in the top row. The t-statistics are in parentheses.

Table 3 shows that the relationship between the past values of the t-statistic of alpha and future values of unleveraged returns is robust across a wide range of parameters of the ranking window and the rebalancing frequency. When ranking was performed using the TASS and the CISDM CTA Indices, results were very similar[1].

To investigate the potential impact of the incubation and backfill biases on persistence result, we repeated our analysis by excluding the first 24 months of the data and then once again by not excluding any data. While results seemed slightly weaker when the first 24 months of data were excluded and slightly stronger when no data were excluded, the overall impact of the incubation and backfill biases on persistence results as measured by the Fama-MacBeth slope coefficients and their corresponding t-statistics was insignificant[2].

1 *Results are available from the author upon request*

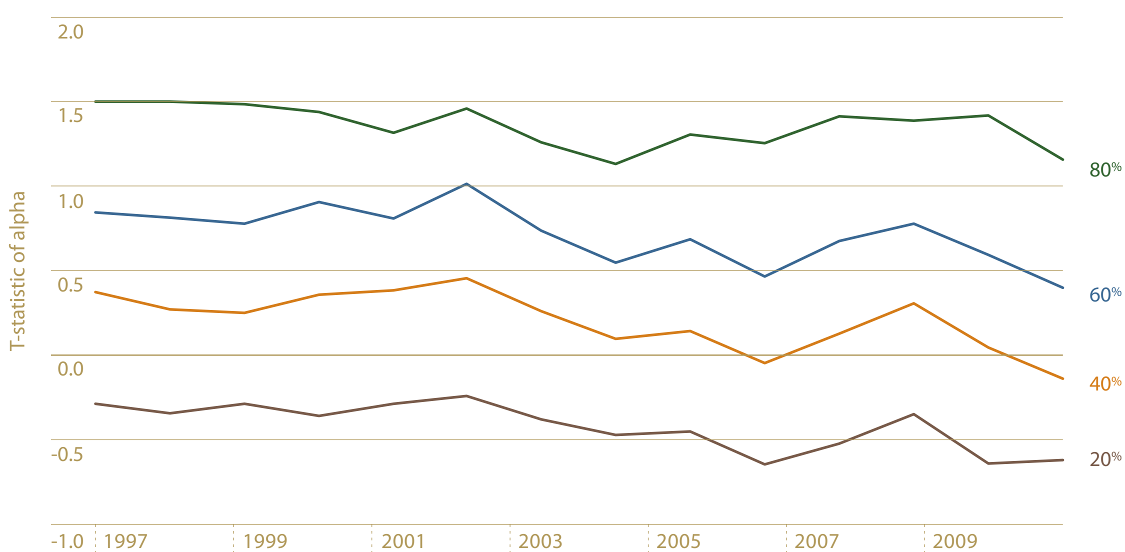
2 *Results are available from the author upon request*

Our empirical results confirm that performance is persistent on average for a wide range of parameters with negligible impact of the backfill and incubation biases or the choice of a benchmark. We further investigate whether that relationship is driven by the top performing funds, worst performing funds or average performers by applying quintile methodology.

B. Quintile Analysis

Each December all funds that have at least 24 months of data and at least US$ 1 million in assets under management are ranked using the t-statistic of alpha with respect to the Barclay CTA index. Figure 1 displays the values of the t-statistics of alpha that serve as the breakpoints of the quintiles. On average funds have positive alphas which can be explained by the choice of the Barclay CTA Index composition.

Figure 1 • Quintiles of the t-statistics of alpha



Each December all funds that have at least 24 months of data and at least US$ 1 million in assets under management are ranked using the t-statistic of alpha with respect to the Barclay CTA index. Figure 1 shows quintiles' breakpoints.

The number of funds in each quintile is presented in Table 4.

Table 4 • Number of funds in each quintile

Year	I	II	III	IV	V
1997	52	54	53	53	53
1998	54	55	54	55	54
1999	59	58	59	58	59
2000	58	59	58	59	58
2001	57	57	57	57	57
2002	63	62	63	62	63
2003	67	68	67	68	67
2004	72	72	71	72	72
2005	80	81	80	81	80
2006	86	85	86	85	86
2007	90	89	90	89	90
2008	99	98	99	98	99
2009	108	107	108	107	108
2010	115	116	115	116	115

Each December all funds that have at least 24 months of data and at least US$ 1 million in assets under management are ranked using the t-statistic of alpha with respect to the Barlcay CTA index. This table displays the number of funds in each quintile by year.

The number of funds in each quintile is sufficiently large and more than doubled throughout the period covered in the study.

Performance of equally-weighted portfolios is presented in Table 5.

Table 5 • Performance of equally-weighted quintile portfolios of funds

Portfolio	Annualized Excess Return	Annualized Std Dev	Sharpe
I (high)	6.44%	5.99%	1.08
II	4.02%	6.42%	0.63
III	4.62%	6.80%	0.68
IV	3.64%	6.95%	0.52
V (low)	1.84%	5.61%	0.33
I-V spread	4.61%	4.75%	0.97
	(3.63)		
IV-V spread	1.80%	3.46%	0.52
	(1.94)		
I-II spread	2.42%	3.20%	0.76
	(2.83)		
I-III spread	1.82%	3.88%	0.47
	(1.75)		

Each December all funds that have at least 24 months of data and at least US$ 1 million in assets under management are ranked using the t-statistic of alpha with respect to the Barlcay CTA index. Equally-weighted unleveraged portfolios are formed for each quintile and rebalanced annually. The t-statistics are in parentheses.

There is very strong evidence that funds from the top quintile outperform funds from the bottom quintile as represented by the difference of annualized return of 4.61% with the corresponding t-statistic of 3.61. The spreads between the top performers and the next two quintiles (I-II) and (I-III) as well as the spread between the bottom two quintiles (IV-V) also seem marginally significant as represented by the t-statistics.

To investigate the potential impact of the incubation and backfill biases on relative performance of quintile portfolios, we repeat our analysis by excluding the first 24 months of the data and then once again by not excluding any data.

Table 6 displays performance of quintile portfolios when first 24 months of data were excluded.

Table 6 • Performance of equally-weighted quintile portfolios of funds (24 months excluded)

Portfolio	Annualized Excess Return	Annualized Std Dev	Sharpe
I (high)	5.87%	6.21%	0.95
II	3.87%	6.93%	0.56
III	4.80%	7.25%	0.66
IV	4.14%	7.09%	0.58
V (low)	1.56%	5.84%	0.27
I-V spread	4.31%	4.88%	0.88
	(3.19)		
IV-V spread	2.58%	3.80%	0.68
	(2.45)		
I-II spread	2.00%	3.22%	0.62
	(2.24)		
I-III spread	1.07%	4.16%	0.26
	(0.93)		

To account for incubation and backfill biases, first 24 months of performance were excluded for each fund. Each December all funds that have at least 24 months of data and at least US$ 1 million in assets under management are ranked using the t-statistic of alpha with respect to the Barlcay CTA index. Equally-weighted unleveraged portfolios are formed for each quintile and rebalanced annually. The t-statistics are in parentheses.

The spread in performance between the top quintile and the other quintiles declined. The t-statistic of the spread between the first and the third quintiles (I-III) declined from 1.75 to 0.93 making outperformance statistically insignificant. On the contrary, the spread between the bottom two quintiles (IV-V) widened with the corresponding t-statistic increasing from 1.94 to 2.45.

Table 7 displays performance of quintile portfolios when there was no exclusion of data to account for the incubation and backfill biases.

Table 7 • Performance of equally-weighted quintile portfolios of funds without accounting for incubation and backfill biases

Portfolio	Annualized Excess Return	Annualized Std Dev	Sharpe
I (high)	7.05%	5.88%	1.20
II	4.07%	6.23%	0.65
III	4.56%	6.39%	0.71
IV	3.45%	6.69%	0.52
V (low)	2.10%	5.60%	0.37
I-V spread	4.96%	4.61%	1.07
	(4.02)		
IV-V spread	1.35%	3.07%	0.44
	(1.65)		
I-II spread	2.98%	3.05%	0.98
	(3.66)		
I-III spread	2.49%	3.75%	0.66
	(2.48)		

Each December all funds that have at least 24 months of data and at least US$ 1 million in assets under management are ranked using the t-statistic of alpha with respect to the Barclay CTA index. Equally-weighted unleveraged portfolios are formed for each quintile and rebalanced annually. The t-statistics are in parentheses.

The spread in performance between the top quintile and the other quintiles increased substantially. The t-statistic of the spread between the first and the third quintiles (I-III) increased to 2.48 making outperformance statistically significant. On the contrary, the spread between the bottom two quintiles (IV-V) declined with the corresponding t-statistic decreasing from 1.94 to 1.65.

The results are striking. Not properly adjusting for the incubation and the backfill biases could significantly overstate relative performance of the funds from the top quintile and understate relative underperformance of the worst performers. However, the spreads between the top performers and the bottom performers as well as the spreads between the fourth and fifth quintiles are consistently significant in all three cases. Our results are consistent with those reported previously suggesting strong persistence of worst performing funds and weak persistence of the top performers.

We further examined transition probabilities. Table 8 displays estimated transition probabilities.

Table 8 • Estimated Transitional Probabilities

	I	II	III	IV	V	Liquidated	Unknown	Closed
I	24.27%	17.56%	19.55%	17.00%	14.35%	4.63%	1.13%	1.51%
II	18.19%	17.81%	19.04%	17.91%	18.38%	5.47%	2.26%	0.94%
III	17.09%	18.51%	18.13%	18.04%	17.00%	7.74%	3.12%	0.38%
IV	13.68%	19.25%	18.40%	19.15%	17.26%	9.53%	1.98%	0.75%
V	14.53%	14.43%	12.26%	15.47%	20.75%	19.72%	2.17%	0.66%

Each row represents 5 original states (quintiles I-V) calculated during portfolio formation. Each column represents 8 future states: quintiles I-V, 'Liquidated' funds that stopped reporting due to bad performance, 'Closed' funds that stopped reporting due to lack of interest in attracting new investors and 'Unknown' funds that stopped reporting for an unknown reason.

There is a very definite pattern of increasing attrition rates with decline in relative performance. For example, the worst ranked funds have a probability of 19.72% to liquidate during the next 12 months, while funds from the top quintile have the attrition rate of only 4.63%.

We test the hypothesis of the funds from the bottom quintile V having the same attrition rate as the funds from the next worst quintile IV[3] by focusing on two future states: staying in business (by combining states of future rank I-V) or liquidating during the 12 months following the portfolio formation. The Chi-square one-tailed test gives the t-statistic value of 6.64 which yields the p-value of 0. The Fisher's exact test also gives the p-value of 0. Thus, both tests reject the hypothesis that funds from quintiles IV and V have the same attrition rates which means that funds from the lowest quintile have statistically higher probability to liquidate than funds from the other four quintiles.

3 *If we reject that hypothesis, we would also reject the hypothesis that the attrition rate for funds from quintiles I, II and III have the same attrition rate as that of the funds from quintile V.*

We perform an additional test of persistence focusing on two states: quintile I and quintile V. We test the hypothesis whether a fund from quintile I has the same conditional probability of staying in quintile I over the next 12 months as a fund from quintile 5. The Chi-square one-tailed test gives the t-statistic value of 6.06 which yields the p-value of 0. The Fisher's exact test also yields zero p-value. Thus, both tests reject the hypothesis that funds quintiles I and V have the same conditional probabilities of transitioning to quintile I versus transitioning to quintile V during the next 12 months.

To investigate the potential impact of the incubation and backfill biases on the above results, we repeated our analysis by excluding the first 24 months of the data and then once again by not excluding any data. In both cases we got similar results confirming that funds from the bottom quintile have a higher probability to liquidate than funds from the other four quintiles and funds from quintile I have higher conditional probability of transitioning to quintile I versus transitioning to quintile V during the next 12 months than funds from quintile V.

V. Concluding Remarks

We have examined returns of 835 active CTAs and 1760 defunct funds from the Barclay database, the largest publicly available single source of CTA data, for 1994-2010 period. After adjusting the data for the survivorship, incubation and backfill biases, we find that ranking funds using the t-statistic of alpha is predictive of future unleveraged returns. Sorting on the t-statistic of alpha yields a 4.6% annual spread of unleveraged returns between equally-weighted portfolios of the top and bottom quintiles. This finding is robust to the choice of CTA benchmark and model parameters.

We additionally investigated the impact of incubation and backfill biases on the performance persistence results by repeating the analysis after excluding the first 12 and 24 months of data for each fund. We find no impact on average relationship between previous rankings and future unleveraged returns, and on persistence of the worst performing funds. However, we find that the identified strong persistence of the best performing funds can potentially be driven solely by not appropriately accounting for the incubation and backfill biases.

We further use the Chi-square and Fisher tests to confirm that the worst performing funds have a significantly higher probability of liquidation than those of the other quintiles, and the top performing funds have a higher conditional probability of staying top performers versus becoming worst performers than that of the worst performing funds.

Appendix A

Fund Status

There are three categories of funds that stop reporting. The first category ('Liquidated funds') includes funds with returns that are insufficient to cover operational expenses because of either poor performance or inability to raise assets. Second category ('Closed funds') includes successful funds that are not interested in attracting more investors, either because they have already reached their capacity level, or because they have a good client network sufficient to reach that level. The third category ('Unknown') includes funds closed for reasons not related to performance (for example, owners decide to close the fund and retire, etc). Barclay database provided reason for discontinued reporting for only 403 funds, 36 of which were 'Closed' and 367 were 'Liquidated'. We tried to make some reasonable assumptions to categorize the remaining 1,357 funds with uncertain status. First, we assigned 'Closed' status to 98 funds that had Sharpe ratio greater than 1 with AUM exceeding US$ 10 million and length of drawdown below 6 months because we assumed they stopped reporting due to lack of interest in attracting more investors. Second, we assigned 'Liquidated' status to 1050 funds that stopped reporting either while being in drawdown for over 24 months, or at the depth of drawdown in excess of their annual volatility, or AUM below US $5 million or with track record shorter than 12 months.

After performing cleaning procedures, we had 835 active CTAs, 1,417 funds with 'Liquidated' status, 134 funds with 'Closed' status and 209 with 'Unknown' status.

Appendix B

Unleveraged Returns

The concept of leverage can be illustrated in a simple example. Consider CTA A with expected annual return of 20% and expected annual standard deviation of 10%. If an investor's risk appetite, measured in terms of expected annual standard deviation (or volatility), is equal to 15%, then the investor can request that the CTA increase its position size by 50%. For the investor, leveraged A's expected annual return becomes 30% and the expected annual standard deviation is 15%. Consider two funds A and B. The annual return of A is 20% and its annual standard deviation is 10%. The annual return of fund B is 25% and its annual standard deviation is 30%. The investor can use leverage to scale both CTAs to 15% volatility. Leveraged A has annual return of 30% and annual standard deviation of 15%, leveraged B has annual return of 12.5% and annual standard deviation of 15%.

This approach is commonly used by practitioners in the managed futures industry. However, it has a limitation. If a CTA's volatility is very low, the leverage coefficient may be too high to scale a CTA to a target volatility level because of margin requirement constraints. If a CTA's volatility is below a certain level, that assigned minimum level should be used for leverage calculation.

Without loss of generality, we define the un-leverage factor as:

$$\lambda_t^i(k) = \min\left[\lambda_{\max}, {}^{TVol}\!\big/_{Vol_t^i(k)}\right]$$

where $TVol$ is the target volatility, $Vol_t^i(k)$ is the annualized standard deviation of the CTA *i* calculated at point *t* using *k* most recent monthly returns, and $\lambda_{\max}$ is the maximum un-leverage factor. In this study $TVol$ is considered 15% and $\lambda_{\max}$ is equal to 3.

References

Billingsley, R., and D.M. Chance. "Benefits and limitations of diversification among commodity trading advisors." *Journal of Portfolio Management*, Vol. 23, No. 1 (1996), pp. 65-80.

Brorsen, B.W., and J.P. Townsend. "Performance Persistence for Managed Futures." *Journal of Alternative Investments*, Vol. 4, No. 4 (2002), pp. 57-61

Carhart, M.M. "On Persistence of Mutual Fund Performance." *Journal of Finance*, Vol. 52, No. 1 (1997), pp. 57-82

Capocci D. "CTA Performance, Survivorship Bias and Dissolution Frequencies", *Commodity Trading Advisors: Risk, Performance Analysis, and Selection*, Wiley Finance, Chapter 4 (2004), pp. 49-78.

Diz F. (1999) "How Do CTAs' Return Distributions Characteristics Affect Their Likely-hood of Survival?" *Journal of Alternative Investments*, Vol. 2, No. 2 (1999), pp. 37-41.

Edwards, F.R., and M.O. Caglayan. "Hedge Fund and Commodity Fund Investment Styles in Bull and Bear Markets." *Journal of Portfolio Management*, Vol. 27, No. 4 (2001), pp. 97-108.

Edwards, F.R., and J.M. Park . "Do Managed Futures Make Good Investments?" *Journal of Futures Markets*, Vol. 8, No. 5 (1996), pp. 589-616

Elton, E., M. Gruber, and J. Rentzler. "The Performance of Publicly Offered Commodity Funds." *Financial Analysts Journal*, Vol. 46, No. 4 (1990), pp. 23-30.

Fama E.F., and J.D. MacBeth. "Risk, Return, and Equilibrium: Emprical Tests." *Journal of Political Economy*, Vol. 81, No. 3 (1973), pp. 607-636.

Fung, W. and D.A. Hsieh. "Survivorship Bias and Investment Style in the Returns of CTAs: The Information Content of Performance Track Records." *Journal of Portfolio Management*, Vol. 24, No. 1 (1997), pp. 30-41.

Fung, W., and D.A. Hsieh. "Asset –based Hedge Fund styles and portfolio diversification." F*inancial Analyst Journal*, Vol. 58, No. 5 (2002), pp. 16-27.

Georgiev, G. "Benefits of Commodity Investments." *Journal of Alternative Investments*, Vol. 4, No. 1 (2001), pp. 40-48.

Gregoriou, G.N., G. Hubner, N. Papageorgiou, and F. Rouah. "Survival of Commodity Trading Advisors: 1990-2003." *Journal of Futures Markets*, Vol. 25, No. 8 (2005), pp. 795-816.

Hendricks, D., J. Patel, and R. Zeckhauser. “Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance”, *Journal of Finance*, Vol. 48, No. 1 (1993), pp. 93-130.

Kat, H.M. “Managed Futures and Hedge Funds: A Match Made in Heaven”, *Journal of Investment Management*, Vol. 2, No. 1 (2004), pp. 1-9.

Kosowski, R., N.Y. Naik, M. Teo. “Do hedge funds deliver alpha? A Bayesian and bootstrap analysis”, *Journal of Financial Economics,* Vol. 84, No. 1, pp 229-264.

McCarthy D., T. Schneeweis, and R. Spurgin. “Informational content in historical CTA performance.” *Journal of Futures Markets*, Vol. 17, No. 3 (1997), pp. 317-339.

Schneeweis, T., U. Savayana, and D. McCarthy. “Alternative Commodity Trading Vehicles: A Performance Analysis.” *Journal of Futures Markets,* Vol. 11, No. 4 (1991), pp. 475-490.

Schneeweis, T., R. Spurgin. “A Study of Survival: Commodity Trading Advisors, 1988-1996.” *Journal of Alternative Investments*, Vol. 2, No. 3 (1999), pp. 16-22.

Schneeweis, T., R. Spurgin, and D. McCarthy. “Survivor Bias in Commodity Trading Advisor Performance.” *Journal of Futures Markets*, Vol. 16, No. 7 (1996), pp. 757-772.

Schneewies, T., and G. Georgiev. “The Benefits of Managed Futures.” *CISDM Working Paper*, University of Massachusetts, Amherst, MA (2002).

Schneewies, T. “The Benefits of Managed Futures”. European Managed Futures Association, London (1996).

Schneeweis T. and R. Spurgin “Comparison of Commodity and Managed Futures Benchmark Indices.” CISDM Working Paper *Series*, University of Massachusetts, Amherst, MA (1996).

Schneeweis T., R. Gupta, and J. Remillard “CTA/Managed Futures Strategy Benchmarks Performance and Review.” *CISDM Working Paper*, University of Massachusetts, Amherst, MA (2007).

Schwager, J.D. “Managed Trading: The Myths and Truths.” New York: John Wiley & Sons Inc (1996).

Worthington, R.L. “Alternative Investments and the Semi-affluent Investor.” Research Report, Undiscovered Managers, Dallas, TX (2001).

Please Be Advised

Notes & Disclaimers

This document does not constitute a solicitation in any jurisdiction in which such a solicitation is unlawful or to any person to whom it is unlawful. Moreover, this document neither constitutes an offer to enter into an investment agreement with the recipient of this document nor an invitation to respond to the document by making an offer to enter into an investment agreement. Opinions expressed are our current opinions only.

This material is furnished by Efficient Capital Management, LLC ("Efficient"). This material is for information purposes only, is intended for your use only and is not an offer or a solicitation to subscribe for or purchase any product. This material is not intended to provide a sufficient basis on which to make an investment decision.

Benchmark data was obtained from various internal and external sources, such as CTA databases, Bloomberg, International Traders Research Inc., Hedge Fund Research, and Newedge. Efficient believes the benchmark data to be reliable, but can make no warranty as to its accuracy. Efficient has not and cannot verify the accuracy of all such information and the recipient should be aware that the information is presented on an informational "as-is" basis and is subject to change without notice.