YALE
INTERNATIONAL
CENTER
FOR FINANCE

Yale ICF Working Paper No. 08-21

October 2008

Fooling Some of the People All of the Time: The Inefficient Performance and Persistence of Commodity Trading Advisors

Geetesh Bhardwaj
AIG Financial Products; Rutgers University - Department of Economics

Gary B. Gorton
Yale School of Management; National Bureau of Economic Research (NBER)

K. Geert Rouwenhorst
Yale School of Management, International Center for Finance

Fooling Some of the People All of the Time: The Inefficient Performance and Persistence of Commodity Trading Advisors*

Geetesh Bhardwaj
AIG Financial Products

Gary B. Gorton
Yale School of Management and NBER

K. Geert Rouwenhorst
Yale School of Management

October 4, 2008

Abstract

Investors face significant barriers in evaluating the performance of hedge funds and commodity trading advisors (CTAs). The only available performance data comes from voluntary reporting to private companies. Funds have incentives to strategically report to these companies, causing these data sets to be severely biased. And, because hedge funds use nonlinear, state-dependent, leveraged strategies, it has proven difficult to determine whether they add value relative to benchmarks. We focus on commodity trading advisors, a subset of hedge funds, and show that during the period 1994-2007 CTA excess returns to investors (i.e., net of fees) averaged 85 basis points per annum over US T-bills, which is insignificantly different from zero. We estimate that CTAs on average earned gross excess returns (i.e., before fees) of 5.4%, which implies that funds captured most of their performance through charging fees. Yet, even before fees we find that CTAs display no alpha relative to simple futures strategies that are in the public domain. We argue that CTAs appear to persist as an asset class despite their poor performance, because they face no market discipline based on credible information. Our evidence suggests that investors' experience of poor performance is not common knowledge.

* This paper has benefited from comments and suggestions from Martijn Cremers, Ned Elton, Bill Fung, Mila Getmansky, Will Goetzmann, Marty Gruber, Raj Gupta, David Hsieh, Jon Ingersoll, Bing Liang, Jonathan Macey, Roberta Romano, Ken Scott, and seminar participants at UMass Amherst. The views expressed are those of the authors and do not necessarily reflect the official position of AIG Financial Products Corp.

1. Introduction

Hedge funds command hefty fees because they allege that they can earn above average risk-adjusted returns, based on their skills. This means that the returns generated by their trading strategies must not be easily replicated by lower cost alternatives such as passive indices, mutual funds or ETFs. According to the Government Accountability Office (2008), based on industry estimates, the number of hedge funds has grown from 3,000 to more than 9,000 between 1998 and early 2007, and their assets under management have grown from $200 billion to more than $2 trillion globally. Investors appear to have concluded that these funds are worthwhile investments.

Are hedge funds worthwhile investments? Do they earn above average risk-adjusted returns? What benchmarks should be used for the risk-adjustment? How should investors determine which funds to invest in? It has proven very difficult to answer these questions, because it is difficult to obtain reliable performance data and to determine the relevant benchmarks. Hedge funds are prohibited from direct advertising, but are allowed to indirectly market themselves by reporting their past (possibly paper) returns to private vendors who then sell this performance information through databases to potential investors, news sources, consultants, and researchers. While past returns may be useful for investment choices, the available hedge fund databases are contaminated by a number of biases that affect the ability of investors to make proper inferences. The academic literature has recognized many of these biases (e.g. selection bias, survivor bias, and backfill bias), but the proposed adjustments are often crude or difficult to implement ex-post.

Even when the performance data is available, the issue remains of how to adjust the returns for risk. It is not clear what benchmarks hedge funds should be evaluated against because they are an extremely heterogeneous group and can employ time-varying, state-contingent, and leveraged strategies. In fact, it is not obvious that hedge funds form an "asset class" since their strategies are so diverse. All they have in common is that they have chosen to organize themselves so as to be exempt from various U.S. legal requirements, hence becoming "hedge funds" – a legal definition of the "asset class."[1] Faced with this heterogeneity problem, the literature on hedge funds is not so much performance analysis as it is a descriptive, positive, analysis of these funds' returns. The focus has been less on whether funds add value for investors than on empirically characterizing fund strategies.

A central point of our work is that biased data and a lack of benchmarks are problems faced by investors and researchers alike. We separate the question of whether fund managers exhibit skill from the question of whether investors receive positive risk-adjusted returns, by looking at both returns net of fees and estimated gross returns. To the extent that fund managers exhibit skill, we ask how the value added is divided between the funds and its investors. We narrow the set of funds to be evaluated to commodity trading advisors (CTAs).[2] There are four reasons for our choice. First, the strategies that CTAs employ are relatively well-known compared to many hedge fund strategies. CTAs report in surveys that they are trend followers and momentum traders. In a survey in 2000, 75 percent of CTAs responded that they are trend followers and 71 percent

[1] Hedge funds and CTAs are organized so as to qualify for exemptions from regulations, and disclosure requirements of certain federal securities laws, including the Securities Act of 1933 and the Securities Exchange Act of 1934. For purposes here, to qualify hedge funds must not advertise to the general public and can only solicit participation in the fund from certain large institutions and wealthy individuals. For details see Hall (2008).

[2] We use the term CTA (Commodity Trading Advisors) to refer to the legal form of investment vehicles that trades in futures markets and consequently registers with the U.S. Commodity Futures Trading Commission.

responded that they used momentum as a signal in their trading approach.[3] Second, the commodity and financial futures comprise a smaller strategy space for CTAs compared to equity hedge funds, event-driven funds, or multi-strategy hedge funds, for example. This simplifies the choice of benchmarks for evaluating CTA performance. Third, the available performance history of CTAs is relatively long. While most major hedge fund databases do not provide samples that are free from backfill bias prior to 1994, there exists an early academic literature on public CTAs in the 1980s in which this bias is effectively eliminated (Elton, Gruber, and Renzler (1987, 1989, 1990)). Finally, although CTAs are a subset of the hedge funds universe, they control a significant amount of assets. While there are no official measures of the size of the CTAs' money-under-management (MUM), BarclayHedge estimates that as of the end of 2007, MUM was $206.6 billion, having grown from $50.9 billion five years earlier – a 306 percent increase.[4]

We analyze the performance of all CTAs that voluntarily report to the Lipper-TASS database. To eliminate the influence of various biases induced by strategic returns reporting and database construction, more than 80% of the available observations are excluded.[5] We show that these corrections greatly influence inference about CTA performance. We estimate that between 1994 and 2007 the average bias-adjusted CTA returns after fees have been statistically indistinguishable from the average return on an investment in US T-bills. The average CTA has therefore not created value for their investors. This conclusion mirrors the finding by Elton, Gruber and Rentzler (1987, 1989, 1990) (EGR) who – almost two decades ago – found that publicly traded commodity funds did not create positive returns for investors. The combined evidence is therefore one of 20 years without performance. The surprising finding therefore is that the considerable attention that the Elton Gruber and Renzler studies received at the time of publication does not seem to have influenced the ability of CTAs to attract assets.

The poor net returns for investors are not necessarily inconsistent with CTA managers possessing skill. For example, it is possible that managers generate excess returns, but capture the rents of outperformance through charging fees. We present some evidence consistent with this view. Using standard procedures to estimate gross returns (i.e., returns before fees), we estimate that the average CTA return has exceeded T-bills by more than 5 percent per annum between 1994 and 2007, but only by 0.85 percent per annum after fees. In order to evaluate whether these gross excess returns are abnormal, we develop a number of simple performance benchmarks. We find that relative to these benchmarks CTAs display no significant skill (alpha). However, the benchmarks can explain relatively little of the variance of CTA returns. It is difficult to explain variation in ex-post gross returns of CTAs, despite the fact that the majority of funds describe their style as trend-following. A regression of individual fund returns on our benchmarks produces an R-squared below 30% for seven out of ten funds. We show that exposure to simple trend following strategies can explain the most of the average outperformance before fees.

The poor performance track record of CTAs raises the question of why the asset class has continued to grow – apparently despite a long history of poor performance. The supply side of the market is easy: CTAs generate fee income of about 4% on assets under management, which also explains the high rates of entry into a market with high attrition rates. Why investors continue to allocate to CTAs is more difficult to answer. Did investors ignore the conclusions of the EGR papers despite the publicity they received at that time?

---

[3] See Waksman (2000). Academic research is in agreement with these CTA self-assessments. Fung and Hsieh (1997) argue that CTAs have one dominant style factor, namely, trend following.

[4] See Hhttp://www.barclayhedge.com/research/indices/cta/Money_Under_Management.htmlH .

[5] As explained below, we exclude 83,201 of the 102,393 available monthly observations on fund performance post-1993, and all returns prior to 1994.

We explore several broad explanations. First, average fund performance may not be sufficient as an overall indicator of the attractiveness of an asset class. To the extent that CTAs offer option-like payoffs that exhibit positive skewness, investors may prefer to allocate to CTAs despite poor average returns. However, our data shows that CTAs are equally likely to exhibit positive or negative skewness. It seems unlikely that CTAs are attractive because of the portfolio properties of their performance. While correlations of managed futures programs with traditional asset classes have historically been low, it seems unlikely that investors would allocate $200 billion to an asset class that offers T-bill returns with a standard deviation that is comparable to equities.

An alternative explanation is that investors are unable to overcome the information asymmetry to properly evaluate CTA performance. Although it is difficult to provide direct evidence, several observations are consistent with this view. For example, when academic researchers do not seem to agree on how to properly adjust CTA track records for various biases introduced by strategic reporting, it seems unlikely that investors who often lack access to comprehensive databases can do a substantially better job. Especially since there is no mechanism to create common knowledge about historical CTA performance – it is difficult to learn from the investment experience of others when information is not aggregated, either through market prices, disclosure, or regulatory oversight. In this context it is illustrative that in response to the EGR studies in the 1980s which revealed poor performance of public commodity funds, the industry has reorganized itself into a form that requires less disclosure and regulatory oversight. And while in theory funds can attempt to signal quality through the contract terms they offer investors, we find no systematic relationship between contract terms and fund performance.

Finally, investors may simply be unaware that there is an information asymmetry and the history of poor CTA performance may not be common knowledge. Such an information setting differs from the failure of Akerlof's (1970) lemons market, in which it is common knowledge that there is an information asymmetry. It appears that CTAs strategically report their performance data to maintain this information environment. We discuss these issues towards the end of the paper. But, note that it puts researchers in a somewhat delicate position. Simply put, we do not have all the data we would like and the available data must be treated with great care, precisely because of the strategic desires of the CTAs. We alert the reader to these difficulties as we proceed.

The literature most directly related to our work is about CTAs. In addition to the Elton, Gruber and Rentzler (1987, 1989, 1990) papers, our work is closely related to Fung and Hsieh (1997, 2001). Fung and Hsieh (1997) argue that the dominant investment style of CTAs is trend following. Fung and Hsieh (2001) construct dynamic factor portfolios to capture this trend following behavior. We show that while the Fung-Hsieh (FH) factors are useful for style analysis, they are less useful for answering the question of whether CTAs create alpha.[6] In particular we show that that the FH factors tend to impound an upward bias in fund alphas, because they are inefficient replications of trend-following styles.

The paper proceeds as follows. In Section 2 we introduce the data set used for this study and briefly discuss the various, well-known, biases that exist in CTA and hedge fund data sets. In

[6] There is also a literature on hedge funds. For example, Ackerman, McEnally, and Ravenscroft (1999) analyze hedge funds, comparing hedge fund returns, volatility, and Sharpe ratios to the returns and characteristics of the S&P 500 and eight standard market indices. They conclude that hedge funds outperform mutual funds, but not standard market indices. Brown, Goetzmann and Ibbotson (1999) also look at hedge funds and find little evidence of outperformance. Brown and Goetzmann (2003) used a classification algorithm to group hedge funds into similar styles, which then becomes the benchmark for out-of-sample performance evaluation. There are many other papers (e.g., Brown and Goetzmann (1997)).

addition we construct a performance index for CTAs net of fees, and estimate their (gross) investment returns before fees. In Section 3 we discuss a variety of benchmarks to evaluate the style and performance of CTAs. Given the strategy space of CTAs, we focus is on simple futures based strategies in equity, commodity and currency markets that are in the public domain. We find that CTAs do not add value, in the sense of producing alpha relative to these benchmarks. In Section 4 we review the historical performance of commodity funds in light of the earlier work by Elton Gruber and Rentzler. In section 5 we explore explanations for why CTAs persist despite two decades of poor performance. Section 6 concludes.

2. Fund Performance Data

A CTA is a hedge fund which has registered to trade futures with the Commodity Futures Trading Commission. Like hedge funds, CTAs are essentially prohibited from advertising.[7] Faced with this restriction, a primary way to reach potential investors is for the hedge fund or CTA to voluntarily report performance information to private companies, data vendors, which then sell the data. Individual funds can release their own performance data, but not comparative data for advertising purposes.[8] For making an investment decision, comparing individual funds to other funds, the vendor data is the only publicly-available source of information for evaluating these funds. The data are purchased by the news media and published in a variety of locations, such as Barron's or ManagedFutures.com, for investors to observe.[9]

Because the decision to report performance data by CTAs and hedge funds is entirely voluntary, it introduces a strategic element in the reporting process. The resulting biases lead to an overstatement of performance of hedge funds, which contributes to the inference problem for investors and researchers alike. While many of the biases are well-known, there seems to be less agreement on how to handle these biases when evaluating hedge fund performance. Without reviewing the entire literature, we illustrate some of the major biases in the context of CTAs, and discuss why some attempts to adjust for the biases are suspect.

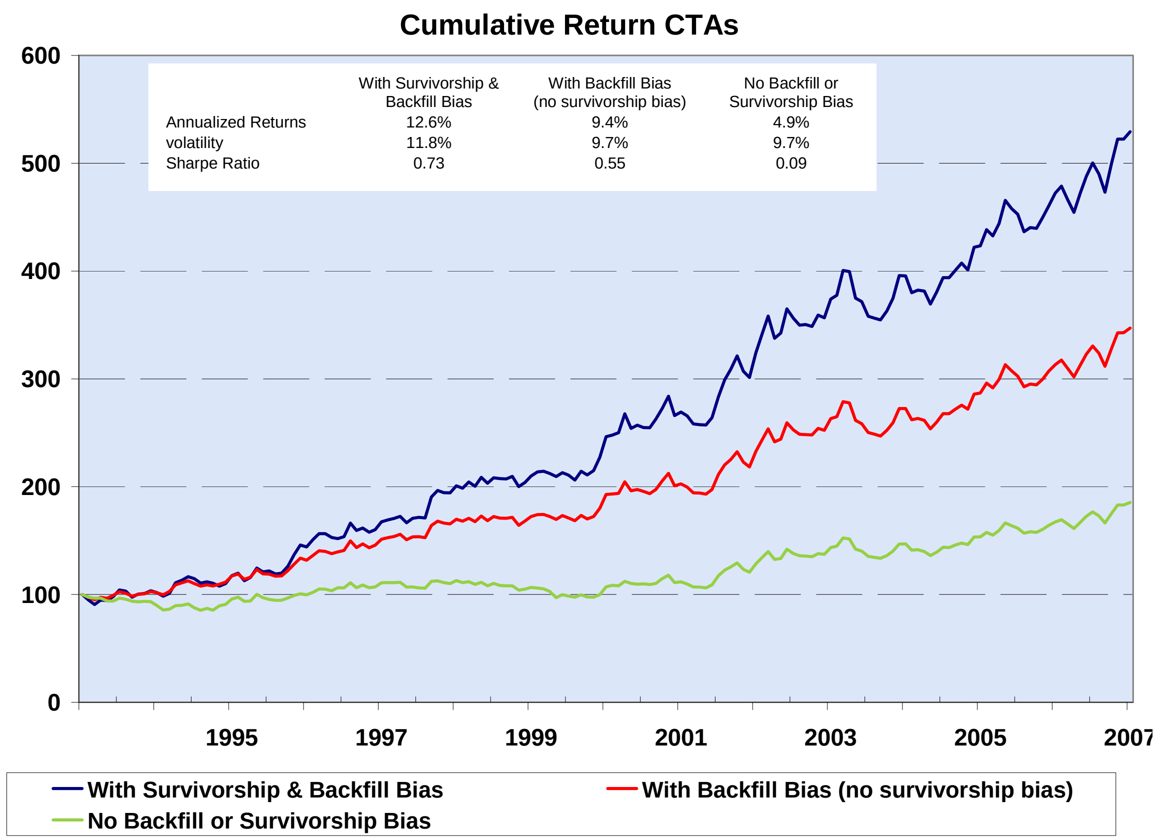
Consider a naïve investor who is contemplating an investment in CTAs and decides to examine the track record of all currently investable funds. In order to simplify the data collection process, the investor uses the Lipper-TASS database to calculate the average return to CTAs that are currently in existence, going back to 1994. The resulting performance series is given by the top line in Figure 1, which shows the cumulative total returns to an equally-weighted (EW) portfolio of CTAs over this period. The average return (net of fees) on this portfolio was 12.6 % which exceeds the return on T-bills which was about 4.0 % per annum over the 14-year period between 1994 and 2007. Our naïve investor might conclude that CTAs are an attractive investment: they provide an absolute return over T-bills which is significant economically (8.6% per annum) as well as in a statistical sense (t-stat = 2.73). However, this calculation does not correct for various

[7] The prohibition on advertising seems problematical with the internet. One need only type "commodity trading advisor" into Google to get a sense of what this means as a practical matter. There are 93,900 hits.
[8] Individual CTAs can publicly present performance data. The CFTC under Regulation 4.41(a) adopted "a rule that leaves to the discretion of the [CPO, CTA, or principal] advertising results –whether actual, simulated or hypothetical—the format of that presentation, so long as that format is not false, misleading or deceptive." See Federal Register Vol. 71, No, 163 (Wednesday, August 23, 2006), p. 49388.
[9] See the "Market Lab" section of Barron's which provides "Commodity Traders Advisors Performance." Barron's provides the current monthly return, year-to-date, 12-month return, 3-year return, and 5-year return, the 12-month annualized standard deviation, the 12-month maximum drawdown (%), and the assets under management. Not all of this available for every fund listed. The performance data comes from the CASAM CISM Database (formerly the MAR Database); see Hhttp://www.casamhedge.com/H .

biases in the database. Figure 1 previews our discussion in the remainder of this section that a correction for survivorship bias and backfill bias would lower the average return to CTAs by about 7.7% to 4.9% per annum, which is only 85 basis points above the average return to T-bills. The correct inference from the data ought to be that the average CTA does not offer absolute returns but merely adds risk.

Figure 1: Measures of CTA Performance

The figure shows the cumulative performance of an investment in an equal weighted portfolio of Commodity Trading Advisors that report to the Lipper-TASS Database. The portfolio labeled *With Survivorship Bias and Backfill Bias* consists of all Funds that were alive at the end of our sample. The portfolio labeled *With Backfill bias (no survivorship)* includes all monthly return observations in the "live" and "graveyard module" of the database. The portfolio *No Backfill or Survivorship Bias* includes only fund-returns after the first date of a fund reporting to the database.



2.1 Sources of Bias in Lipper-TASS

There are at least four sources of bias in the Lipper-TASS database:

Selection Bias

The selection bias stems from the strategic reporting decision by a fund. Funds that experience poor performance may decide not to report to the database. Funds that look to attract new investors are more likely to report, while successful funds may stop reporting to the database as their need to advertize may have diminished. This issue has been widely discussed in the literature.

Look-back Bias

Look-back bias refers to ex-post data withholding by a fund after observing performance. This can take several forms. For example, a fund is unlikely to not report the return(s) prior to liquidation due to poor performance. More generally it is likely that funds delay reporting poor returns. If performance improves subsequently, it may report the delayed returns, or alternatively drop out of the database when fund returns continue to be low. This option to withhold poor performance has been discussed in the literature. What seems to have gone unnoticed in the literature is that funds can ex-post remove their entire performance record from the database. Comparing two versions of the Lipper-TASS database, we find several instances where the entire track record of a fund disappears. Conversations with the vendor confirms that funds can indeed request to have their entire historical track record removed, based on the view that "reporting is entirely voluntary and at the discretion of the funds." This "look-back bias" affected about 2% of the CTAs between the October 2007 and April 2008 versions of the database. It seems plausible that unsuccessful funds have a larger incentive to remove their performance data ex-post, which would lead to an upward bias in the performance of the funds that remain in the database. Quantification of the magnitude of this bias would require a full record of these deletions, which is unfortunately unavailable.[10]

Survivorship Bias

The survivorship bias occurs when a fund disappears from the database after it dies. By focusing only on funds that are currently in existence, the naïve investor in our example excluded funds that were dissolved. Because the surviving funds have outperformed their peers, this leads to an upward bias. Malkiel (1995) estimated the size of this bias by comparing the (annualized) returns for the live funds (those funds that still exist at the end of the data sample) to the whole data set of returns (including funds that exited during the sample period).[11] Since 1994 Lipper-TASS has maintained a record of non-surviving funds in the "graveyard module" of the database. The top two lines in Figure 1 compare the average return of CTAs that were in existence at the end of 2007 to an equally-weighted performance of all funds in the "live" and "graveyard" modules of the database. Figure 1 illustrates that surviving funds have outperformed the average fund in the database by 3.2% (12.6% minus 9.4%) between 1994 and 2007.[12] When we discuss the next source of bias, induced by backfill we will include all CTAs from both the live and graveyard modules.

Backfill Bias

Also known as "instant history," backfill bias is created when funds are allowed to submit a performance history at the time of first reporting to the database. Because managers are more likely to report funds with a good history, and avoid reporting funds with poor histories, this

[10] A by-product of the look-back bias is that it makes it difficult to exactly replicate results of other researchers unless the exact same version of the database is used. Lipper-TASS only distributes the most recent version of the database to current subscribers.

[11] Fung and Hsieh (2000), Brown, Goetzmann, and Ibbotson (1999), Ackerman, McNally, and Ravenscraft (1999), and Liang (2000), among others, use this method. The estimates of the bias range from 3.0 percent (from Fung and Hsieh) to 0.2 percent (from Ackerman, McNally, and Ravenscraft). Malkiel and Saha (2005) report that the average difference between live hedge funds and defunct hedge funds is more than 830 basis points over the period 1996-2003.

[12] These calculations do not exclude backfilled returns.

creates an upward bias in the returns prior to the first live reporting date.[13] A comparison of the bottom two lines in Figure 1 illustrates the magnitude of this bias that results when "instant histories" of returns before the first reporting date are excluded.

The figure illustrates the wide difference between the average performance of funds when they report to the database in real time (4.9%) and the average performance of all funds including backfill (9.4%). The former return is lower because the average backfilled return of 11.3% considerably exceeds the "live" average return of 4.9%. The backfill bias in CTA returns mirrors the observation by Elton, Gruber and Rentzler, that publicly traded commodity funds in the 1980s generally failed to beat the historical performance reported in their prospectuses.

Early hedge fund studies starting with Park (1995) attempted to correct for this bias by excluding the first portion of the track record of each fund before calculating performance, typically a fixed number of months reflecting the estimated backfill for the "average" fund. This "*x*-month screen" is a crude measure that leads to overstatement of the measured returns of funds that have a longer backfill period than *x* months. Recent versions of the Lipper-TASS database contain a field for each fund indicating the date of first reporting to the database. The backfill bias can therefore simply be eliminated by discarding returns prior to the first reporting date. Perhaps surprisingly, many studies continue to apply the *x*-month screens to account for backfill bias.[14] The following table illustrates that *x*-month screens lead to very different conclusions about the magnitude of the backfill bias for CTAs.

Table 1: Backfill Bias and CTA Performance

The table gives the average return expressed as % per annum on an equally-weighted portfolio of CTAs between 1994 and 2007 using different screens for inclusion of funds in the portfolio. *Backfill not removed* includes all funds and months for which data are available in Lipper-TASS. *Backfill removed* only uses firm-month observations for funds after their first reporting date to the database. The *x-month screen* removes the first x months from the performance record of a fund before it enters the portfolio.

EW CTA Index	Average Return (% p.a.)
Backfill removed (first reporting date)	4.9
Backfill not removed	9.4
12-month screen	8.3
24-month screen	7.8
36-month screen	7.7

Applying a 12-month screen across all funds lowers the average CTA return by only 1.1 % per annum, as compared to 4.5% using the first day of reporting as a screen. Longer screens lower average returns but not to the extent of eliminating returns prior to the first live reporting date. The reason is that there is a great deal of dispersion in the number of backfilled returns across funds. In the Lipper-TASS data set, the average number of backfilled months for all hedge funds is 28 with a standard deviation of 33.86. For CTAs the average number of backfilled months is 43 with a standard deviation of 47.41. This explains why even a conservative 36-month screen is not sufficient to eliminate the backfill bias. Throughout our analysis we will only use funds for which

[13] Several papers have quantified this bias, including Posthuma and Van der Sluis (2003), and Malkiel and Saha (2005).

[14] Recent examples include Koslowski (2007) and Ter Horst and Verbeek (2007).

we have an observation on the date of first reporting, and exclude performance data prior to that date.

2.2 Sample

There is a choice of data vendors of Hedge Fund and CTA data. We elect to use Lipper-TASS because it has relatively broad coverage of CTAs and includes flags for the date of first reporting by funds, thus allowing for backfill biases to be taken account of.[15]

Our sample of CTAs is taken from the April 5, 2008 version of the database. The Lipper-TASS database consists of 10,179 hedge funds. CTAs appear under the primary category "Managed Futures," which includes 827 funds (327 live and 500 in the graveyard module).[16] To avoid the backfill bias we select only those funds for which the date of first reporting is available, which excludes the separate CTA module for which this information is not available and 134 funds in the hedge fund module. Of the remaining funds, 108 were discarded because they did not have any returns after the first date of live entry. Finally we exclude funds (3) that do not report returns net of fees. The resulting sample consists of 582 funds, of which 201 were in existence as of the publication of the database. As of December 2007, our sample covers approximately 20% of all CTAs in terms of money-under-management (MUM).[17]

2.3 The Cross-section of Performance of CTAs

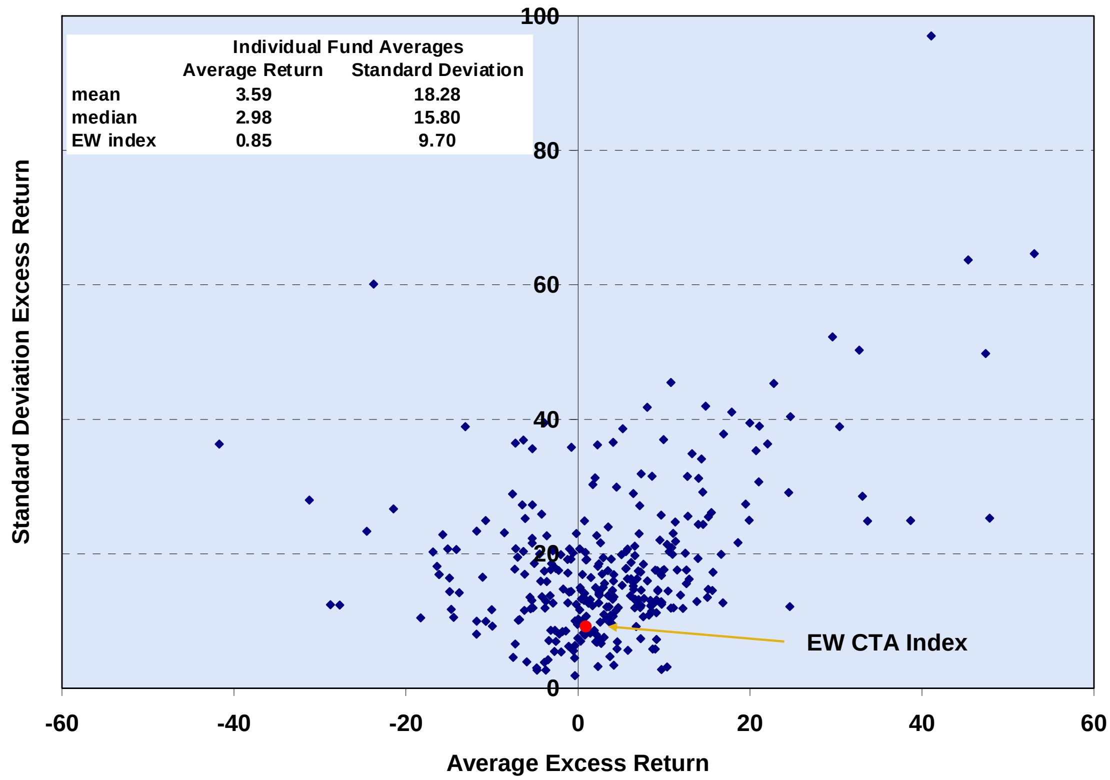
The poor performance of the average fund, as measured by the average return on the equally-weighted (EW) index, may mask the presence of stellar performers. Figure 2 provides a scatter plot of the average excess net returns and standard deviations of the individual funds in the database. In order to allow for a sufficient number of observations to calculate the average net return by fund, we restrict ourselves to CTAs that report at least 24 monthly observations (excluding backfill) in the database. This limits the number of observations to 312 (down from 582).

For comparison we include the EW CTA net return index. The graph shows large cross-sectional variation among individual CTAs. Annualized average excess net returns range from -42% to +53%, and standard deviations range from 1.9% to 97%. The V-shape of the graph reflects the intuition that funds that take more risk are more likely to exhibit extreme performance. The figure shows that the average standard deviation among individual funds (18.28%) is about double the standard deviation of the EW CTA index (9.70%), which suggests some diversification benefits to holding a portfolio of CTAs. Perhaps surprisingly, the average and median CTA has outperformed the EW index. However, this is caused by an increasing number of funds reporting to Lipper-TASS during the second half of our sample, which is also the period when the average fund performance was higher. In the remainder of the paper we will concentrate on the performance of the EW index rather than individual funds to further analyze the asset class. First, few individual funds have a long time-series to analyze, because the attrition rate of CTAs is high.

[15] The CISDM database, while potentially broader in scope lacks such a flag, which prevents us from indentifying backfilled returns. The Barclays database also lacks a backfill flag. The HFR database contains flags for backfilled returns but its coverage of CTAs is not as extensive as Lipper-TASS.
[16] Lipper-TASS contains a separate CTA module covering 2,149 funds which overlaps with the hedge fund module. All hedge funds classified as Managed Futures are in the CTA Module, and CTAs that are in the CTA Module but not "Managed Futures" do not appear elsewhere in the Hedge Fund Module.
[17] At the end of 2007, our sample contains 205 funds, of which 189 reported a combined MUM of $43.98 billion. BarclayHedge estimated industry-wide MUM to be $206.6 billion.

Figure 2: Individual CTA Risk and Return

The figure shows the annualized average excess return and standard deviation for all CTAs that have at least 24 months of reported returns in the Lipper-TASS database after excluding backfilled returns. Excess returns are calculated as total returns minus the three month T-bill rate.



Second, high individual fund volatility further complicates the inference about skill and style. In addition, the portfolio approach naturally takes into account the correlations among individual CTAs which are hard to model. We note that the performance of the EW index is lower than the average fund; this could bias out findings against finding average CTA skill if investors could have forecast this performance. This seems unlikely, but we will present separate results for the recent sub-period.

2.4 Robustness of Performance to Equal-Weighting of Funds

The fact that individual CTA performance has been higher during the second half of the sample, a time when the asset class experienced substantial inflows, suggests that investors may have rationally forecast the performance of successful managers. Also, in light of high entry and attrition rates of funds – discussed in more detail in section 5 – it is possible that the performance of the equally-weighted index is weighted down by a large number of small funds that briefly enter the database. It this is the case, an equally-weighted index would underperform an asset-weighted measure of performance. Unfortunately, this proposition is difficult to test due to incomplete data on asset under management in Lipper-TASS. For those funds that report a history of assets under management, we compared the performance of an equally-weighted index to an asset weighted index of funds, and find that an asset weighted index would have outperformed an equally weighted index by about 3% per annum between 1995 and 2007. This difference is not significant in a statistical sense ($t = 1.30$).

We also separately analyzed the performance of large CTAs. Perhaps CTAs signal with their "pedigree." In other words, new managers that spin-off from established, well-known, funds or who were trained at established, well-known, funds may use the name of the fund where they worked as an advertisement for their trading acumen. Insofar, as these spin-offs have a track record at their prior fund, it is not public, but still there may be a kind of "seal of approval" from the original fund. Many of these might find it easier to attract assets and turn into large funds. To address this issue, we constructed sample of large CTAs in our data set. These tend to include many of the well-known names.[18] In any given month a CTA is categorized as large if it had at least $250 million under management over the last 12 months. Then we constructed a Big CTA Index, which consists of the equally-weighted returns of the large CTAs, apart from the above mentioned cut-off we also constructed an equally weighted index of CTAs that had more then $100 million under management. Table 2, below, compares the performance of the equally weighted CTA index with the Big CTA indices discussed above. The comparison is for the period of 1998 to 2007.[19]

The Big CTA Index contains many of the large, well-known, CTAs – but their non-backfilled performance suggests that pedigree is not a signal, though it may be successful as an advertisement. More importantly, we find very little difference between the performance of large and small CTAs in our sample. For this reason, and because of the relative small number of funds for which data on assets is available, we decided to use the equally-weighted index of CTAs in the remainder of the paper.

Table 2: Big CTA Index Returns 1998-2007

The table gives the annualized average return, standard deviation and Sharpe ratio of the Equally-Weighted portfolio of CTAs in Lipper-TASS, and two portfolios of Big CTAs. A CTA is classified as Big in a year if it reports assets under management at some point during the prior 12 months that exceed USD 250 MM, or 100MM. All return calculations exclude reported returns prior to the date of first reporting to the database.

	EW CTA Index	CTAs > 250MM	CTA s > 100MM
Average Return	5.6	6.5	7.7
Standard Deviation	9.9	13.3	13.5
Sharpe Ratio	0.20	0.22	0.30

2.5 CTA Performance Before and After Fees

In addition to net (of fees) returns, we are interested in gross returns for two reasons. First, gross returns measure the payoffs to the fund's portfolio investments and speak to the question of whether a manager has the ability to generate positive investment returns. A comparison of gross and net returns indicates how the returns to skill are shared between the fund and its investors. The discrepancy is potentially large, because CTA fees resemble those of hedge funds: in our sample fixed fees on money-under-management range from 0.167% to 8.0% per annum while variable performance fees range from 0% to 50%. The average fixed fee is 2.15% and the variable fee averages 19.5% across funds. The second reason to study returns before fees is that gross returns are potentially better suited for performance analysis because the fee structure may induce additional nonlinearities in the post fee returns.

---

[18] Some of the well known names included in the sample are: Campbell & Company Inc., Graham Capital Management, Man Investments Ltd, Winton Capital Management Ltd, and Aspect Capital Ltd.

[19] There are very few big CTAs prior to 1998.

Table 3: CTA Excess Returns and Fees

The table gives the annualized average excess return and standard deviation of the equally-weighted portfolio of all CTAs in the Lipper-TASS database before and after fees, between 1994 and 2007. Before fee returns are estimated using net of fee data and fee information using the methodology outlined in French (2008).

	Average	Standard Deviation	t-statistic
After Fees	0.85	9.70	0.33
Before Fees	5.37	9.79	2.05

Brown et al. (2004) and French (2008) estimate gross returns for hedge funds from net returns and fee information. We follow French (2008) in the construction of gross returns for Managed Futures funds in Lipper-TASS, using the reported net returns. We make two assumptions implementing French's model, namely that fees accrue on a monthly basis, and that high watermarks, when applicable, increase at the rate of return on T-bills. Table 3 summarizes the effect of fees on performance.

The table shows that:

1. As a consequence of fees, the estimated average return on a fund's investments of 5.37% exceeds the return earned by investors (85 bps) by 4.52% per annum. Although not included in the table, of this difference 2.19% can be attributed to the fixed component of the fee structure, and 2.33% to variable performance fees.

2. We can reject the hypothesis that the average CTA has no ability to outperform T-bills. The gross excess return is marginally significantly different from zero (t = 2.05). However, most of this outperformance accrues to the fund management through levying fees, leaving on average 85 bp per annum for fund investors, an amount that is indistinguishable from zero in a statistical sense.

2.6 Performance Summary

The conclusion from this section is that the properly bias-adjusted average return to investors from CTAs has been poor between 1994 and 2007. Relative to T-bills, the average value added after fees – which is what investors care about – has been 85 basis points per annum. And in order to earn these returns, investors had to accept volatility at the fund level that has been comparable to investing in equity indices. Our observations closely resemble the central conclusions of the EGR studies (1987, 1990) which document poor performance of public commodity pools between 1979 and 1988. Why is it that CTAs not only have survived since the EGR publications, but have thrived as measured by the growth of money managed by the industry? We will return to a discussion of these issues in Sections 5 and 6.

The poor returns to CTAs do not imply an absence of skill of CTA fund managers. Our results are consistent with a world in which CTAs produce "alpha" before fees but successfully capture most of the rents they generate through charging (high) fees. In the next section we will attempt to identify particular investment strategies of CTAs. This is of interest because CTAs describe their style as predominantly trend-following, and academic research has documented that certain trend following (or momentum) strategies are profitable. Do CTAs extract fees from following simple strategies that are in the public domain? Or does a substantial component of their fees come from other sources that generate alpha?

In the next section, we will examine the correlation of CTA returns with various versions of simple dynamic strategies, which we will use as benchmarks for performance analysis. This will provide the answer to two questions. First, is there a predominant style for CTAs and how pervasive is this style? Second, how does CTA performance compare relative to these benchmarks?

3. Normative Asset Based Benchmarks

A central characteristic of hedge fund strategies is that they invest in active strategies, take both long and short positions and generally use leverage. For these reasons it has been difficult to specify appropriate benchmark returns that are comprised of passive strategies that capture the potential non-linear nature of hedge fund returns (see, for example, Hasanhodzic and Lo (2007) for a discussion). In the first subsection we illustrate the difficulties of developing benchmarks by looking at risk factors developed by Fung and Hsieh. Then, in subsection 3.2, we set out our own "Normative" benchmarks, factors that we think CTAs ought to reasonably outperform. In subsection 3.3 we analyze CTA gross return performance against the Normative benchmarks. Subsection 3.4 looks at subperiods. Subsection 3.5 summarizes our analysis of individual fund performance, as opposed to the EW index.

3.1 Fung and Hsieh Factors

Fung and Hsieh (FH) (2001) demonstrate that CTAs actively engage in trend-following strategies which generate option-like characteristics in their payoff structures. This motivates FH to conduct a style analysis in which they compare CTA returns to a dynamically traded portfolio of look-back (options) straddles. Fung and Hsieh (2004) label their approach "Asset Based Style Analysis".

We follow a similar approach in this paper, and construct a set of active strategy returns for each of three asset classes for which there exist liquid futures markets: commodities, foreign exchange, and equities. Our focus is slightly different from FH in that we are not merely interested in creating "positive" benchmarks that successfully describe the *style* of hedge funds. In addition we want our benchmarks to be "normative" and useful in evaluating the *performance* of hedge funds against these benchmarks. In particular, when the benchmarks are dynamic trading strategies themselves, there can be a tradeoff between the objective of capturing style and measuring performance. To illustrate this issue consider the following regression of the EW Index of before-fee CTA (gross) excess returns on the FH factors (using their notation):[20]

[20] The five factors (PTFSBD, PTFSFX, PTFSCOM, PTFSIR, and PTSSTK) are factors that have been constructed by Fung and Hsieh (2001) to represent nonlinear trading strategies designed to capture "trend following" by CTAs. Each acronym starts with the prefix "Primitive Trend-Following Strategy" and then includes Bonds (BD), Foreign Exchange (FX), Commodity Markets (COM), Interest Rates (IR), and Stocks (STK). Construction of these factors involves rolling a pair of lookback straddles for various asset classes. Applying the analysis to CTAs, Fung and Hsieh interpret their results as supporting the view that CTAs follow nonlinear, option-like, strategies. Fung and Hsieh (2001) conclude that the use of their nonlinear factors "supports our contention that trend followers have nonlinear option-like strategies" (p. 337).

$$R_{EW} = \underset{(3.81)}{0.77} + \underset{(2.40)}{0.02} \times PTFSBD + \underset{(4.32)}{0.04} \times PTFSFX + \underset{(3.06)}{0.05} \times PTFCOM - \underset{(-2.91)}{0.02} \times PTFSIR + \underset{(2.90)}{0.04} \times PTFSSTK$$

$$\overline{R}^2 = 0.25$$

where the dependent variable represents the excess gross returns of the equal weighted portfolio of CTAs and the independent variables are the excess returns of the FH style factors corresponding to bonds (PTFSBD), currencies (PTFSFX), commodities (PTFCOM), interest rates (PTFSIR), and equities (PTFSSTK). As explained above, we analyze returns gross of fees because that return series captures the talent of the average manager. The regression shows that the various style factors explain about 25 percent of the variance of CTA excess gross returns. And controlling for exposure to the various styles, the average CTA earns an excess return of 0.77 percent per month (t = 3.81), which is about 9.2 percent annualized. The regression seems to indicate that the style factors are somewhat successful in capturing various aspects of CTA return variance, and provides evidence of positive excess gross returns after controlling for style ("alpha").

The interpretation of the regression alpha is complicated by the fact that the style factor returns correspond themselves to dynamic trading strategies, which may be inefficient replications of that particular style. Although the payoffs to trend-following rules can mimic those of look-back options strategies described by FH, it is likely that CTAs will achieve these payoffs by directly trading in futures markets rather than options markets. The return on trading look-back straddles would understate the achievable returns to the trend-following style. In what follows, we will show that trend-following characteristics are as easily captured by simple momentum strategies, which outperform the FH style factors and change the inference about the presence of "alpha."[21] Table 4 gives the excess returns on the FH factors between 1994 and 2007.

Table 4: Performance of the Fung and Hsieh Style Factors

The table gives the annualized average excess returns and standard deviations of the style factors reported to capture "trend following" in FH (2001). These factors are constructed as the returns on look-back option straddles in bonds (BD), Currencies (FX), Commodities (Com), Interest Rates (IR) and Stocks (STK). Sample consists of monthly data between 1994 and 2007. Data Source: http://faculty.fuqua.duke.edu/~dah7/HFRFData.htm

	Mean		**Standard Deviation**
	Arithmetic	**Geometric**	
PTFSBD	-19.5%	-27.2%	51.1%
PTFSFX	-4.0%	-20.7%	64.9%
PTFSCOM	-9.3%	-17.5%	46.1%
PTFSIR	5.2%	-17.0%	85.8%
PTFSSTK	-64.7%	-53.6%	44.0%

The table shows that the (geometric) average excess returns of the FH style factors has been negative (and highly volatile) over the 14-year period between 1994 and 2007. The issue is more dramatically illustrated in Figure 3, which plots the cumulative return for the five FH factors: a dollar invested in each of these factors at the end of 1993 would have lost more than 90 cents of its value by 2007. Unlike passive benchmarks, dynamic rebalanced portfolios may require frequent trading. And because the style returns are measured before transactions costs, accounting for trading costs in options markets would further lower the reported averages in the table.

[21] See also Cremers, Petajisto, and Zitzewitz (2008) of a discussion of the effect of nonzero alphas of benchmark indices on performance attribution.

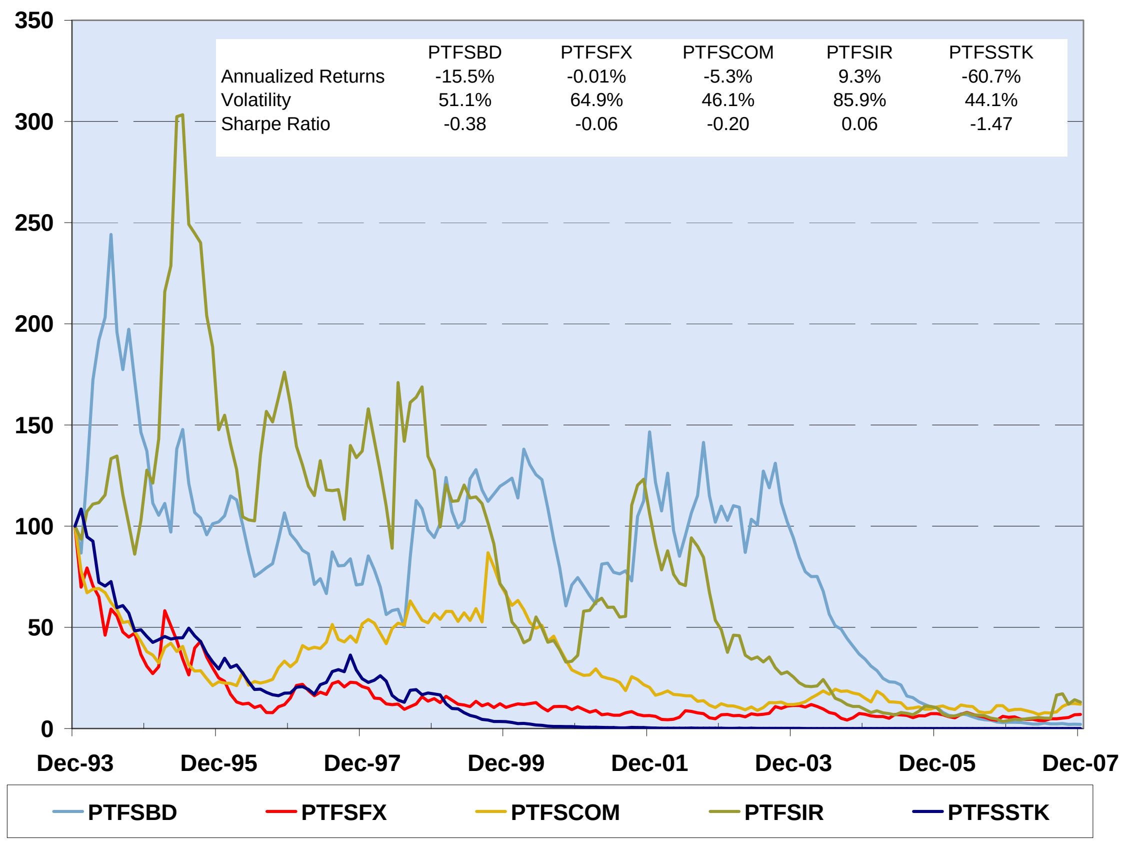
Measurement error in the style returns induces measurement error in the alpha of CTAs relative to these style benchmarks. Because the average CTA positively loads on the FH style portfolios, the resulting alpha will exceed the raw excess gross return to CTAs. To the extent that the style returns reflect inefficient replication of the trading strategies followed by CTAs, this will lead to an upward bias in the alpha. It seems unlikely that CTAs would choose to follow styles that have earned negative returns over a 14-year period. Instead, it seems more plausible that the negative style returns and the apparent positive alpha are merely a reflection of inefficient benchmarks.

In fairness, the FH style factors were not, of course, intended for the purpose of performance evaluation, yet the example illustrates the tradeoff between capturing style and performance evaluation when the style portfolios are not passive benchmarks. To the extent that options are expensive, a strategy that buys straddles to mimic trend-following behavior will exhibit negative excess returns.

In the spirit of the FH analysis we propose to evaluate the style of CTAs by correlating their returns to those of dynamic trading strategies in equities, currencies and commodities. Our strategies differ in two respects from FH. First they are relatively cheap to trade, and therefore are more useful for performance evaluation. For example we evaluate the performance of CTAs against a set of simple momentum strategies which are likely to capture the basic characteristics

Figure 3: Cumulative Performance Fung-Hsieh Trend Following Factors

The figure shows the cumulative total return, standard deviation, and Sharpe Ratio of the trend following factors (PTFS) reported by Fung and Hsieh (2001). These factors are constructed as the returns on look-back straddles in bonds (BD), currencies (FX), Commodities (COM), interest rates (IR) and Stocks (STK). Sample consists of monthly data from 1994 to 2007 Data Source: http://faculty.fuqua.duke.edu/~dah7/HFRFData.htm



of trend-following but are likely to be cheaper than option straddles. Second, our choice of benchmarks is not just based on what CTAs self-purportedly do (trend-following), but also on what they ought to be doing, in the sense that the strategies are dynamic strategies in the public domain. In addition to momentum, we select for each asset class a second style factor that is in the public domain and has been documented to be correlated with average returns. These factors are value (price-to-book) for equities, interest rate differentials (the carry trade) for currencies, and the basis (backwardation) for commodity futures. We call these benchmarks "Normative" benchmarks.

3.2 Normative Benchmarks Performance

We construct the Normative Benchmarks by constructing rules-based active strategies using primitive assets that include currency futures, commodity futures and country equity indices. The active strategies are intended to capture known sources of return as well as self-declared styles of CTAs. In our selection of benchmark portfolios, we are guided by the academic literature, to ensure a reasonable expectation that these benchmarks are indeed in the public domain and therefore available to CTAs. In line with the previous evidence of CTA trend-following, we construct a momentum factor for each of the three major asset classes: currencies, commodities, and equities. In addition, we construct actively traded portfolios based on the forward bias in currencies ("carry trade"), a factor to capture inventory effects ("backwardation") in commodities markets, and a factor related to cross-country value in equity markets (the price-to-book ratio (PB)). A detailed discussion of the construction of these factors is contained in the Appendix.

Table 5: Annualized Average Excess Returns and Standard Deviation of Normative Benchmarks 1993/12 – 2007/12

The table gives the average excess return, standard deviation, and t-statistic for a test of non-zero average excess return for the Equally-Weighted portfolio of CTAs and portfolios of dynamically traded futures of Commodities, Equities, and Currencies. Dynamic portfolios are constructed by monthly sorting commodity, equity and currency futures on past performance (Momentum), end of prior month futures Basis (Commodities, and Currencies) or Price-to-Book (Equities). Long-Only indices take long positions in the top half of the relevant assets in this ranking, while Long-Short takes a long position in the top half and a short position in the bottom half of the futures in the ranking.

Panel A: Long-only

		Average	Volatility	*t*-stat (Average)
EW CTA Index After Fees		0.9%	9.7%	0.33
EW CTA Index Before Fees		5.4%	9.8%	2.05
Commodities	**Hi Momentum**	15.1%	11.6%	4.88
	Hi Basis	13.0%	11.7%	4.17
Equities	**Hi Momentum**	9.4%	14.5%	2.43
	Low PB	8.7%	15.0%	2.15
Currencies	**Hi Momentum**	2.6%	7.2%	1.34
	Hi Basis	4.1%	6.8%	2.27

Panel B: Long-Short

		Average	Volatility	*t*-stat (Average)
Commodities	**LS Momentum**	15.9%	14.3%	4.17
	LS Basis	11.9%	12.6%	3.52
Equities	**LS Momentum**	3.9%	8.6%	1.71
	Low minus Hi PB	3.0%	6.5%	1.73
Currencies	**LS Momentum**	0.7%	5.2%	0.51
	LS Basis	4.0%	5.8%	2.60

Table 5 summarizes the excess returns to the Normative Benchmarks between 1994 and 2007. For each asset class and strategy, we report both the excess returns of the High ("HI") characteristic portfolio (Long-Only), as well as the excess return difference of High minus Low (Long-Short or "LS").

The table illustrates the following points:

1. Contrary to the FH factors, all of our Normative Benchmarks have earned positive risk premiums. Long-Only excess returns range from 2.6% per annum (FX Momentum) to 15.1% (Commodities Momentum). With the exception of the FX Momentum premium, all Normative factor premiums are significantly different from zero.

2. The Normative factor excess returns exceed the average return of the equal-weighted return CTA index after fees and, with the exception of currencies, exceed the average return of the equal-weighted return CTA index before fees.

3. The Long-Short excess returns in Panel B are slightly lower than the Long-Only excess return, but with the exception of LS FX exceed the excess returns on the EW CTA after fees index. Only the commodities strategies exceed the EW CTA before fees returns.

The table is compelling in the sense that average CTA performance (before or after fees) is poor relative to most of our Normative Benchmarks. The full sample Sharpe Ratio of CTAs after fees is only 0.09, and the Sharpe ratio before fees is 0.55, as compared to the Sharpe ratios of the active strategies which exceed 0.94 in the case of commodities, and 0.14 in the case of currencies.

The finding of relatively low gross returns already suggests that CTAs follow strategies that are different from those embedded in the benchmarks. A formal performance evaluation is the subject of the next subsection of the paper. Regression of estimated gross excess returns on the Normative Benchmarks addresses the question whether CTAs earn alpha relative to a set of strategies that are in the public domain.

3.3 The Performance CTAs and Asset Based Style Benchmarks

Table 6 contains the regression results of the gross excess return of the EW CTA index on the excess returns of the various factors using data from 1994 to 2007, as well as two sub-sample periods. The slope coefficients and R-squared of these regressions are informative about average CTA style, while the constant term provides us with the estimate of alpha conditional on the style factors. All specifications include the S&P 500, the Lehmann Aggregate Bond Index, and the Gorton and Rouwenhorst (2006) Equally-Weighted Commodity Index (GRCI). The motivation for including these three indices is twofold: first, they benchmark the CTA returns relative to the basic passive asset class exposures. Second, we find that including passive benchmarks seems to alleviate the problems outlined in section *3.2*, where the regressions alphas are biased by inefficient replication of style factors.

In addition to the passive benchmarks, we contrast three dynamic style benchmarks. The first is the Mount Lucas Index (MLM), a commercially-produced index that equally-weights 25 different

Table 6: The Abnormal Performance of CTAs

The table gives the results of a regression of the excess return of the Equally-Weighted portfolio of CTAs, gross of fees and corrected for survivorship bias and backfill bias on three groups of style factors. *MLM* refers to the excess return of the Mount Lucas Index, *F&H* are excess returns of the five trend following style factors of Fung and Hsieh (2001), and the LS are the excess returns of asset based style factors based on long-short positions in Commodities (based on momentum and the basis), Stocks (based on price-to-book and momentum), and Currencies (momentum and the basis). In addition to the sets of style factors each regression includes the excess return of three passive benchmarks: SP500, Lehman Aggregate Bond Index, and the Equally-Weighted Commodity index described in Gorton and Rouwenhorst (2006). In parentheses below the coefficients are *t*-statistics, corrected for heteroskedasticity.

	Sample: 1994-2007				Sample: 1994-2000				Sample: 2001-2007			
		MLM Index	F & H Factors	LS		MLM Index	F & H Factors	LS		MLM Index	F & H Factors	LS
Constant	0.265	0.147	0.582	0.028	0.000	-0.277	0.117	-0.313	0.484	0.651	1.069	0.562
	(1.22)	(0.74)	(2.85)	(0.13)	(0)	(-1.01)	(0.44)	(-1.15)	(1.52)	(2.41)	(3.23)	(1.74)
Commodities LS Momentum				0.367				0.403				0.369
				(4.23)				(2.46)				(3.53)
Commodities LS Basis				-0.110				-0.196				-0.078
				(-1.06)				(-1.03)				(-0.62)
Equities Hi minus Low Momentum				0.106				-0.319				0.504
				(0.7)				(-1.57)				(1.92)
Equities Low minus Hi PB				0.039				0.028				-0.256
				(0.19)				(0.1)				(-0.88)
FX LS Momentum				1.331				1.272				1.615
				(5.5)				(3.95)				(3.72)
FX LS Basis				0.048				0.203				-0.220
				(0.23)				(0.82)				(-0.52)
MLM		0.571				0.485				0.767		
		(5.4)				(3.23)				(3.9)		
PTFSBD			0.025				0.034				0.033	
			(2.45)				(2.27)				(2.27)	
PTFSFX			0.040				0.028				0.053	
			(4.22)				(2.86)				(3.26)	
PTFSCOM			0.046				0.048				0.044	
			(2.63)				(1.73)				(2.24)	
PTFSIR			-0.012				-0.017				-0.012	
			(-1.78)				(-1.46)				(-1.26)	
PTFSSTK			0.038				0.016				0.061	
			(2.68)				(0.86)				(2.97)	
Rbar-squared	0.097	0.201	0.296	0.233	0.036	0.106	0.204	0.155	0.119	0.282	0.343	0.329
Durbin-Watson	1.908	1.846	1.882	1.923	2.090	2.008	2.028	2.182	1.714	1.795	1.878	1.788
Num of Obs	168	168	168	168	84	84	84	84	84	84	84	84

futures contracts that cover foreign exchange, energy, financials, metal and agricultural futures.[22] It is rebalanced monthly, as follows. If the 200 day moving average is greater then the closing price of the future then it takes a short position otherwise it takes a long position. The MLM index is a widely used benchmark for CTAs, which is why we include it here. We compare including MLM to the FH style regression and our Normative Benchmarks. White heteroskedasticity-adjusted *t*-values are in parentheses. The Table shows that:

1. The estimated alpha of CTAs after controlling for passive exposure to stocks, bonds and commodities CTAs is 0.265% per month (3.2% annualized) between 1994 and 2007, which is insignificantly different from zero. Although we estimated earlier that CTAs earn positive excess returns before fees, we cannot reject absence of skill after controlling for passive asset class exposures.

2. The only instances in which we can reject absence of alpha are for the FH and MLM benchmarks during the 2001-2007 sub-period, and FH factors for 1994-2007. A comparison with the constant term of a regression on only the passive benchmarks shows that the MLM and FH alphas are driven by negative realized excess returns to the MLM and FH factors during the second half of the sample. CTAs do not add value relative to the Normative Benchmarks.

3. Style benchmarks can explain up to about 30% of the full sample variance of the EW CTA gross of fees index (as measured by adjusted R-squared). However, the explained variation remains low, which suggests that there may be potentially important, omitted style factors.

4. Among the two styles, CTAs tend to have higher average exposure to Momentum factors than Basis or Value. This is consistent with previous studies that identified trend-following as the major CTA style. Among the three asset classes, Momentum exposure is highest in Currencies and Commodities.

The overall conclusion of this section is that the average CTA – as measured by the EW gross of fees performance index – has failed to deliver alpha to investors. The predominant style has been one of trend-following, most pronounced in currencies, but the combined factors have a maximum explanatory power of 33% (over the period 2001-2007). This evidence does not support the hypothesis that CTAs on average adhere to the trading strategies as embodied in the Normative Benchmarks and capture the apparent profits of these strategies through charging fees. Instead, it suggests that CTA performance has a large idiosyncratic component, and that the combination of poor performance and high fees has on average resulted in absence of alpha for investors.

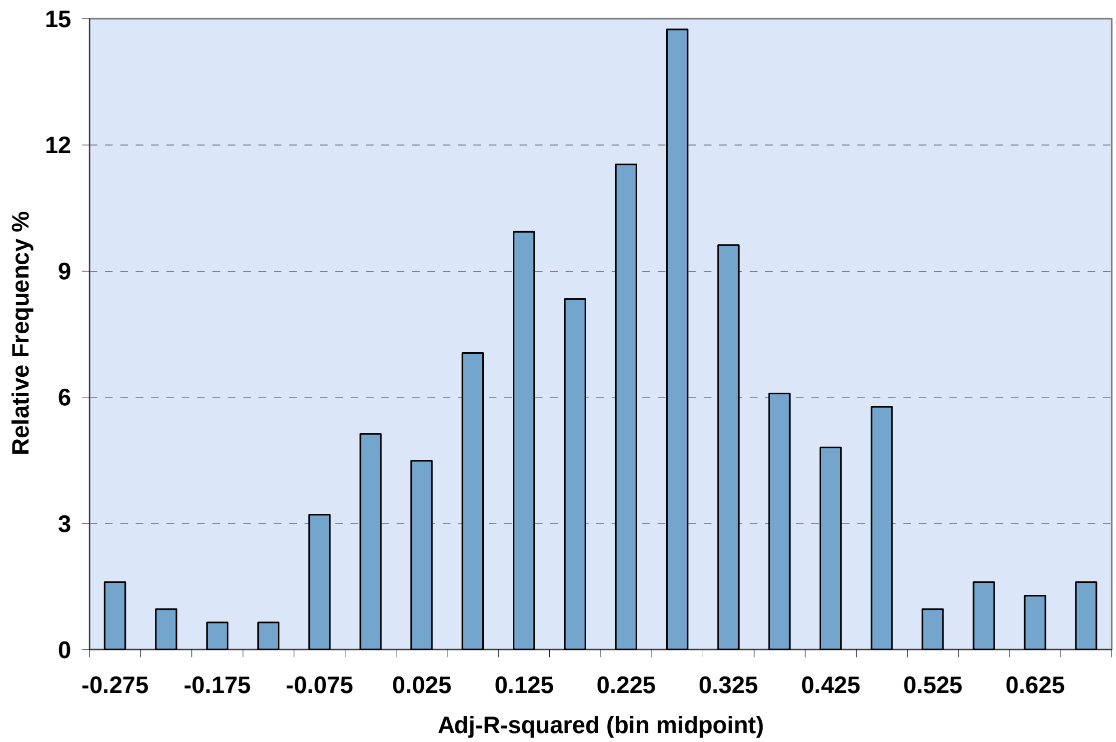
3.4 Individual Fund Analysis

Looking at the alphas on individual CTAs, with at least 24 months of returns, relative to the Normative Benchmarks, we find that 21 percent of the CTAs have an alpha which is significant at the five percent level. These alphas are about evenly divided among positive and negative, and the average of these significant alphas is −0.14 percent per month. So, CTAs perform poorly on average, but even those with individually significant alpha are not particularly good performers.

[22] For the Mount Lucas Index (MLM), see Hhttps://www.mtlucas.com/about.aspxH . On MLM, see Mulvey, Kaul, and Simsek (2003).

Given the relatively small strategy space of CTAs it is perhaps surprising that style analysis can explain only about a third of the variance of the returns on the EW CTA gross of fees index. As shown in Figure 4, however, a style analysis conducted at the individual fund level reveals a similar picture. The histogram shows the distribution of the R-squared of a regression of individual fund level returns on the Normative Benchmarks. The figure shows that 74% of the funds have an R-squared that is below the index-level regression (0.33). And about 24% of regression R-squareds is below 10%. This further complicates the inference problems that investors in CTAs face. Even if investors are able to obtain clean performance data, an analysis of the self-proclaimed style of CTAs can explain less than 30% of the return variance for most of the funds. For 12% of the CTAs the adjusted R-squared is not positive. The overall conclusion of our style analysis is therefore that proclaimed style explains very little of the variance of individual CTA returns as well as of the returns to the broader asset class.[23]

Figure 4: Distribution of R-squared of Individual Fund Returns on Normative Style Factors
For each CTA with at least a 24-month return history after controlling for backfill and survivorship bias, we regress the excess fund gross return on the excess return of the 6 style benchmarks that capture Momentum (Commodities, Currencies and Equities), Basis (Commodities, Currencies) and Value (Price-to-Book, Stocks). The figure provides a histogram of the distribution of the R-squareds of the fund-level regressions.



4. Why do CTAs Persist? An Historical Perspective

Our conclusion about the lack of return and alpha of CTAs is surprising because the asset class has experienced substantial inflows over time. For example, BarclayHedge estimated money-under-management of CTAs to be $206.6 billion as of December 2007, which is an increase of

[23] Results are qualitatively similar for net-of-fee fund level returns.

306 percent from 2002 when assets were estimated to be $50.9 billion. Apparently investors increased allocations following years of poor performance. This finding is perhaps even more striking considering the broader historical context of commodity funds provided by the studies of Elton, Gruber and Renzler (EGR 1987, 1989, 1990). They studied the performance of publicly traded commodity funds between 1979 and 1985, and document a similarly poor performance in their earlier sample. EGR attributed the existence and persistence of poorly-performing CTAs to misinformation. They find that no fund is able to outperform its prospectus track record. EGR concluded that the historical returns series provided in the prospectuses of public commodity funds were misleading. The past performance was unreasonably upward biased and investors had no other information to rely on.

The EGR studies were widely reported on in the press, sometimes in scathing terms. See, for example, Newswire (September 18, 1986), the Wall Street Journal (September 29, 1986), the Toronto Star (October 5, 1986), Chicago-Sun Times (November 17, 1986), the Washington Post (March 1, 1987, September 28, 1987), the St. Petersburg Times (September 26, 1987), the San Francisco Chronicle (October 5, 1987), the New York Times (August 20, 1988), Business Week (November 28, 1988), The Economist (December 1, 1990), Forbes (September 2, 1991). As an example, here is one part of an article on their studies by Jane Bryant Quinn in the San Francisco Chronicle (February 21, 1989):

> The larger - and more intractable - scandal lies in the entirely legal deceptions that surround the selling of commodities funds in the first place. Brokerage firms mislead you as a matter of course, with the full approval of the market's so-called regulators.
>
> The problems lie in the sales brochures and prospectuses for new commodities funds. They "disclose" the portfolio manager's past performance, which is never anything less than spectacular. Gains may be claimed of 50 percent, 60 percent, even 70 percent a year.
>
> But those astonishing track records can be a clever form of fiction. They're not wrong, exactly. But they're biased and misleading. They greatly exaggerate the manager's chance of success.
>
> For proof, I give you a study by three New York professors - Edwin Elton and Martin Gruber of New York University's Graduate School of Business, and Joel Rentzler of the Baruch College of the City University of New York. They took 77 new commodity funds, and compared the managers' past performance with how well the funds actually did in practice. The verdict: disaster.

Given the widespread publicity, it is hard to believe that investors would continue investing. But, things did change, in two important ways. First there was a regulatory reporting change. Second, the form of the investment vehicle changed.

On the regulatory front, subsequent to the EGR papers, the Securities and Exchange Commission put out a STATEMENT OF THE COMMISSION REGARDING DISCLOSURE BY ISSUERS OF INTERESTS IN PUBLICLY OFFERED COMMODITY POOLS SECURITIES AND EXCHANGE COMMISSION, Release Nos. 33-6815; 34-26508 [S7-1-89]; 17 CFR Parts 231 AND 241, February 1, 1989, which said in part::

> Certain recently published studies suggest that the actual performance of publicly held commodity pools was significantly lower than the performance disclosed in

> the prior performance tables included in commodity pool disclosure documents.* While the findings and issues raised in these studies are currently being reviewed by the staff of the CFTC, the Commission believes that it should provide guidance to issuers of publicly offered commodity pools at this time. Although the positions expressed in this release and the CFTC's interpretive statement currently reflect the respective agencies' views regarding appropriate disclosure in commodity pool disclosure documents, the Commission is interested in receiving views on the interpretive positions expressed in those statements. Commentators may wish to make the same submission to both agencies. The Commission expects to consult with the CFTC concerning the comments received in response to their respective statements with a view towards determining whether further action is necessary or appropriate.
>
> * See Elton, Gruber & Rentzler, New Public Offerings, Information and Investor Rationality: The Case of Publicly Offered Funds, 62 J. Bus. 1-15 (January, 1989). The authors hypothesized that the findings of the study were at least in part due to the following factors: 1) public commodity pools have larger transaction costs and management fees than private commodity accounts; 2) only trading advisers with recent successful track records are likely to go public; and 3) trading advisers can select the period of time for disclosing their prior performance, resulting in an upward bias in performance results. See also Edwards & Ma, Commodity Pool Performance: Is the Information Contained in Pool Prospectuses Useful? Working Paper Series No. 16, Center for the Study of Futures Markets, Columbia Business School (January, 1988). [Footnote in original.]

Subsequently, filings of public commodity funds included the SEC statement in their filings.[24] Also, the CFTC did change the reporting requirements.

The second change appears to have been a response from the commodity fund industry to the EGR publicity. After the EGR studies and ensuing publicity, and after the CFTC reporting requirement changes, commodity trading advisors appear to have stopped the frequent use of publicly-offered funds, which required a prospectus following the new rules. Rather, commodity fund managers began to structure themselves like hedge funds, which require less disclosure. One possibility is that this change in organizational form was enough to entice investors to continue to invest.

5. Explaining the of Persistence CTAs

Data sets that have been strategically manipulated not only make it hard for econometricians to draw inferences, investors have the same problem. It is difficult to evaluate performance and, as we have discussed, even to determine CTAs' style. In this section we delve into two related issues. First, we ask whether CTA return distributions have desirable characteristics that are not captured by means and variances. In particular we examine whether individual CTA returns exhibit skewness and coskewness with other asset classes, which might explain why the asset class can persist despite offering poor returns on average.

The second set of explanations focuses on the information asymmetry between investors and funds. We first investigate whether there is evidence that talented CTAs try to overcome the

[24]For example, the JWH Global Trust S-1 on Nov. 26, 1996 (see Hhttp://www.secinfo.com/dRqWm.9rzv.2.htmH , section Ex-99.01). Other examples include Kenmar Global Trust, July 25, 1996 (Hhttp://sec.edgar-online.com/1996/07/25/00/0001005477-96-000208/Section21.aspH), and also MAN-Ahl 130/LLC S-1/A Nov. 11, 2005, Ex-99.01: Hhttp://www.secinfo.com/dsvRm.zcZk.8.htm#1stPageH .

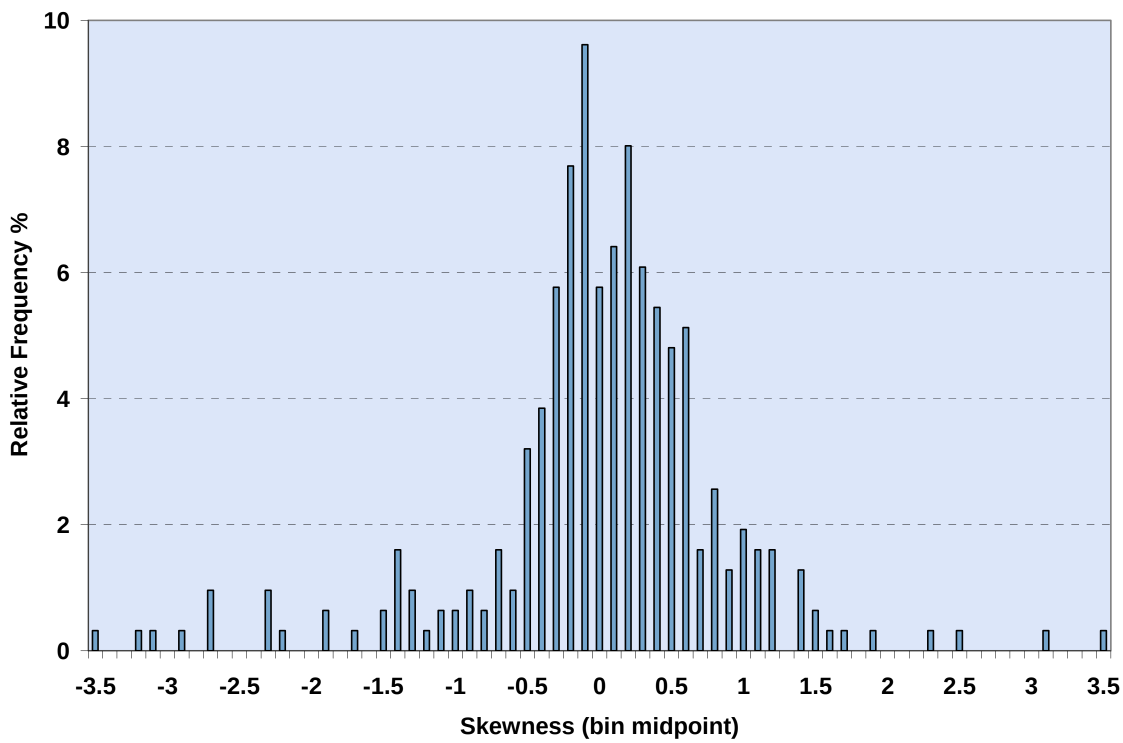
information problem by signaling. Secondly, is there evidence that investors are aware of the information issues, concerning, for example, biased performance data?

5.1 Zero Alpha, but Positive Skewness: Are CTAs Lotteries?

Perhaps CTAs generate desirable skewed returns. It is well-known that if investors have utility functions that display decreasing absolute risk aversion, then they will have a preference for positively skewed returns (see, e.g., Markowitz (1952) and Arrow (1971)). Also, the literature in behavioral finance suggests that investors may have a preference for skewed payoffs (see, e.g., Barberis and Huang (2007)). There is some evidence that investors indeed do have a preference for skewness. For example, Levy and Sarnat (1984) find a strong preference for positive skewness in a study of mutual funds. Also, see Polkovnichenko (2005). It is possible to reconcile the poor performance of CTAs with the growth in assets-under-management if CTA returns exhibit positive skewness. We calculate the skewness of individual CTAs using the full-sample of available returns for all the funds that have at least 24 months of returns, 312 funds. Figure 5 below provides a histogram of the sample skewness of individual funds.

Figure 5: The Distribution of Individual CTA Return Skewness

For each CTA with at least a 24-month return history after controlling for backfill and survivorship bias, we calculate the skewness of monthly returns. The figure provides a histogram of the distribution of the calculated fund-level skewness.



The histogram in Figure 5 shows that the number of CTAs with positive skewness is about the same as the number of funds with negative skewness. The median estimated skewness is 0.056 and the average is 0.0213. The figure also shows that some CTAs do exhibit large skewness, but these are equally divided over the positive and negative tails of the empirical distribution. Absent skewness, it is unlikely that CTAs attract assets because they offer lottery-like payoffs to investors.

A related issue concerns coskewness with other asset classes. Fung and Hsieh (2001) argue that CTAs funds are attractive because they do well when other asset classes are not doing well, in particular when there are "tail events" in which the other asset classes are doing particularly poorly. Vice versa, perhaps the other asset classes are doing well when CTAs are doing particularly poorly. We look at this issue in a very simple way. We look at the performance of the S&P500, the Lehman Aggregate Bond Index (LABI), and the equally-weighted CTA Index in the months where each had the 5 percent worst months of performance and the 5 percent best months of performance in our sample period. For those months we ask how the other asset classes performed. Table 7 shows the results.

Table 7: Tail Correlation during Extreme Events

For Each asset class, including the S&P500, the Lehman Aggregate Bond Index, and the Equally-weighted CTA Index, we compute the average annualized return for the other indices in the months where the specified index had its 5 percent worst performing months and the 5 percent of the best performing months.

Panel A: Best and Worst 5% of S&P500 Months

	Worst 5% S&P500 Months		
	CTA	**S&P500**	**LABI**
Monthly Average ER	4.1%	-9.3%	0.7%
	Best 5% S&P500 Months		
	CTA	**S&P500**	**LABI**
Monthly Average ER	-1.0%	7.9%	0.2%

Panel B: Best and Worst 5% of Lehman Aggregate Bond Index Months

	Worst 5% LABI Months		
	CTA	**S&P500**	**LABI**
Monthly Average ER	-1.0%	-1.1%	-4.7%
	Best 5% LABI Months		
	CTA	**S&P500**	**LABI**
Monthly Average ER	2.3	2.9%	4.5%

Panel C: Best and Worst 5% of the EQ CTA Index Months

	Worst 5% CTA Months		
	CTA	**S&P500**	**LABI**
Monthly Average ER	-4.9%	2.1%	-1.8%
	Best 5% CTA Months		
	CTA	**S&P500**	**LABI**
Monthly Average ER	6.6	-2.4%	1.6%

Looking at Panel A of the table, CTAs do well in the months when the S&P500 is doing very poorly and conversely do poorly when the S&P500 is doing well. Fung and Hsieh (2001) also make this point. This pattern is also true when we select the worst months and best months for CTAs, Panel C of the table. While there is nothing that stands out in this regard for bonds; see Panel B. What is less clear is whether this tail behavior is sufficient to justify an investment in

CTAs despite the poor performance. This would not seem to be justification since this type of diversification can be achieved at much lower cost using passive indices of commodity futures; see Gorton and Rouwenhorst (2006). The correlations of CTA returns with traditional asset classes in the tails seem unlikely to justify investors allocating $200 billion to an asset class that offers T-bill returns with a standard deviation that is comparable to equities.

We conclude that there is no compelling evidence to justify investing in CTAS in a portfolio context.

5.2 Signaling

It is possible that some CTAs are talented and want to signal their ability, in the face of the lack of credible information and relevant benchmarks. More onerous contract terms for a CTA may signal a more talented fund manager, who is confident in his abilities. This private information may be conveyed contractually by agreeing with the investor to have a high-water mark and no lockup, as opposed to weak contract terms like no high-water mark and a lockup period, for example. A high-water mark (HWM) imposes discipline on the manager's performance and if this performance is poor, the investor with no lockup can disinvest quickly.

Our data set contains information about the contract terms and fee structure. The contract terms for which we have data are high water mark (HWM) and lockup. It turns out that very few CTAs have lockup provisions, likely reflecting the fact that futures are very liquid markets. CTAs do show variation with respect to whether their contract includes a high water mark and their fees differ, although the fee structure of 2% fee on money-under-management and 20 percent of the gains above the high water market ("2-20') predominates.

We look at this signaling hypothesis in Table 8, which shows the annualized average net-of-fees returns, the standard deviation of the those returns, and the Sharpe Ratio (excess return/standard deviation) (SR) for an equally-weighted CTA index of the CTAs with and without high water marks and for those with 2-20 fee terms

We restrict attention to the period starting in 2001 because prior to that most CTAs did not have high water marks, an observation discussed further below. The top part of the table shows the results for the period 2001-2007. The relevant comparison is between one of the two categories (ALL and 2-20) with a high water mark (HWM) to the same category with no high water mark (No-HWM). For example, in the case of ALL, the average CTA with a high water market had a Sharpe ratio of 0.48 while those with no high water mark had a Sharpe ratio of 0.41.

Table 8: Returns by Contract Type

The table gives the annualized average returns (TR), standard deviations (Vol), and Sharpe Ratio (SR) of CTAs sorted by the presence of contract terms, High water Mark (HWM) and whether there is a 2-20 fee structure. The sample period is 2001/10 to 2007/12 and includes subperiods.

		HWM		**NO-HWM**	
		ALL	**2-20**	**ALL**	**2-20**
2001-2007	**TR**	8.2%	3.4%	7.2%	3.3%
	Vol	10.8%	10.1%	10.4%	11.2%
	SR	0.48	0.06	0.41	0.04

The table shows that the difference between the average returns on HWM and No HWM are not statistically different, for either ALL of the 2-20 fee structure. Overall, there appears to be no signaling. [25] The table also shows that the funds offering the 2-20 contract terms do not do as well as the population as a whole. It is not clear why this is the case. CTAs increasingly adopted hedge fund-like contracts starting around 2000, but it is not clear why those adopting 2-20 should be worse performers.

We noted above that CTAs have increasingly included high water marks in their contracts. During the period 1994 through 1996 no CTAs in our sample had high water marks in their contracts. Thereafter, the percentage rises almost monotonically to 71 percent by the end of 2007. This period also happens to coincide with an increase in the importance of hedge funds, which predominantly have high water marks in their contracts. One interpretation of the CTA behavior is that they were forced by hedge fund competition to include high water marks. This is consistent with investors receiving slightly better returns during this period. One way to measure this is to look at the performance fee as a percentage of gross returns, that is, what fraction of the gross return that is charged by the CTA as a performance fee? If hedge funds are a source of competitive pressure for CTAs, then this fraction should be going down. Looking only at the years where CTA gross returns are positive (which eliminates the years 1994 and 1999), the performance fee as a percentage of gross return averaged 28 percent for the period 1994 through 2000, and averaged 20 percent for the period 2001 through 2007. It appears that CTAs changed their contract terms to be more like those of hedge funds, and were forced to share a bit more with outside investors.

Without signaling, investors may not be able to distinguish talented CTAs from those without talent. But, the evidence suggests that there is a dearth of talent in the asset class generally. Perhaps talented CTAs do not enter the industry because they cannot differentiate themselves via signaling.

5.3 Information About the Information Problems

Investors may simply be unaware of the poor CTA performance. But, the information setting of investors is different during the period we study compared to the earlier period that Elton, Gruber and Rentzler (EGR) studied, the period from July 1979 to June 1985. EGR attributed the existence and persistence of poorly-performing CTAs to misinformation. Namely, the historical returns series provided in the prospectuses of public commodity funds were misleading and investors believed this misinformation. The past performance was unreasonably upward biased and investors had no other information to rely on.

Combining our evidence with that of EGR suggests that CTAs have been successful in taking money from investors without adding value for about twenty years. How can CTAs persist for so long despite their poor performance? To persist, there must be a demand for CTAs and a supply of CTAs. With regard to demand, we hypothesize that while investors are rational, acquiring information, overcoming the biases and lack of benchmarks, is costly and there is no common

[25] A potential difficulty with looking at the period 2001-2007 is that different CTAs are in existence at different times, so an equally-weighted index is reflecting a potentially varying population. To address this we also looked at subperiods, where this issue is mitigated. Looking at the subperiods, there are no real differences and these results are omitted.

knowledge about the experience of poor returns. Understanding the supply of CTAs is perhaps easier. CTAs can earn fees on money-under-management, even if they have no ability to generate alpha. If they fail, they can easily restart, after erasing their prior history from the data bases. This suggests that they should have very high attrition rates, entering the industry to collect the fees while having difficulty surviving. We examine these hypotheses below.

5.3.1 Costly Information and a Lack of Common Knowledge

The information available to investors via formal offer documents ('prospectuses') and publicly available performance databases is fraught with biases and it is difficult to determine the true performance of CTAs. The databases that are available are not uniform, and not all of them allow for backfill bias to be corrected (see footnote 15, above). Evaluation of risk adjusted returns is further complicated by the absence of clear relevant benchmarks. The task of producing risk-adjusted performance evaluation of unbiased returns falls on the investor, a task that is costly, time consuming and requires analytical skills. Thus it is very costly for investors to have a view of CTAs that is different from that portrayed in Figure 1, where their performance looks very attractive. Given the available information, fraught with biases, and lack of relevant benchmarks, investors may simply believe that CTAs are a good investment.

But, costly information production can be only part of the story. For at least two decades investors have, on average, received poor returns. But, individually each investor may believe that his experience is simply bad luck. Since individual investors have no way of learning of the investment experience of other investors, information is never aggregated, so a true picture of the industry never emerges from actual experience. Although the publicity surrounding the EGR studies revealed poor performance, as discussed above, commodity fund mangers changed their organizational form, which may have allowed the industry to continue. The lack of aggregated actual experience of poor performance is necessary to keep investors from revising their view of Figure 1, which they may believe to be true. It seems unlikely that the same set of investors has been involved in investing in CTAs over that period. There must be new investors arriving, so that even when investors experience poor returns and withdraw their money, there are other investors willing to invest.

Table 9: A Comparison of CTA and Hedge Fund Fees.

The table summarizes the fixed and variable component of fees for CTAs and Hedge Funds. For each category the table gives the average and the standard deviation expressed in percent per annum.

		Management Fee	Incentive Fee
CTAs	**Average**	2.15	19.50
	Standard Deviation	1.22	6.32
Hedge Funds	**Average**	1.42	16.33
	Standard Deviation	0.51	6.84

Another, non-mutually exclusive, possible explanation could be the performance sensitivity of investors. Christoffersen and Musto (2002) analyze the money-fund industry and find that poorly-performing funds increase their fraction of performance-insensitive investors over time. They conclude that funds with bad performance should charge more from these investors, as their demand is price inelastic and a reduction in after-fee performance will not result in large outflow of money. Perhaps performance-insensitive investors in hedge funds and CTAs end up disproportionately at CTAs. If so, rational CTAs should charge higher fees. Are there differences in fees between hedge funds and CTAs? Table 9 provides a look at the data. CTAs do appear to have higher management fees and slightly higher incentive fees than hedge funds. This suggests that the demand for CTAs is possibly less performance-sensitive and more price inelastic. Investors might be investing in CTAs for perceived diversification benefits and mandates for alternative investments (e.g. pension funds) and end up staying invested even in the face of poor performance.

5.3.2 Entry and Exit of CTAs into Fund Management

We saw above that CTA managers make money, even if investors do not, on average. This creates an incentive for new CTAs to enter the industry. There is little cost to entering the industry, only registration, which costs an almost trivial amount.[26] There is no certification or any kind of screening. Poor performance results in liquidation of the fund, at little cost. Moreover, because a nonsurviving CTA can eliminate his history from the data bases retrospectively, there is no stigma to not surviving. It seems clear then that even if CTAs have no talent, in the sense of ability to generate alpha, they should persistently enter the industry because they earn fees on money-under-management, as long as they survive. If a CTA does not survive, he can restart with no history. We can shed some light on this by looking at CTA attrition rates.

Table 10: CTA Entry and Exit.

The table summarizes by year-end the number of funds in the Lipper-TASS database (after correction for backfill and survivorship bias), the percentage of firms disappearing in the subsequent 12 and 24 months, the excess returns of the firms exiting in the subsequent 12 months, and the relative performance of the exiting funds compared to the sample average in the year of exit.

Date	# Active funds	12-Month Attrition %	24-Month Attrition %	ER Exits Next 12 Month	% Exits below avg ret
1994	16	38	63	-3.74	66.67
1995	74	34	50	-12.20	60.00
1996	81	23	52	-18.86	52.63
1997	83	33	61	-26.01	70.37
1998	68	41	71	-19.10	42.86
1999	52	50	77	-17.25	50.00
2000	34	44	59	-12.98	46.67
2001	153	21	31	-10.71	56.25
2002	149	11	20	-19.38	76.47
2003	166	10	22	-21.65	56.25
2004	203	18	32	-7.51	58.33
2005	211	19	36	-8.80	57.50
2006	227	19	NA	-7.11	40.48

Table 10 shows CTA attrition rates and the excess net returns for the funds that exit. The second column reports the number of active funds as of the date in the first column. The third and fourth columns report the attrition rates for these funds over the next twelve months and twenty four months, respectively. During a year, some funds exit. “ER Exits” is the annualized excess return, with respect to all the reporting funds, for the funds that exit during the year indicated. The

[26] Prospective CTAs need to register with the National Futures Association. There is a nonrefundable fee of $200 and a fee of $85 for fingerprinting for each individual principal. See Hhttp://www.nfa.futures.org/registration/cta.aspH .

returns are calculated on a monthly basis, including all CTAs in existence during that month. The column entitled "% survivors below avg ret" shows the percentage of the exiting funds that had annualized returns below the equally-weighted annualized average of CTA month returns.

Table 10 indicates that CTA attrition rates are high, although they have declined recently. E.g., of the CTAs present on 31-Dec-05 (i.e., during 2006), 36 percent were gone by the end of 2007. Exiting funds are poor performers. E.g., during 2006, the exiting funds averaged an excess return of -8.8 percent annually.

In summary, the evidence is broadly consistent with the view that investors, facing high information costs with regard to evaluating CTA performance, believe that Figure 1 represents the investment opportunity. The poor performance experience of individual investors is not widely known. As a result investors continue to invest with CTAs and, recognizing this, CTAs continue to enter the industry, earning fees on money-under-management even though failure rates are very high.

6. Summary and Conclusions

Consumers and investors need information to rationally allocate their resources. Normally, we think of the price system as guiding these decisions. But, hedge funds are not publically traded, so there are no prices. There is only past performance data. In the case of hedge funds the available vendor data about their performance is biased, and there are few credible benchmarks for performance analysis. For these reasons, it has proven very difficult to evaluate the performance of hedge funds. These issues pose problems for investors as well as researchers as to whether hedge funds are an attractive asset class to invest in. They also potentially pose issues for public policy, to the extent that the hedge fund industry is sufficiently large to pose systemic risks.

We illustrate these issues by narrowing the universe of hedge funds to CTAs, because they are fairly homogeneous, their strategies are better known, and their strategy space is smaller. Using data between 1994 and 2007 from Lipper-TASS, we show that survivorship and backfill bias overstate the reported average return of CTAs by more than 8 percent per annum. Bias-corrected annualized average returns to investors were 4.9 percent, which is merely 85bp over the return on T-bills during this period. However, we estimate that gross average CTA returns (before fees) significantly exceed Tbill returns, which implies that funds retain most of their outperformance by charging fees. We propose simple dynamic futures-based trading strategies for performance evaluation. Because these strategies are in the public domain, they provide a natural hurdle that CTAs ought to overcome. Yet we find that the average CTA exhibits no skill (alpha) relative to these benchmarks. Combining our results with earlier studies by Elton, Gruber, and Rentzler, we conclude that poor CTA performance has persisted for at least twenty years. CTAs are a kind of market failure. Normally, asymmetric information is viewed as leading to an absence of a market. But, in the case of CTAs, the absence of information has led to the persistence of the market.

In Akerlof's (1970) celebrated work, lemons problems result in market failure, the market does not exist. Our results suggest that CTAs are lemons and that this lemons market can persist. How can the CTA market persist? In Akerlof's model, the information asymmetry is common knowledge; both sides of the market understand that there is an information asymmetry, namely, that car sellers have private information about the value of their cars. There is no way to signal car quality and because all cars sell at the same price there is an externality, namely, if a used car is sold some of the gains that should accrue to the sellers of good used cars accrue to the sellers of bad used cars. This causes sellers of good used cars not to enter the market. Buyers can rationally make this calculation, and they do not buy used cars, knowing that any car in the used car market is a lemon. If there is a trade in the used car market, the price – in the limit –will be the value of a lemon. Like the used car market, CTAs cannot signal ability, but in other respects the situation is fundamentally different for CTAs. There appears to be no common knowledge of an information asymmetry. There are no prices to convey information. Investors appear to believe that Figure 1 represents an accurate portrayal of the performance history relative to a benchmark. As EGR point out, there is also an inability to short CTAs. Somewhat paradoxically the market for CTAs appears to be an example of a persisting lemons market. We argue that CTAs persist as an asset class despite their poor performance, because they face no market discipline based on credible information. There is no required disclosure as with SEC filings for firms or bank Call Reports. There is no regulation like that for mutual funds or banks. There are no private institutions that certify the managers' competence (like the American Medical Association for doctors), or that certify their performance (like the Good Housekeeping seal of approval), and, as we have seen, no private repository of credible information for comparison purposes. Further, investors' individual experience of poor performance is not common knowledge. In such a setting, it seems that some people can be fooled all of the time.

Appendix: Construction of Normative Benchmarks

Foreign Exchange

The literature on the forward bias in currency markets is the among the earliest studying "anomalies" in financial markets, and dates back to Rogoff (1979) and Bilson (1981) – see also Froot and Thaler (1990) for a summary of this early literature. The market for FX is a natural place to look for CTA trading strategies, as the futures market is large, and the surge in market activity since 2001 corresponds to a period when interest rate differentials favored investments in high interest rate currencies financed by short positions in low interest rate currencies (see Galati and Melvin (2004) on the increase in FX trading activity). Or, alternatively said, the environment favored the "carry trade," in which an investor borrows in the low interest rate currency, and takes a long position in a high interest rate currency, speculating that the exchange rate will not change so as to offset the interest rate differential. Galati and Melvin (2004) show that FX turnover growth increases in interest rate differentials and with the magnitude of prior year's exchange rate changes.[27]

In order to construct FX factors we employ the data for spot and one month forward prices against the US dollar for 15 currencies.[28] The excess return from the end of the month *t* to the next is calculated as $\frac{S_{t+1} - F_{t,t+1}}{F_{t,t+1}}$, where $F_{t,t+1}$ is the forward price at time *t* on a contract that expires at the end of month *t+1*,[29] and S_{t+1} is the spot price at time t+1. The basis at the end of the month is defined as the difference between the current spot price and the current one month ahead forward, expressed as a ratio to the current spot price: $\frac{S_t - F_{t,t+1}}{S_t}$.

At the end of each month, we construct currency basis portfolios by ranking all currencies on their basis (interest rate differential) relative to the US dollar. Currencies in the top half of this ranking are assigned to the high basis portfolio and the bottom half of the currencies to the low basis portfolio. Both portfolios are equally weighted. The positions are rebalanced monthly so that the high (low) basis portfolio represents a dynamically rebalanced portfolio of currencies with the highest (lowest) interest rate differential relative to the US dollar.

Currency momentum portfolios are similarly constructed by ranking currencies by their on prior 3-month excess returns. At the end of each month currencies are assigned to High and Low

[27] Also, see Galati and Heath (2007) and Galati, Heath and McGuire (2007).

[28] The currencies used are: AUD (Australian Dollar), CAD (Canadian Dollar), CHF (Swiss Franc), DKK (Danish Kroner), DEM (Deutsche Mark), EUR (Euro), FRF (French Frank), GBP (British Pound), IEP (Irish Pound), ITL (Italian Lira), JPY (Japanese Yen), NLG (Netherland Guilder), NZD (NZ Dollar), NOK (Norwegian Kroner), and SEK (Swedish Kroner).

[29] For some cases the forward contract trading at time t does not expires at the last day of next month, for example if the last day of next month is a Friday the forward contract might expire on the Monday. In such cases an interpolation rule following interest parity is used to figure out the forward price $F_{t,t+1}$. Let $F_{t,t+k}$ be the forward price at time *t* on a contract that expires on *t*+*k*, and we are interested in $F_{t,t+j}$ which corresponds to the last day of next month, then:

$$\ln\left(F_{t,t+j}\right) - \ln\left(S_t\right) = \left(\frac{j}{k}\right) * \left\{\ln\left(F_{t,t+k}\right) - \ln\left(S_t\right)\right\}.$$

currency momentum portfolios, which are equally-weighted and held for one month subsequent to ranking after with time they are rebalanced.

Commodities

The information content of the futures basis for expected risk premia has been documented empirically by Fama and French (1987) and more recently by Erb and Harvey (2006) and Gorton and Rouwenhorst (2006). Commodity price momentum has been documented by Pirrong (2005), Erb and Harvey (2006), Miffre and Riallis (2007), and Gorton, Hayashi and Rouwenhorst ((2007). Our construction of dynamic commodity portfolios mirrors Gorton, Hayashi, and Rouwenhorst (2007) who argue that the excess returns to portfolios sorted by the basis and prior returns in part stems from selecting commodities when inventories are low. Based on the Theory of Storage, GHR show that prior returns ("momentum") and the futures basis ("backwardation") are price-based signals of low physical inventory levels, and the excess returns are a compensation for the increased volatility of commodity prices.

At the end of each month, available commodities futures are ranked on the basis, defined as the annualized slope of the futures curve between the nearest and the next-to-nearest to maturity contracts. High and low basis portfolios are constructed from the top and bottom half of the commodities in this ranking. All portfolios are equally weighted and rebalanced monthly. Similar to the Basis portfolios, we construct monthly rebalanced equally-weighted High and Low Commodities momentum portfolios by ranking commodities on prior 1-year returns.

Equities

We construct a momentum and a value factor by sorting country index returns on prior return and book-to-market. Momentum in country equity index returns has been documented by Asness et al. (1997), and Chan et al. (2000), and studied more recently by Bhojraj and Swaminathan (2006). The profitability of value strategies has been documented by Asness (1997). Related papers on performance reversals include Richards (1997) and Balvers et al. (2000).

At the end of each month we sort available country equity index futures by country-wide measures of book-to-market (Value) or prior 12-month return (Momentum). For each of these sorts we construct High and Low Value and Momentum portfolios containing the top and bottom half of constituents of this ranking.

References

Ackermann, Carl, Richard McEnally, and David Ravenscraft (1999), "The Performance of Hedge Funds: Risk, Return, and Incentives," Journal of Finance 54(3), p. 833-874.

Akerlof, George (1970), "The Market for 'Lemons": Quality Uncertainty and the Market Mechanism," Quarterly Journal of Economics 84(3), p. 488-500.

Arrow, K. J. (1971), Essays in the Theory of Risk-Bearing (Chicago: Markham Publishing Co.).

Asness, Clifford S. (1997), "The Interaction between Value and Momentum Strategies," Financial Analyst Journal 53(2), pp 29-36

Asness, Cliff S., John M. Liew, and Ross Stevens (1997), "Parallels Between the Cross-Sectional Predictability of Stock and Country Returns," Journal of Portfolio Management (Spring), p.79-87.

Balvers, Ronald, Yangru Wu and Erik Gilliland (2000), "Mean Reversion across National Stock Markets and Parametric Contrarian Investment Strategies," Journal of Finance, 55(2), pp. 745-772.

Barberis, Nicholas and Ming Huang (2007), "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices," Yale School of Management, working paper; American Economic Review, forthcoming.

Bhojraj, Sanjeev and Bhaskaran Swaminathan (2006), "Macromomentum: Returns Predictability in International Equity Indices," Journal of Business 79(1), p. 429-451.

Bilson, John (1981), "The Speculative Efficiency Hypothesis," Journal of Business 54, 433-451.

Brown, S. and W.N. Goetzmann (1997), "Mutual Fund Styles," Journal of Financial Economics 43(3), p. 373-99.

Brown, Stephen and William Goetzmann (2003), "Hedge Funds with Style," Journal of Portfolio Management 29(2), p. 101-112.

Brown, Stephen J., William N. Goetzmann, Roger G. Ibbotson (1999), "Offshore Hedge Funds: Survival and Performance 1989-1995," Journal of Business 72, p. 91-118.

Brown, Stephen J., William N. Goetzmann, Roger G. Ibbotson, and Stephen A. Ross (1992), "Survivorship Bias in Performance Studies," Review of Financial Studies 5(2), p. 553-580.

Brown, Stephen J., William N. Goetzmann, and Bing Liang (2004), "Fees on Fees in Funds of Funds," Journal of Investment Management 2, 39-56.

Chan, Kalok, Allaudeen Hameed and Wilson Tong (2000), "Profitability of Momentum Strategies in the International Equity Markets," The Journal of Financial and Quantitative Analysis 35(2), pp. 153-172.

Christoffersen, Susan and David Musto (2002), "Demand Curves and the Pricing of Money Management," Review of Financial Studies 15, 1499-1524.

Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz (2008), "Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation," working paper Yale School of Management.

DeRoon, Frans, Jenke TerHorst, and Theo Nijman (2003), "Evaluating Style Analysis," Journal of Empirical Finance vol. 11, no. 1 (January), p. 29-53.

Edwards, Franklin and M.O. Caglayan (2001), "Hedge Fund Performance and Manager Skill," Columbia University, working paper.

Edwards, Franklin and Jimmy Liew (1999), "Managed Commodity Funds," Journal of Futures Markets Vol. 19, No. 4, p. 377-411.

Elton, Edwin, Martin Gruber, and Joel Rentzler (1990), "The Performance of Publicly Offered Commodity Funds," Financial Analysts Journal (July/August), p. 23-30

Elton, Edwin, Martin Gruber, and Joel Rentzler (1989), "New Public Offerings, Information, and Investor Rationality: The Case of Publicly Offered Commodity Funds," Journal of Business 60(2), p. 175-199.

Elton, Edwin, Martin Gruber, and Joel Rentzler (1987), "Professionally Managed, Publicly Traded, Commodity Funds," Journal of Business 60(2), p. 175-199.

Erb, Claude and Campbell Harvey (2006), "The Strategic and Tactical Value of Commodity Futures," Financial Analysts Journal 62: 69-97.

Fama, Eugene F. and Kenneth R. French (1993), "Common Risk Factors in the Returns on Stocks and Bonds," Journal of Financial Economics 33, p. 3-56.

Fama, Eugene and Kenneth French (1987), "Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage," Journal of Business 60: 55-73.

Fama, Eugene F. and Kenneth R. French (1992), "The Cross-Section of Expected Stock Returns," Journal of Finance 47(2), p. 427-465.

Farhi, Emmanuel and Xavier Gabaix (2008), "Rare Disasters and Exchange Rates," NBER Working Paper No. 13805.

French, Kenneth (2008), "The Cost of Active Investing," Dartmouth College, working paper, Journal of Finance, forthcoming.

Froot, Kenneth A., and Richard H. Thaler (1990), "Anomalies: Foreign Exchange," The Journal of Economic Perspectives, 4(3), 179-192.

Fung, William and David Hsieh (1997), "The Information Content of Performance Track Records: Investment Style and Survivorship Bias in the Historical Returns of Commodity Trading Advisors," Journal of Portfolio Management, 24(1), p. 30-41.

Fung, William and David Hsieh (2000), "Performance Characteristics of Hedge Funds and Commodity Funds: Natural vs. Spurious Biases," Journal of Financial and Quantitative Analysis vol. 35(3), p. 291-307.

Fung, William and David Hsieh (2001), "The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers," Review of Financial Studies 14(2), p. 313-341.

Fung, William and David Hsieh (2002a), "Asset-Based Style Factors for Hedge Funds," Financial Analysts Journal 58(5), p. 16-27.

Fung, William and David Hsieh (2002b), "Hedge Funds: An Industry in Its Adolescence," Federal Reserve Bank of Atlanta Economic Review (Fourth Quarter), p. 1 - 34.

Fung, William and David Hsieh (2004), "Hedge Fund Benchmarks: A Risk-Based Approach", Financial Analysts Journal, 60(5) p. 65-80.

Galati, Gabriele and Michael Melvin (2004), "Why has FX Trading Surged? Explaining the 2004 Triennial Survey," Bank for International Settlements, Quarterly Review (December), p. 67 - 74.

Galati, G, A Heath and P McGuire (2007): "Evidence of Carry Trade Activity", Bank for International Settlements, Quarterly Review (September), pp 27–41.

Galati, Gabriele and Alexandra Heath (2007), "What Drives the Growth in FX Activity? Interpreting the 2007 Triennial Survey," Bank for International Settlements Quarterly Review (December), p. 63 - 72.

Gorton, Gary and K. Geert Rouwenhorst (2006) "Facts and Fantasies about Commodity Futures," Financial Analysts' Journal 62(2), p. 47-68.

Gorton, Gary, Fumio Hayashi and K. Geert Rouwenhorst (2007), "The Fundamentals of Commodity Futures Returns," http://papers.ssrn.com/sol3/papers.cfm?abstract_id=996930, working paper.

Hall, Dustin (2008), "The Elephants in the Room: Dangers of Hedge Funds in Our Financial Markets," Florida Law Review 60, p. 183-227.

Hasanhodzic, Jasmina and Andrew Lo (2007), "Can Hedge Fund Returns Be Replicated? The Linear Case," Journal of Investment Management 5(2), p. 5-45.

Jensen, Michael (1968), "The Performance of Mutual Funds in the Period 1945-1964," Journal of Finance 23(2), p. 389-416.

Levy, Haim and Sarnat, Marshall (1984), Portfolio and Investment Selection: Theory and Practice (Prentice-Hall International).

Liang, Bing (2000), "Hedge Funds: The Living and the Dead," Journal of Financial and Quantitative Analysis 35, p. 309-326.

Malkiel, Burton and Atanu Saha (2005), "Hedge Funds: Risk and Return," Financial Analysts Journal 61(6), p. 80-88.

Markowitz, H. (1952), "The Utility of Wealth," Journal of Political Economy 60(2), p. 152–158.

Miffre, Joëlle and Georgios Rallis (2007), "Momentum in Commodity Futures Markets," Journal of Banking and Finance 31(6), p. 1863- 1886.

Mulvey John., S.S.N. Kaul and Koray Simsek (2004), "Evaluating a Trend-Following Commodity Index for Multi-Period Asset Allocation," Journal of Alternative Investments 7(1), p. 54-69.

Park, James (1995), "Managed Futures as an Investment Asset," PhD. Dissertation, Columbia University.

Pirrong, Craig (2005), "Momentum in Futures Markets," Bauer College of Business, University of Houston, working paper.

Polkovnichenko, V. (2005), "Household Portfolio Diversification: A Case for Rank-Dependent Preferences," Review of Financial Studies 18, 1467–1502.

Posthuma, Nalke and Pieter Jelle van der Sluis (2003), "A Reality Check on Hedge Fund Returns," available at SSRN: http://ssrn.com/abstract=438840 .

Richards, Anthony J. (1997), "Winner-Loser Reversals in National Stock Market Indices: Can They be Explained?," Journal of Finance 52(5), pp 2129-2144.

Rogoff, Kenneth S. (1979), "Essays on Expectations and Exchange Rate Volatility," PhD. Dissertation, Massachusetts Institute of Technology.

Rosenthal, Leonard, and Colin Young (1990), "The Seemingly Anomalous Price Behavior of Royal Dutch/Shell and Unilever N.V./PLC," Journal of Financial Economics 26, 123-141.

Sharpe, William (1992), "Asset Allocation: Management Style and Performance Measurement," Journal of Portfolio Management (Winter), p. 7-19.

Siri, Erik and Peter Tufano (1998), "Costly Search and Mutual Fund Flows," Journal of Finance 53(5), 1589-1622.

Waksman, Sol (2000), "Commodity Trading Advisor Survey: Adding Equities to A Managed Futures Portfolio," Journal of Alternative Investments (Fall), p. 43-44.