Master Thesis October 29^{th}, 2004

MBF, HEC Lausanne

A Quantitative Analysis of CTA Funds

Simon Vuille, Corneliu Crisan[1]

Abstract

Our research studies various properties of commodity trading advisors (CTAs) from a quantitative point of view. Our investigation is based on a commercial database of 549 funds and focuses on the period 1990 to present. Firstly, CTAs' return distributions are analyzed and strong evidence of non-normality is found. Secondly, relative persistence in return distribution parameters is studied. We review the major benchmarks available to the industry and build new benchmarks from our dataset. This allows us to infer the magnitude of various biases. We study homogeneity of 2 CTA subsets, namely trend-followers and non-trend-followers, and study the diversification possibilities in a CTA portfolio. In the second part of the study, we focus on linking CTAs returns with that of traditional assets. After showing that a "buy and hold" multi-factor linear model fails to explain CTAs returns, we point out the presence of option-like payoffs in CTAs return patterns. Lastly, using simple trading algorithms based on moving averages, we propose a linear model in which factors capture the dynamic nature of CTA managers' strategies. Our model leads to significant improvements over the classical model.

Keywords: Commodity Trading Advisors, Trend-Followers, Performance Persistence, Market Model

JEL Classification: C40, G10, G11

[1]Simon Vuille & Corneliu Crisan, Master in Finance and Banking, University of Lausanne. We wish to thank our supervisor, Professor François-Serge Lhabitant for his precious help and Professor Fred Feinberg for some Matlab hints. We are also grateful to Dr. Bart Janssen, Dr. Alexander Passow, and Bruno Veras De Melo from Gottex, for their continuous support and insightful comments. Many thanks to Magda Bogusz and Simon Hogg for lending us their English skills. The usual disclaimer applies. Correspondence: svuille@umich.edu

Executive Summary

Commodity trading advisors (CTAs) are professional money managers who invest their clients' funds using global futures and options markets as a medium. Often referred to as "managed futures", CTAs take part in all liquid futures markets, and as such, offer investors an efficient way of gaining exposure to markets otherwise not easily accessed.

Literature has pointed out the positive effects on the efficient frontier derived from holding CTAs in a diversified portfolio. More recently, it has been argued that the peculiar nature of CTAs' return patterns can bring additional diversification effects that are not fully captured by a mean-variance approach. Authors have suggested that CTAs are characterized by both nonlinear correlations with traditional assets, and positive skewness. It is thought that such features can be used advantageously for the purpose of downside protection and capital preservation. Aware of CTAs' potential benefits, investors have been increasingly trusting CTA managers. Assets under management in the industry have soared to an estimated 80 billion, with the most recent inflows coming from funds of hedge funds as well as wealthy investors who renewed their confidence in CTAs.

Our research studies various properties of CTAs from a quantitative point of view. The investigation is based on a commercial database of 549 funds and focuses on the period 1990 to present.

The first goal of the study is to shed light on the exact nature of CTAs' return distributions. Analyzing CTAs' historical return distributions, we find strong evidence of non-normality. We find that CTAs are characterized by both positive skewness and excess kurtosis. This stresses the need for portfolio allocation techniques that account for higher order moments, such as Omega. Additionally, we find that the presence in our dataset of various biases (survivorship, instant history and selection) renders the formulation of an estimate for CTAs long run expected return very difficult. We proceed to study the autocorrelation structure characterizing CTAs. We find the presence of slightly negative autocorrelation for a 2 months lag.

In the second part of the research, we try to answer the following question: to what extent is it reasonable to rely on parameters of past return distributions to infer funds' future performance and risk profile. Using contingency tables, we look for the presence of relative persistence in various return distribution parameters, both in the short and long term. While standard deviation is found to persist even for long term time horizons, we fail to find evidence of persistence for measures such as average return, sharpe ratio, skewness and kurtosis. These findings lead us to think that analysis of CTA funds should focus on understanding the underlying risks, rather than seeking potential out-performers in an asset class where returns are highly volatile and relative performance unstable.

The next analysis we run is that of historical risk premium characterizing CTAs. This analysis uncovers an interesting fact. While risk premium seems to be generally low for CTAs, we fail to identify a significant risk premium for non-trend-followers category.

In the last part of the research, we focus on linking CTAs' returns with that of traditional assets. We first show that a "buy and hold" multi-factor linear model fails to explain CTAs' returns, despite a slight exposure of CTAs to bond and commodity markets. Then, we point out the presence of non-linear, option-like payoffs in CTAs' return patterns. Trend-followers are characterized by straddle-like payoffs, while non-trend-followers exhibit payoff structures reminiscient of simple call options on major asset classes. Especially for non-trend-followers, this finding confirms the strong potential of CTAs as a mean to achieve capital preservation in times of market turmoil. In a second step, we develop a linear model in which much of the dynamism of CTAs' trading strategies is built into the return-generating factors. Using simple trading algorithms based on moving averages, our indices model trend-following trading, basic stop-loss rules, pyramidal trading schemes and capital allocation. The model leads to significant improvements over the classic "buy and hold" model. We find it possible to replicate a CTA index out of sample using only 5 global trend-following indices, even when the in sample calibration period is small. Studying the historical exposure of the CTA universe to our 5 indices, we find

that long term interest rates make up most of CTAs' exposure. Exposure to currency and commodity trends share the second place, while exposure to stock trends has diminished monotonically for the past 10 years. Lastly, we apply 2 selection methods which use our model as a starting point. Both attempts are unsucessful at picking outperforming funds. Nevertheless we find evidence that CTAs with low overall exposure to market trends are characterized by very low returns, and should be avoided.

Contents

List of Figures

List of Tables

1 Introduction

Commodity trading advisors (CTAs) are professional money managers who invest their clients' funds using global futures and options markets as a medium. Historically, CTAs were limited to trading commodity futures, hence the nomenclature. Nowadays, highly liquid futures markets exist for interest rates, bonds, stock indices, currencies, precious metals, energy, and agricultural products. Often referred to as "managed futures", CTAs take active part in all of them. Through their ability to take both long and short positions, CTAs offer investors an efficient way of gaining exposure to markets otherwise not easily accessed.

From a legal point of view, and according to the U.S Commodity Exchange Act (Title 7, Chapter 1, Section 6n), CTAs have to register with the Commodity Futures Trading Commission. For firms located outside of the U.S.A, similar obligations exist (for example, the Commodity Investment Regulations in Japan). CTAs are typically organized as Limited Parternships and have offshore structures reminiscent of the ones set up by hedge funds.

1.1 CTA classification

CTAs can be classified along two dimensions: the markets they trade in, and the techniques on which their trading strategies rely.

With respect to the markets traded, CTAs are either fully diversified or focused on specific markets. Whereas diversified CTAs sometimes claim to trade as many as 300 different futures contracts, it is safe to assume that, given the liquidity constrains faced, positions are taken in only 50-100 contracts on a regular basis. Non-diversified CTAs specialize in a particular market, or a set of related markets. The following is a non-exhaustive list of markets for which specialized CTAs exist: currencies, agricultural commodities, precious metals, energy, stocks.

The trading approach for CTAs can be classified as systematic or discretionary, even though some CTAs base their actions on a mix of the two. Systematic approaches rely on quantitative models to perform technical or fundamental analysis, generating buy or sell signals. Purely systematic approaches are fully computerized. In such cases, the role of the CTA is to fine-tune the model, keep it up to date, or develop additional models in order to cope with the evolution of the financial markets. Non-systematic CTAs, also known as discretionary traders, base their strategies on fundamentals and underlying economic factors. Since experience is key for discretionary traders, they often specialize on a particular sector or market.

Most trading systems used by CTAs can be classified as either trend-following or counter trend-following. Trend-following is by far the most widespread strategy among CTAs. Often fully automated, such programmes tend to be diversified across a range of markets. Most trend-followers refrain from trying to predict trends, and rather take positions that will benefit if the current market trend persists. Trend-followers look at various indicators in order to eliminate market noise and find the current direction of a market. Widespread indicators include moving averages, exponential smoothing and momentum. Trend-followers differ from each other with respect to the time horizon they use to determine the existence of a trend. Funds focus on short, medium, or long-term trends, or a combination thereof. Counter-trend systems look for trend reversals using methods such as rate of change indicators (oscillators, momentum) or head and shoulders patterns. The use of trading systems relying on highly quantitative techniques, such as neural networks, genetic algorithms, or chaos theories has also generated much interest in the recent past.

Risk management is a key part of any trading strategy, and most systematic CTAs will typically cut losses as soon as they materialize, while they will try to let the profits run, often adding to winning trades. Additionally, various filters will be applied to the signals in order to determine capital allocation. Such filters include volatility, volume, as well as various forms of risk/reward ratios.

Figure 1, adapted from Habib (2004) summarizes the different approaches used by CTAs.

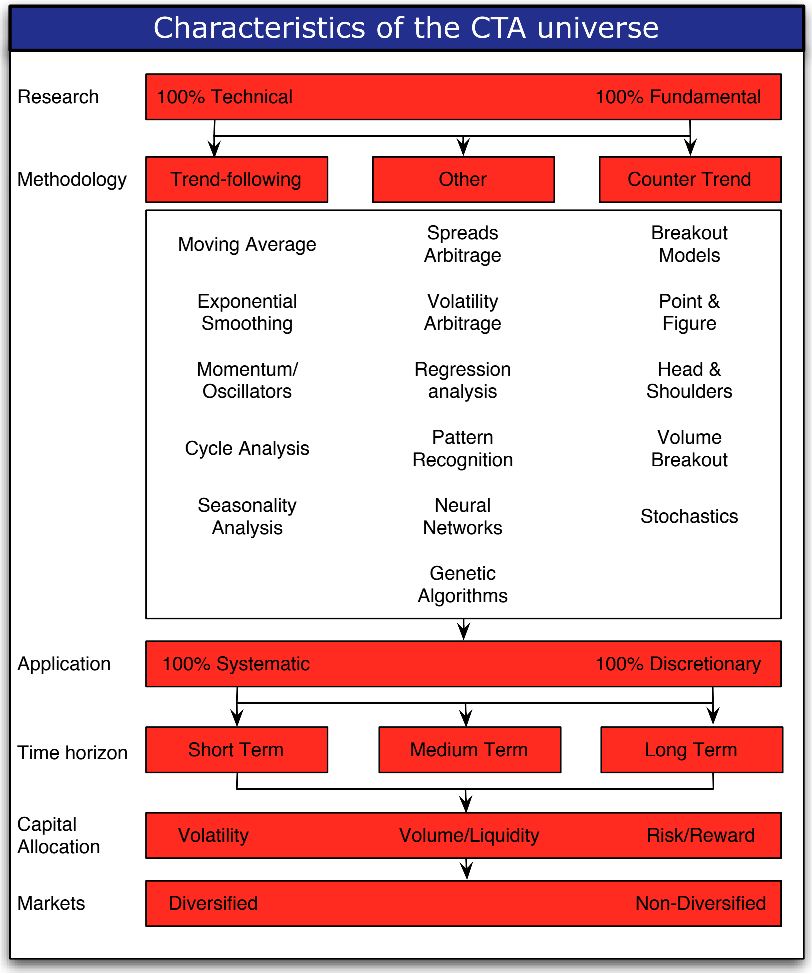

Figure 1: A model of the different investment styles among CTAs

1.2 Market growth and prospects

Futures and options have been used for centuries, both as a risk management tool and a return enhancement vehicle. However, managed futures as an investment alternative have been available only since the late 1960s. More recently, institutional investors such as pension funds, endowments and trusts, family offices and funds of hedge funds have been including managed futures in their portfolios.

As shown in Figure 2, the dollars under management in the CTA industry have experienced tremendous growth during the past few years. Even though we lack precise numbers on the state of the assets under management in the industry as a whole, some figures report as much as a 100 billions being currently managed by CTA funds.

After a very good year in 2002, CTAs have attracted large amounts of funds. High net worth individuals have renewed their confidence in managers after having experienced disillusion in the stock market. In addition, improvements in the integrity and safety of trading in organized exchanges for futures/options contracts have provided further assurances of investor's safety. Currently, funds of hedge funds represent the major source of inflow in the industry. Funds of hedge funds are allocating money to managed futures as one of their strategies and see CTAs as a subset of hedge funds. Other, more conservative investors like pensions and endowments are showing interest for CTAs, despite a perception of extreme risk still being associated with futures strategies.

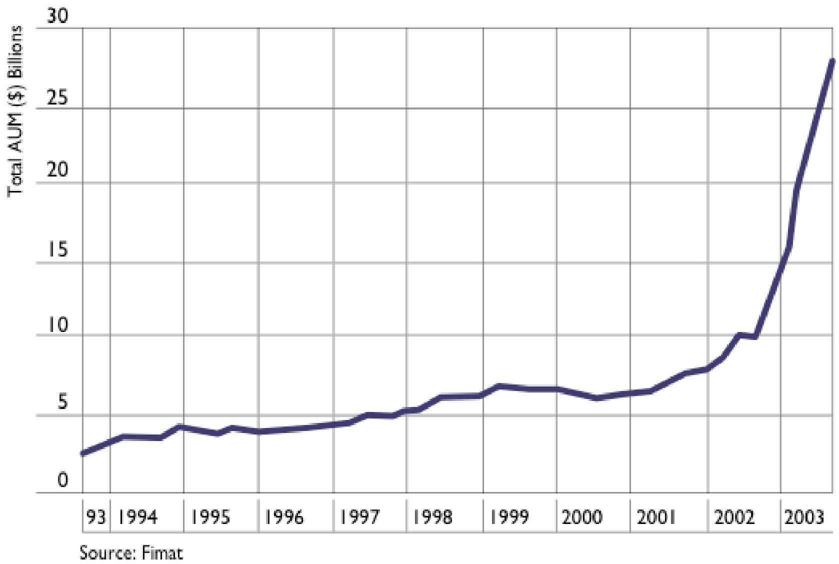

Figure 2: Assets under management growth for the CSFB/Tremont CTA index

The recent growth of the managed futures industry has not gone without raising questions concerning the industry's capacity to absorb new inflow. Unlike some other alternative investment strategies though, it is estimated that 75% of CTAs' assets are in financial futures or currency markets, where even big trades have relatively little impact.

1.3 Benefits of managed futures

The growth in demand for managed futures products indicates investor's appreciation of the potential benefits of CTAs. Numerous studies have been done on the subject of managed futures, especially with respect to the diversification effects they have on portfolios of various types of assets. When added to stocks or bonds, CTAs have been shown by CISDM (1999) to positively affect the efficient frontier, as depicted in Figure 3. It has been argued that the peculiar nature of CTAs' return patterns have the potential to bring additional diversification effects that are not fully captured by a mean-variance approach. Specifically, Cerrahoglu (2004) shows that correlations between CTAs and stock markets are positive in bull markets and negative in bear markets. While it is not clear why trends and other profit opportunities tend to develop when stock markets are experiencing turmoil, this feature can be used advantageously in the context of portfolio construction for the purpose of downside protection and capital preservation. Lastly Kat (2004) introduces the possibility of combining both CTAs and hedge funds in a portfolio. In this setup, CTAs' positive skewness was shown to be helpful in reducing the impact of negative skewness, which is otherwise a problem with hedge funds strategies.

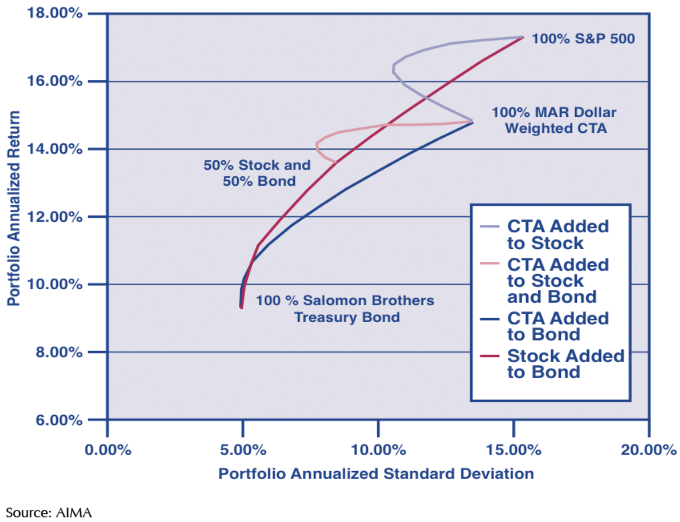

Figure 3: The diversification effects of adding CTAs to stocks and bonds

2 The data

We base our study on Barclay's database of 549 CTAs as of July 2004. This data was kindly provided by Gottex Fund Management. The database contains monthly returns, funds under management, as well as a short description of the strategy used for each CTA. The strategy description is usually provided by the partners of the funds themselves and thus, should be treated with caution. The oldest fund in the database started its operation in 1975 and most of the returns series stop in June 2004. Of the 549 CTAs, some differ only by the currency in which they report their returns (ie. some funds report returns in more than one currency), or by the amount of leverage used (several funds offer different levels of leverage for the same underlying strategy). In order to keep a universe of truly unique funds, we choose to keep only the US dollar denominated funds when more than one currency is available, and to keep only the lowest leverage level for companies offering multiple leverage alternatives. This leads to a final sample size of 498 funds.

2.1 Subsamples

Based solely on the funds description, we build 2 subsamples, one for the funds that acknowledge using trend-following or momentum strategies, and one for the remaining funds, that we call non-trend-followers. The number of funds in the 2 subsamples is respectively 272 and 226. The funds that declare using both trend-following and contrarian strategies are, nevertheless, classified as trend-followers. If throughout this study we use any other subset of the data, we will mention it, as well as the filtering rules applied and the resulting size of the subset.

2.2 Biases

It is important to realize the presence of several biases in our data set. Unlike other types of investment vehicles who have disclosure requirements that forces them to report about their activities to regulatory authorities, CTAs usually report to database vendors on a voluntary basis. This leads to several issues which make it difficult to infer the properties of CTAs as an asset class from a commercial database.

2.2.1 Survivorship bias

First and foremost, the data is subject to survivorship bias. Since database vendors discard funds that stopped reporting to them, all the funds for which we have data are funds that are still alive today. The rationale for discarding dead funds seems to be that subscribers of such database services are only interested in funds accepting new capital. A fund can stop reporting for several reasons. Possibly, poor performance led to an outflow of funds, leaving the manager with no other choice but to stop its operation because the fees no longer cover the operational costs. Since this is the most likely reason for a fund to stop reporting, the average return of the funds present in the database is an upward biased estimate of the real score, what's referred to as survivorship bias. Sometimes though, it is because of a totally different reason that a fund disappears from a database. A manager may think that assets under management's size has reached its maximum for the particular strategy implemented and may decide to close the fund to new investors in order not to hurt the incumbents (over-capacity concern). In this case, the sign of the survivorship bias is more difficult to determine. Let us mention that survivorship bias is not unique to CTAs or to other classes of hedge funds, but that it is also an issue when dealing with mutual funds, as pointed out by Brown, Goetzmann, Ibbotson, and Ross (1992).

Several studies of the importance of survivorship bias on returns have been conducted. Fung and Hsieh (1997b) find a survivorship bias of 3.6% per year. Schneeweiss, Spurgin, and McCarthy (1996) find an estimate of 1.4% bias a year. We are not aware of any study that looks at the survivorship bias potentially present in standard deviation or higher moments of CTAs' return distributions.

2.2.2 Instant history bias

Instant history bias is a consequence of the usual incubation period of CTAs. Since an impressive track record is a key element in order for a manager to attract funds, it is a common practice for managers to start trading with friends and relatives' funds. The problem is that when the manager decides to report to a database vendor, typically after having achieved good performance for a certain number of months, the vendor will back fill the database with the incubation period returns, creating instant history bias. Brown, Goetzmann, and Park (1997) Park (1995) both find an average incubation period of 27 months. Based on this, Fung and Hsieh (1997b) estimate instant history bias to be 3.6% per year.

2.2.3 Selection bias

Since reporting to database vendors is done by managers on a voluntary basis, it is likely that only funds with good performance want to be included in a database. This leads to the possibility of a selection bias, which, again, means that a database is not truly representative of the overall universe of funds available for investment.

3 CTAs' return distribution properties

In order to make the reader comfortable with CTAs' return distributions, we present here a preliminary analysis of some basic properties of CTA funds' returns. We start by looking at the first four moments of the distributions for our sample of 498 funds. We then proceed to the analysis of the funds' autocorrelation coefficients. We run the same analysis for our 2 subsamples of trend-followers and non-trend-followers. We relegate the figures pertaining to the analysis of the 2 subsamples to the Appendix, and mention only the properties that differ from those found for the whole sample of funds.

3.1 First 4 moments

Here we compute average return, standard deviation, skewness and kurtosis for each of the 498 funds, regardless of when they started their operations or of how long they have been in business. Figure 4 below presents a histogram for each moment. Average return and standard deviation are expressed in yearly terms. The average return for our sample is 14.5%, with only few funds having negative returns. The average standard deviation is 21%. The average skewness is, as expected, slightly positive at around 0.5. Most CTAs exhibit leptokurtic return distributions, the average kurtosis falling in the $4-5$ area. Let us also mention the presence of outliers on the histograms. These outliers are mainly funds for which only few data points are available, and they should not be regarded as very significant scores. In an attempt to correct for potential biases, and since outliers may have a large effect on skewness and kurtosis, we compute for the 4 moments the average value for all significant scores using a confidence level of 5%. The significance for skewness and kurtosis can be ascertained using the following formulas for their standard deviations:

$$\sigma_{skew} = \sqrt{N^{-1}[\frac{\mu_6}{\mu_2^3} - \frac{3\mu_5\mu_3}{\mu_2^4} - 6\beta_2 + 9 + \frac{\beta_1}{4}(9\beta_2 + 35)]} \tag{1}$$

$$\sigma_{kurt} = \sqrt{N^{-1}[\frac{\mu_8}{\mu_2^4} - \frac{4\mu_6\mu_4}{\mu_2^5} - \frac{8\mu_5\mu_3}{\mu_2^4} + 4\beta_2^3 - \beta_2^2 + 16\beta_1\beta_2 + 16\beta_1]} \tag{2}$$

With:

$$\mu_i = N^{-1}\sum_{n=1}^{N}(R_n - \bar{R})^i$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

In the formulas above, N stands for the number of data points and R_n for the return achieved by the fund during the nth period of its life.

Finally, let us emphasize again that the figures presented here are valid for our dataset only, and would have to be corrected for survivorship, instant history, as well as for selection biases in order to characterize CTAs as an asset class. See Fung and Hsieh (1997b) for a discussion of these biases and an estimation of their importance in the context of CTAs.

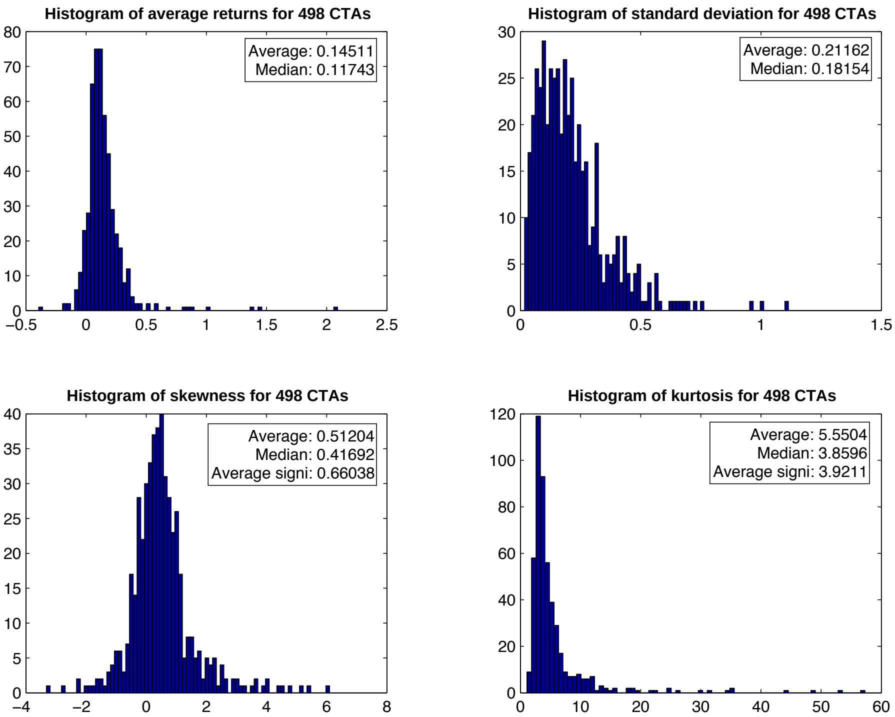

Figure 4: Histograms of first 4 moments of CTAs' return distributions

3.1.1 Trend-followers vs. non-trend-followers

Here we compute the 4 first moments for our 2 subsamples of trend-followers and non-trend-followers. The results can be seen in Figures 33 and 34 in the Appendix. These results do not differ widely from the analysis of the full sample of CTAs as far as average returns are concerned. In terms of standard deviation, it seems that trend-followers are slightly more risky than non-trend-followers, with an average standard deviation of 23% versus 18.5% for non-trend-followers. This greater standard deviation seems to be compensated by a slightly higher average skewness (0.54 vs. 0.48), as well as a lower average kurtosis (5.1 vs. 6), two attributes that are generally deemed to be desirable from the point of view of investor preference theory.

3.2 Autocorrelation

Next we turn to an analysis of autocorrelation coefficients for our sample of CTA funds. Autocorrelation refers to the correlation of a variable with itself over successive time intervals. It may affect the way various analysis are conducted. More specifically, strong positive/negative autocorrelation can lead to underestimating/overestimating yearly standard deviation when scaling monthly standard deviation figures using the square root of time rule. Indeed, the presence of autocorrelation brakes the assumption on which the rule relies, namely that returns are identically and independently distributed. For this analysis, we compute the autocorrelation for each fund in our sample, for lags between 1 and 8 months. Figure 5 shows 8 histograms, one for each lag. The histograms also report the percentage of significant coefficients at a confidence level of 5%, the average of the significant coefficients, as well as the average of all coefficients for a particular lag. One can clearly see from figure 5 that for a lag of 1 month, the frequency of significant coefficients is about two times the theoretical frequency (10.6% vs. 5%). The coefficients' distribution is widely spread around an average coefficient that is very close to 0. This means that, even though some funds show autocorrelation, the autocorrelation analysis should be done on a fund by fund basis. For a one month time horizon some funds show strong positive coefficient and some funds strong negative coefficient. CTAs as an asset class should not exhibit autocorrelation at all

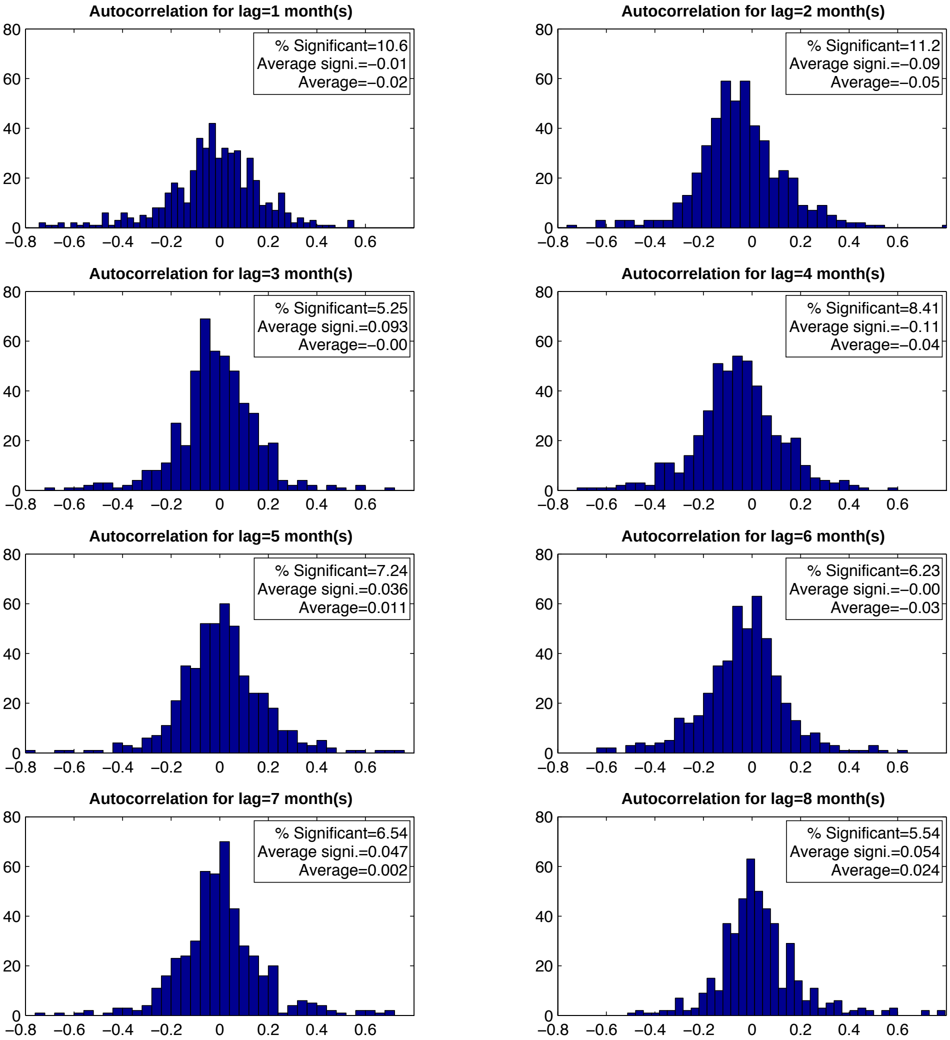

Figure 5: Histograms for CTAs' autocorrelation coefficients for lags between 1 and 8

for this lag. For a lag of 2 months, the frequency of significant coefficients goes up to reach its maximum of 11.2%, much higher than the theoretical frequency. This time though, the coefficients' distribution has a slightly negative mean, whether we look at the significant coefficients only or at all of them. For all higher lags, the observed frequencies are between the theoretical value and 8.5%. The departure from the theoretical frequency for these higher lags is possibly due to the use of net of fees returns to compute the autocorrelation coefficients, a fact that introduces spurious autocorrelation. Ideally, pre-fees returns should have been used when conducting autocorrelation analysis, but such figures are difficult to infer because of the often complex fee structure.

We think it is safe to conclude that a vast majority of funds do not show significant autocorrelation for lags higher than 2 and that for 95% of the funds, as well as for CTAs as an asset class, it is safe to use the square root of time rule to scale standard deviations.

3.2.1 Trend-followers vs. non-trend-followers

Using our 2 subsamples, we run autocorrelation analysis and look for differences between trend-followers and non-trend-followers. The results can be seen in Figures 35 and 36 in the Appendix. Even though higher frequencies of significant coefficients are found for non-trend-followers in each time lag, a clear structure of autocorrelation for this category cannot be inferred from the average coefficients, since most of them are close to zero. For trend-followers, all frequencies are very close to the theoretical 5% of significant coefficients except for the 2 month lag. For this lag, one can see that trend-followers are characterized by negative autocorrelation.

4 Persistence in return distributions: the informational content of track records

A good track record is undeniably the key selling point for a CTA manager. Past accomplishments, if they are outstanding, are sure to attract fresh money to the fund. Virtually all analysts, and more generally investors, rely heavily on track records to build an opinion of what a fund's future performance and risk are likely to be relative to its peers'. Agarwal, Daniel, and Naik (2003) find clear evidence that investors infer a hedge fund manager's ability through past returns. Specifically, they study the relationship between money flows and past relative performance. They find that the top performing quantile's assets under management experience, on average, an inflow of 63% versus a 3% average outflow for the worst performing quantile. This section will attempt to answer the following question: To what extent is it reasonable to rely on parameters of past return distributions to infer the funds' future performance and risk profile? Throughout this section, we will follow closely Kat and Menexe (2003) which focuses on hedge funds, and Brown and Goetzmann (1995), which studies mutual funds, applying similar methodologies to CTAs.

4.1 Data and methodology

4.1.1 Long term persistence

First, in a two period setting and at the individual fund level, we look for persistence of the following four moments of CTAs' distributions: average return, standard deviation, skewness, kurtosis. We also investigate persistence for two risk-adjusted performance measures: a simplified Sharpe ratio computed as $\frac{\mu}{\sigma}$ (the ratio of average return to standard deviation) and, since we recognize that CTAs' returns are not normally distributed, the Omega measure, as described in Keating and Shadwick (2002a) and Keating and Shadwick (2002b). For Omega, we specify a threshold of zero. We conduct the analysis on an 8-year period, 06/01/1994-07/31/2002, that we split into two 4-year sub-periods: 06/01/1994-05/31/1998 (period 1) and 06/01/1998-05/31/2002 (period 2). We filter our dataset to keep only funds that traded over the full 8-year period. We find 113 such funds. We use both parametric (linear regression) and non-parametric (contingency tables) methods to test for the presence of long term persistence in the aforementioned measures. That is, for each measure, we try to find evidence that a relatively high (low) parameter value in period 1 is followed by a relatively high (low) parameter value in period 2. As in the previous section, we run the same analysis on the 2 subsamples of trend-followers and non-trend-followers, for which we find respectively 61 and 52 funds that traded over the full period.

4.1.2 Short term persistence

We test for the presence of short term persistence in the distribution of individual funds' returns using contingency tables. We proceed in a sequential manner, building contingency tables for 2 consecutive 1-year periods. The first contingency table reflects performance persistence for years 1994-1995, the second reflects 1995-1996, and so on and so forth, up to the period 2002-2003. For each table, we select funds that have been trading throughout the whole 2 year period. In order to compute accurately higher order moments, such as skewness and kurtosis, as well as the Omega measure, many more data points would be needed than those from 2 years worth of monthly returns, therefore the short term analysis focuses on average return, standard deviation and their ratio (simplified Sharpe ratio).

4.1.3 Regression

This analysis is done by regressing the excess parameter over median in period 2 on excess parameter over median in period 1. Since we are regressing excess parameters, we do not allow for an α and report the slope of the regression line as being β.

4.1.4 Contingency tables

There are many ways to look for persistence relative to a peer group. Contingency tables are perhaps the most widely used method to do so, as judged by the vast literature (for contingency tables applied to hedge funds see for example Kat and Menexe (2003) or Agarwal and Naik (2000a), for contingency tables applied to mutual funds see Brown and Goetzmann (1995) or Goetzmann and Ibbotson (1994)). The contingency table approach is used to identify the frequency with which funds are defined as winners and losers over successive time periods. In order to build a contingency table, we first compute the parameters under review for each fund in period 1. We then compare each parameter to the median value of the corresponding parameter in the first period, assigning a W for winner (parameter is above peers' median value), and a L for loser (parameter is below peers' median value). We repeat the process for period 2. Consistent out-performers are labeled WW, consistent under-performers LL, while funds showing no persistence are labeled LW or WL. From there, we compute the frequencies with which funds are defined as winners and losers over successive time periods. Several tests can be used to check for significance in persistence/reversal of the parameters. We retain three such statistical criteria, which are used to test for different forms of persistence:

Malkiel's repeat winner test The first statistical test is the repeat winner approach of Malkiel (1995). This test concentrates on the persistence of only one quadrant of the contingency table (WW) by looking at the proportion of repeat winners (WW) to winner-losers (WL). If p is the probability that a winner in one period continues to be a winner in the subsequent period, a value of p less than or equal to $\frac{1}{2}$ indicates no persistence. Thus, a binomial test of $p > \frac{1}{2}$ can be used to test the significance of the proportion of WW to (WW+WL) as follows:

$$Z = \frac{y - np}{\sqrt{np * (1 - p))}} \tag{3}$$

Where: y is the number of repeat winners (WW), n is the sum of repeat winners and winner/losers (WW+WL) and p is the theoretical value when there is no persistence, that is $\frac{1}{2}$. This statistic is approximately normally distributed with zero mean and standard deviation equal to one, when n is reasonably large. Thus, a percentage of WW to (WW+WL) above 50% and a Z-statistic above zero indicates performance persistence for winners, while a percentage value below 50% and a Z-statistic below zero indicates a reversal in performance.

Fienberg's cross product ratio (CPR) This statistical technique, found in Fienberg (1980) and used by Goetzmann and Ibbotson (1994) and Kat and Menexe (2003), is more general, since it tests the persistence of both repeat winners (WW) and repeat losers (LL). The CPR test statistic is the ratio of the product of repeat winners (WW) and repeat losers

(LL) divided by the product of winner-losers (WL) and losers-winners (LW):

$$CPR = \frac{WW * LL}{LW * WL} \tag{4}$$

A CPR of 1 would support the hypothesis that the parameter under review in one period is unrelated to that in the other. A CPR greater than 1 indicates persistence, while a value below 1 indicates that a reversal in the parameter dominates the sample. The statistical significance of the CPR can be determined by using the standard error of the natural logarithm of the CPR given by the square root of the sum of the cell counts' reciprocals:

$$\sigma_{\log CPR} = \sqrt{\frac{1}{WW} + \frac{1}{LL} + \frac{1}{WL} + \frac{1}{LW}} \tag{5}$$

For sufficiently large samples the test statistic is normally distributed with mean $\log(CPR)$. The Z-statistic for this test is given by:

$$Z = \frac{\log(CPR)}{\sigma_{\log(CPR)}} \tag{6}$$

χ^2 The χ^2 test considers persistence of a contingency table as a whole and is used for example in Agarwal and Naik (2000a). The value for the χ^2 test can be computed as:

$$\chi^2 = \frac{(WW - N/4)^2 + (WL - N/4)^2 + (LW - N/4)^2 + (LL - N/4)^2}{N} \tag{7}$$

Where N is the total number of funds in the sample. The χ^2 value must be compared to a theoretical value taken from a χ^2 distribution with $(2-1)*(2-1) = 1$ level of freedom in order to test for significance. Since the χ^2 value is always positive, one must rely on the actual table frequencies to know if we are in the presence of persistence or reversal. Carpenter and Lynch (1999) study the specification and power of various persistence tests. They find the χ^2 test to be well specified, powerful and more robust in the presence of survivorship bias when compared to other test methodologies. Thus, we will give it more importance in the interpretation of the results.

4.2 Long term persistence results

We present the results of the long term persistence analysis in graphical form on Figure 6. The graphs are built so as to show the excess parameter over the corresponding median in

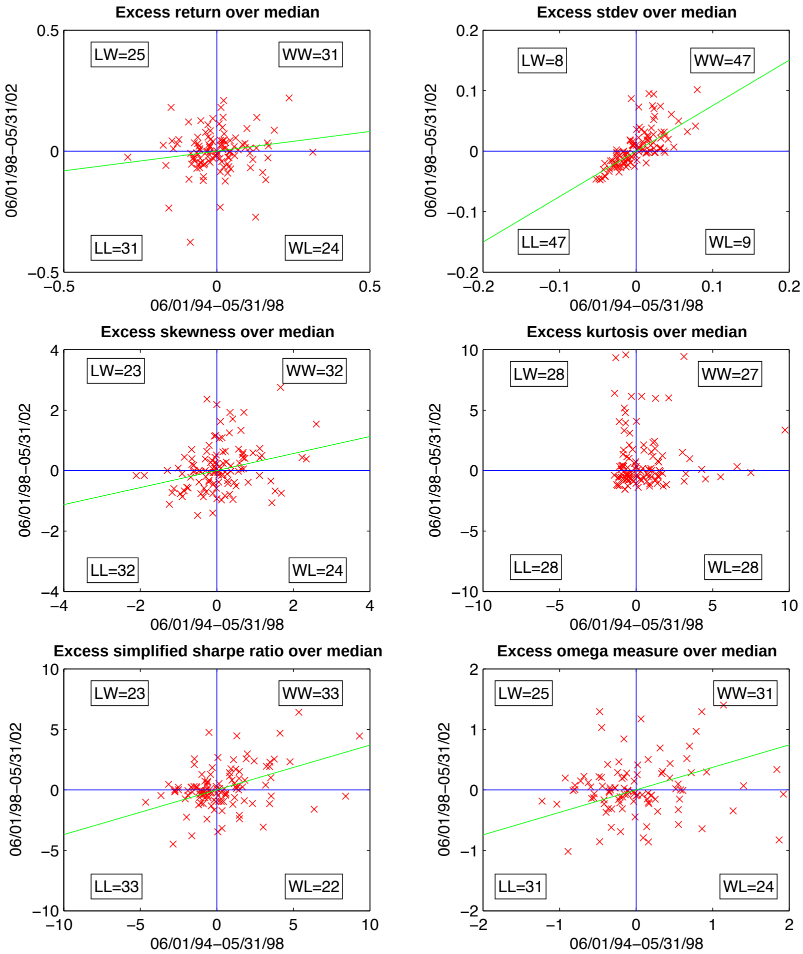

Figure 6: Persistence in CTAs' return distributions parameters

period 1 on the x axis, while the excess parameter over the median in period 2 is plotted on the y axis. We also present the number of funds found in each quadrant of the contingency table, as well as the regression line when it is significant (at a 5% confidence level). The statistics associated with the different tests of persistency can be found in Tables 1 through 4.

From these results, one can clearly see that there is strong evidence of relative standard deviations persistence. Indeed, each of the 3 statistical tests, as well as the high R^2 of the regression back this fact with p-values very close to 0. Interestingly enough, the regression line fits the LL quadrant points better than the WW, indicating that funds with relatively

Table 1: Regression statistics for CTAs' distributions parameters

	$\boldsymbol{\beta}$	**Significant (5%)**	$\boldsymbol{R^2}$
Avg. return	0.16	Yes	0.08
Stdev.	0.75	Yes	0.54
Skewness	0.28	Yes	0.08
Kurtosis	0.06	No	0.10
$\boldsymbol{\frac{\mu}{\sigma}}$	0.37	Yes	0.23
$\boldsymbol{\Omega}$	0.37	Yes	0.59

Table 2: Malkiel winner persistence test statistics for CTAs' distributions parameters

	% Repeat winners	**Z-stat**	***p*-value**
Avg. return	0.56	0.94	0.35
Stdev.	0.84	5.08	0.00
Skewness	0.57	1.07	0.28
Kurtosis	0.49	-0.13	0.89
$\boldsymbol{\frac{\mu}{\sigma}}$	0.60	1.48	0.14
$\boldsymbol{\Omega}$	0.56	0.94	0.35

Table 3: CPR statistics for persistence of CTAs' distributions parameters

	CPR	**Z-stat**	***p*-value**
Avg. return	1.60	1.23	0.22
Stdev.	30.689	6.4	0.00
Skewness	1.86	1.61	0.11
Kurtosis	0.96	-0.10	1.00
$\boldsymbol{\frac{\mu}{\sigma}}$	2.1522	1.98	0.05
$\boldsymbol{\Omega}$	1.6	1.23	0.22

Table 4: χ^2 statistics for persistence of CTAs' distributions parameters

	χ^2	p-value
Avg. return	0.38	0.53
Stdev.	13.13	0.00
Skewness	0.65	0.42
Kurtosis	0.02	0.90
$\frac{\mu}{\sigma}$	0.99	0.32
Ω	0.39	0.53

low standard deviation exhibit more stability for this parameter.

An intermediate case is that of the simplified Sharpe ratio's persistence. Whereas the Malkiel and χ^2 tests' p-values tend to deny persistence in the parameter, regression is significant, with a decent R^2, and the CPR indicates persistence at a 5% confidence level.

None of the 3 tests find persistence for average returns, skewness or the Omega measure, despite significant coefficients for the regressions, as well as numbers of LL and WW being higher than WL and LW across the board. Lack of sufficient data, which renders the estimation of higher order moments difficult, as well as the relatively small number of funds in our sample, may be responsible for this inconclusive evidence.

The numbers in each quadrant for kurtosis indicate a very small reversal for this parameter, one that all tests consider not to be significant.

Let us mention that our results are consistent with those found in Kat and Menexe (2003) and Schwager (1996).

4.2.1 Trend-followers vs. non-trend-followers

The results for the 2 subsamples, which can be found in Figure 37 and Table 14 through 17 for the non-trend-followers, and in Figure 38 and Table 18 through 21 for trend-following funds, confirm the preceding section's findings. Standard deviation shows strong persistence for both subsamples. While nothing can be said about skewness, kurtosis, Omega or average returns, the Sharpe ratio seems to show a decent amount of persistence in the case of trend-followers. The same analysis does not lead to significant evidence of persistence in the case of non-trend-followers.

4.3 Short term persistence results

Here, we present the results of our short term persistence analysis for average return, standard deviation as well as simplified Sharpe ratio. As can be seen in Figure 7, out of the nine 2-year periods, 6 exhibit performance persistence, 2 exhibit performance reversal (1994-1995 being a rather strong reversal), and 1 exhibits neither persistence nor reversal. Even though the corresponding χ^2 statistics never indicate this persistence as being significant at generally accepted confidence levels, the evidence tends towards frequent years of slight to moderate persistence and few years of strong reversals.

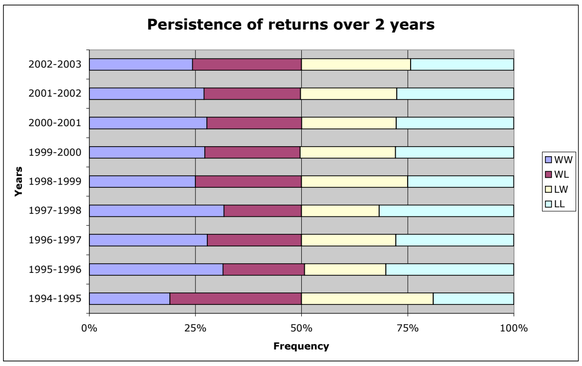

Figure 7: Returns contingency table frequencies for 2-year periods between 1994 and 2003

Figure 8 gives more evidence that it is safe for analysts to forecast a fund's standard deviation relative to its peer group by using its track record. Indeed, for every 2-year period, the relative persistence is very strong and statistically significant.

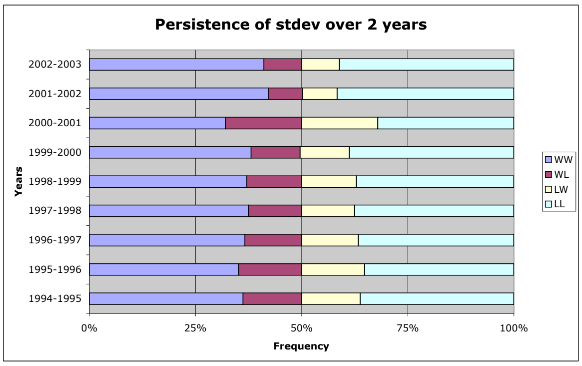

Figure 8: Std. contingency table frequencies for 2-year periods between 1994 and 2003

Figure 9 shows that, even though persistence and reversals are never statistically significant, the Sharpe ratio exhibits persistence for seven 2-year periods out of 9,versus only two periods when there are relatively small reversals.

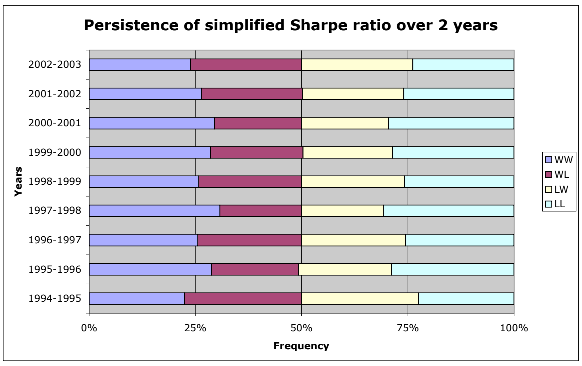

Figure 9: Sharpe ratio contingency table frequencies for 2-year periods between 1994 and 2003

4.4 Summary of results

Our results show that the widespread use of CTA funds' track records by investors is a wise practice. The use of track records makes most sense when it comes to estimating the future risk of a fund, as measured by its standard deviation level relative to its peers. Despite the lack of statistical significance attached to persistence in certain other parameters (average return, simplified Sharpe ratio), it is worth noting that, during the last decade, these parameters have been characterized by year to year persistence more often than by reversals. All in all, even though track records are not a miracle recipe for fund picking, they should definitely be used when evaluating the risk of a portfolio. Their consideration certainly cannot hurt from the point of view of both absolute and risk adjusted performance. These preliminary results have made us confident that an efficient methodology for analyzing CTA funds' returns from a quantitative point of view should focus on understanding the various sources of each fund's risk rather than seeking potential out-performers in an asset class where returns are highly volatile and relative performance unstable.

5 CTA's risk premia

As illustrated by the previous section, CTAs' returns tend to be highly unstable. In order for relatively risky CTAs to attract funds, they have to promise relatively high long-term returns. The analysis we run in this section aims at answering the following questions: did relatively more risky CTAs fulfill their promise of superior long-term returns over the period 1994-2002? And, if so, what is the approximate additional unit of return for each additional unit of risk?

5.1 Data and methodology

The dataset we use for this analysis is the same that was used in the study of long term persistence, that is, a sample of 133 CTAs that have been trading over a full 8-year period from 1994 to 2002, out of which 61 are trend-followers and 52 are non-trend-followers. In order to check for the presence of a relationship between standard deviation and average return, we regress the 8-year period average return of the funds on their standard deviations. We also split the samples in two, creating a sample of relatively more risky funds, for which standard deviation was found to be higher than the median, and a sample of relatively less risky funds, for which standard deviation was found to be lower than the median. We then run a two sample t-test in order to determine if the difference in the mean returns is significantly different from 0.

5.2 Results

As can be seen in the first panel of Figure 10, more risky CTAs clearly earned a higher average rate of return during the period under review. The regression is statistically significant at a 1% level, even though the R^2 is somewhat low at 11.5%. The difference of average return between the low standard deviation and the high standard deviation groups is of approximately 4%, the t-test is highly significant. Our data indicates that a CTA investor can expect to receive approximately 0.27% additional yearly return for each 1% additional risk taken.

The results for trend-following funds, which can be found in panel 2 of Figure 10, are consistent with the results for all CTAs as a group. The regression model is highly significant. The *t*-test significance is hurt by the smaller size of the sample, but is still acceptable, with a *p*-value just above 5%. The difference between the high and low standard deviation groups is of approximately 5.7%, and the risk premium for trend-followers seems to be a little higher than for CTAs as a group, with a 0.34% increase in average return for each unit of risk taken.

Most interestingly, when it comes to non-trend-followers, it seems that only a very weak risk-return relationship can be found. In panel 3 of Figure 10, one can see the very slight incline of the regression line. The model is not significant, even at a confidence level of 10%. Despite the fact that the high standard deviation group's yearly average return is slightly higher that the low standard deviation group's (1.3% difference), the *t*-test clearly rejects the hypothesis of a significant difference between the 2 groups' average returns. This has strong implications for choosing non-trend followers in which to invest: since relative standard deviation has been showed to be very stable for both short term and long term time periods, and since it seems that higher risk is not rewarded by higher return, investors should invest solely in low risk funds in that category of CTA's.

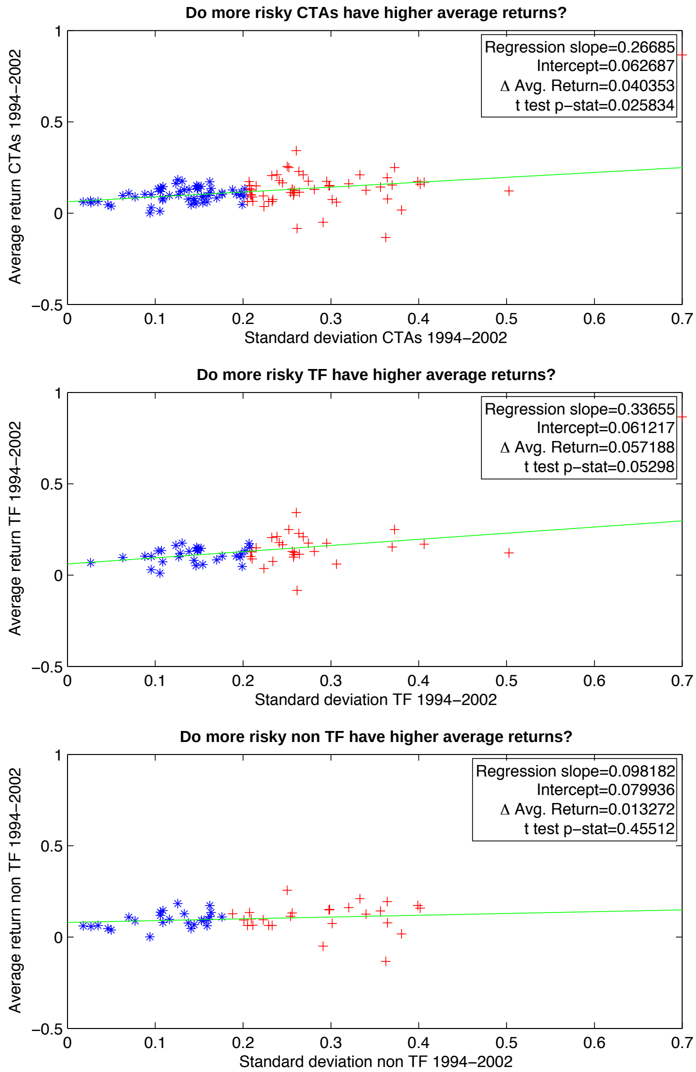

Figure 10: Comparison of average returns for funds with relatively low/high std.

6 CTAs Benchmarks

Benchmark indices are regularly used in studies of investment performance to provide a relative performance index. Even in the alternative investment industry, where it is often argued that performance ought to be compared to an absolute return target, investors still find it reasonable to look at how their picks have performed versus a benchmark. Benchmarks exist in various forms, the most common being a portfolio of similar investment vehicles whose net-of-fees returns can be tracked and used as a proxy for a particular asset class, investment style, or strategy.

Benchmark indices may be either investable or non-investable. A benchmark is termed investable if it is easily replicated through a portfolio of underlying securities. For the purpose of asset allocation between broad asset classes, as well as for tracking relative performance of a portfolio of funds, investable benchmarks should be used. Since the goal of such benchmarks is to represent the actual performance of the asset class under review, biases, such as instant history or survivorship bias, are a major concern when choosing the best index. Let us also mention that, even if a CTA index is deemed investable, it will rarely be possible to fully replicate it, a fact due to the specific conditions (i.e. fees, leverage, etc) at which investors commit their money through individual managed accounts.

A benchmark classification done according to weighting schemes yields two groups: equally weighted and dollar-weighted benchmarks. Equally weighted benchmarks are usually non-investable, since replicating them would require a great amount of rebalancing. Indeed, keeping each fund's weight in the portfolio equal requires selling winners and buying losers frequently. Nevertheless, equally weighted indices are widely used. Technically, they should be used as a measure of relative performance for individual fund managers, since they reflect the investment performance of the average manager in the industry. Dollar-weighted indices are usually investable, but not always, since replicating them may require making investments in funds that are currently closed to new entrants.

In this section, we describe different CTA benchmarks currently available to the industry.

We construct two representative benchmarks (equally weighted and dollar-weighted) from the universe of funds in our dataset and compare them to the industry indices. We then proceed with several analysis relying on benchmarks. We infer the magnitude of the biases present in our dataset, study CTAs' homogeneity and determine diversification possibilities in a portfolio of CTAs.

6.1 CTA industry benchmarks

The major CTA indices are usually constructed by CTA funds database providers. They differ widely in purpose, composition, weighting and re-balancing scheme. Most of CTA indices used to be non-investable and equally weighted. Most of the new indices are designed to be as investable as possible. Some of them contain a broad universe of CTA funds while other focus only on the biggest ones. Some providers offer indices for specific categories within CTAs (i.e. trend-followers, discretionary traders, currency managed futures programs, stock index managed futures, etc)

Barclay indices

The provider of the database for our study, Barclay Trading Group, publishes a CTA index, Barclay CTA, which is currently composed of 386 programs. It is equally weighted and re-balanced at the beginning of each year. To qualify for inclusion in the CTA index, an advisor must have four years of performance history. Additional programs introduced by qualified advisors are not added to Barclay CTA until after their second year. These restrictions should help create a CTA index free of instant history bias, since Park (1995) and Brown, Goetzmann, and Park (1997) found an incubation period for CTA varying between 15 and 27 months. Unfortunately, the restrictions also introduce a small amount of survivorship bias, since no fund that reported for less than 4 years ever was included in the index. Barclay also publishes an index that we believe is fit for benchmarking trend-followers: The Barclay Systemic Traders Index (Barclay TF). Barclay TF is an equally weighted composite of managed programs whose approach is at least 95% systematic. In 2004, there are 333 systematic programs included in the index. A majority of managers using systematic approaches are trend-followers. For discretionary managers (non-trend-followers, for most of them), Barclay issues the Barclay Discretionary Traders Index (Barclay NTF), which is an equally weighted composite of managed programs whose approach is at least 65% discretionary or judgmental. In 2004 there are 85 discretionary programs included in the index. Barclay also offers the following subindices: agricultural, currency, diversified, energy, financial/metal. As a novelty versus other benchmark providers, Barclay publishes the Calyon Financial Barclay Index, which provides the market daily performance of major commodity trading advisors. The index calculates the daily rate of return for a pool of CTAs, equally weighted, and selected from the larger managers that are open to new investments. Selection of the pool of qualified CTAs used in the construction of the index is conducted annually, with re-balancing on January 1st of each year. Currently, there are 22 funds in the index. Finally, Barclay also has an investable index, the Barclay BTOP 50 Index. The index seeks to replicate the overall composition of the managed futures industry with respect to trading style and overall market exposure.

CISDM indices

CISDM (Center for International Securities and Derivatives Markets) also tracks performance of CTAs using equally weighted and dollar-weighted benchmarks. The CISDM indices, originally constructed by MAR (Managed Account Reports) track the performance of individual CTAs as well as CTA funds and pools that invest in individual CTAs. CISDM/MAR asks for a minimum track record of 12 months of trading actual client accounts (as opposed to proprietary accounts). A minimum of $500'000 of assets under management is also required. CISDM main CTA index is the Trading Advisor Qualified Universe Index. Two versions of this index exists: one equally weighted, the other dollar-weighted. In our analysis, we will use the dollar-weighted index (CISDM CTA). We will also use their trend-follower subindex (CISDM TF) as well as the discretionary advisors subindex (CISDM NTF). On top of the aforementioned benchmarks, CISDM publishes subindices for currency, European, stock index, financial and diversified traders. As mentioned before, CISDM offers the CTA Fund/Pool Qualified Universe Index (equally weighted and dollar-weighted formulas) together with the following subindices: guaranteed funds, private pools, public funds, off-shore funds, multi-advisor and single-advisor funds.

CSFB-Tremont indices

CSFB-Tremont, which uses the TASS database, offers historical managed futures performance similar to MAR or Barclays. It publishes a dollar-weighted CTA index and one subindex of currency CTAs. Additional indices and subindices will be available in the future.

S&P index

The S&P Managed Futures Index is constructed to offer investors a standardized and investable benchmark that is representative of systematic managed futures programs. Standard & Poor's used sampling bootstrap simulation techniques to conclude that a portfolio of 12 to 15 programs represents the risk/return characteristics of the broader managed futures universe. To improve the extent to which the index is representative of CTAs, programs with several volatility levels are included. Managed futures programs in the index are investable and have passed a due diligence evaluation. The S&P Managed Futures Index is thus the most easily replicable of the indices we present.

MSCI index

MSCI publishes the MSCI Directional Trading Index, which is designed to capture the performance of 134 funds that employ discretionary trading, systemic trading or tactical allocation strategies. The index is a component of the MSCI Hedge Fund Composite Index which serves as an investable benchmark for the whole hedge funds universe. The index weights are determined using a "median asset-weighted scheme", which is a combination of equally weighted and dollar-weighted schemes. MSCI believes this procedure of computing the weights insures both the funds' representativeness and minimum biases. MSCI also imposes maximum limits for the weights. MSCI Directional Trading Index is reviewed on quarterly basis. Weights are re-calculated, new funds are added, and dead/non-investable funds are deleted.

Tuna index

Hedgefund.net is constructing a series of indices (Tuna indices) for hedge funds and CTAs. The Tuna CTA/Managed Futures is an equally weighted benchmark currently composed of 300 managed futures programs. Hedgefund.net acknowledges that the Tuna indices are affected by survivorship bias.

EDHEC indices

EDHEC publishes the CTA Global index which is a combinaison of 5 major CTA indices: Barclay CTA, CISDM CTA, CSFB-Tremont CTA, Hedgefund.net's Tuna CTA index and S&P Managed Futures Index. Indeed, a combination of indices is prone to be more exhaustive than any of the competing indices it is extracted from. The key feature of EDHEC Alternative Indices is to be found in its weighting scheme. Based on Amenc and Martellini (2002), EDHEC derives weights alloted to each CTA index by using the first component of a principal component analysis. Using principal component analysis insures that no other linear combination of competing indices leads to a lower information loss, while minimizing the extent to which each index's bias affects the EDHEC CTA Global index. Unfortunately, EDHEC has to wait for all other indices to be published before releasing figures for its own index, which renders it less useful for practitioners.

ITR index

The last index we present here is the ITR Premier 40 CTA Index. Each month ITR ranks their proprietary database of over 400 programs to find the top 40 composite CTA programs based on end-of-month assets under management. Based on these 40 composite programs, a dollar-weighted index is constructed. There are no limits as to the number of composite programs that an individual advisor may have in the index in any given month. This index represents a good proxy for the overall performance of leading CTAs.

Choosing the right benchmark for an investor depends mainly on the composition of his/her portfolio. Should it contain a large number of small CTAs, an equally weighted benchmark would be preferred. Should it contain a small number of large CTAs, a dollar-weighted benchmark would be more appropriate. For a more quantitative comparison of available CTA benchmarks, refer to Schneeweiss and Spurgin (1997).

6.2 Database benchmarks

In order to analyze how our data set relates to generally accepted CTA indices, we aggregate our whole universe of funds into representative benchmarks.

6.2.1 Construction

We build both an equally weighted and a dollar-weighted benchmark for all CTAs (CTA-EW and CTA-DW), for trend-followers only (TF-EW and TF-DW) and for non-trend-followers only (NTF-EW and NTF-DW). The benchmarks start on January 1994 and run for more than 10 years until March 2004. The equally weighted benchmarks are constructed as an arithmetic average of the existing funds returns at a certain date. Since returns are monthly, re-balancing takes place every month. For the dollar-weighted index, we use the assets under management provided in the database as a proxy for the fund's market value. We acknowledge, in the spirit of Fung and Hsieh (1997b), that our benchmarks are affected by survivorship and instant history biases.

6.2.2 First 4 moments

We compute the average return, standard deviation, skewness and kurtosis for both the equally weighted and the dollar-weighted benchmarks. The results are shown in Table 5. Like in the individual CTAs return distribution analysis, the average return and the standard deviation are expressed in yearly terms.

Firstly, we notice a bigger average return and skewness for the equally weighted benchmarks than for the dollar-weighted indices. This fact supports the intuition that smaller funds outperform bigger ones, since they are not affected by overcapacity (overcapacity refers to returns diminishing as assets under management's size increases). There is another aspect we can infer from this result, and that could appear counter-intuitive at first: equally weighted benchmarks yield better returns than passive, dollar-weighted ones even though they sell "winners" and buy "losers" in order to keep the weights equal. This piece of information is consistent with the presence of 1 month lag negative autocorrelation for individual funds. Thus, if liquidity allows for monthly re-balancing, a CTA contrarian strategy could potentially be profitable.

Secondly, the benchmark's average returns are quite similar to the individual CTA's average return (15% per year) shown previously in Figure 4. We cannot make the same statement for standard deviation since the benchmark's standard deviation is 10%, while the individual CTA's average standard deviation is approximately 20%. Since our benchmarks are constructed as portfolios of individual funds that are not perfectly correlated, this is diversification at work. We can thus infer that there is quite a large amount of diversification achievable when combining a a number of CTAs in a portfolio. Such diversification effect appears to be present in kurtosis too, but not in skewness. When comparing trend-followers versus non-trend-followers, we reach the same conclusions as we did earlier: trend-followers are more risky than non-trend-followers (12% vs. 7% in standard deviation), a fact that is compensated this time by bigger average return and skewness. In conclusion, we outline that the "market portfolio" of CTAs has a similar return and skewness but a smaller risk (standard deviation) and kurtosis than the average individual CTA.

Table 5: First 4 moments of database benchmarks' distributions

	CTA-EW	**CTA-DW**	**TF-EW**	**TF-DW**	**NTF-EW**	**NTF-DW**
Mean (%)	17.10	12.30	18.46	13.84	15.42	10.86
Stdev. (%)	9.560	10.73	11.84	13.31	6.925	8.512
Skewness	0.5424	0.3495	0.5919	0.4216	0.4078	0.2349
Kurtosis	3.408	3.232	3.430	3.355	3.188	2.953

6.2.3 Benchmarks' autocorrelation

We study benchmarks' correlation for the same reasons that we studied individual CTA's returns autocorrelation. The results of the autocorrelation analysis are presented in Table 6. The autocorrelation coefficients for the six benchmarks are slightly negative but very close to zero. The significance test at a confidence level of 5% shows that none of these autocorrelation coefficients are significant, even though for the 2-month lag the p-values are quite low. This supports our earlier intuition that CTAs, as an asset class, do not exhibit significant autocorrelation. For some lags (2, 4 and 7 months) we notice that the coefficients are slightly more negative than in the other months, which may be due to elements like management fees or other commissions that can lead, as previously stated, to spurious autocorrelation.

Table 6: CTA benchmarks' autocorrelation coefficients for lags between 1 and 8 months

Lag	**CTA-EW**	***p*-value**	**CTA-DW**	***p*-value**
1	0.018	0.848	0.056	0.542
2	-0.165	0.071	-0.172	0.059
3	-0.057	0.539	-0.070	0.448
4	0.104	0.260	-0.098	0.288
5	0.018	0.848	-0.039	0.677
6	-0.043	0.645	-0.049	0.596
7	-0.110	0.239	-0.119	0.203
8	-0.005	0.958	0.005	0.961

6.2.4 Trend-followers vs. non-trend-followers

The results for the autocorrelation analysis are shown in Tables 22 and 23 in the Appendix. The interesting thing to notice here is the significant negative coefficient for TF-EW 2-month lag autocorrelation. This reinforced our previous findings and the possibility of implementing a 2 months contrarian strategy for trend-followers. Concerning the non-trend-followers, as expected, we do not observe any significant auto-correlation.

6.3 Comparison: database benchmarks vs. industry benchmarks

In order to see how the data we use is representative of the CTA industry as a whole, and also to measure the biases that affect it, we proceed with a comparison between the benchmarks built from the database in the previous section and the industry indices.

6.3.1 Evolution over time and bias correction

The easiest way to assess how our benchmarks performed versus Barclay CTA and CISDM CTA over the 10 years our study focuses on is to plot them. Figure 11 includes 2 graphs. The first one presents the evolution of the database benchmarks along with the Barclay and CISDM indices. The second presents the same benchmarks, but with the database benchmarks corrected for survivorship and instant history biases. The correction is based on the results obtained by Fung and Hsieh (1997b). We subtract 3.6% per year for the survivorship bias and another 3.6% for the instant history bias. We notice from the first plot in Figure 11 that the database benchmarks evolve in the same way as the industry ones. They are affected by the same shocks, but it is obvious that they are upward biased, which is our rationale for the bias-correction. The bias-corrected benchmarks are much closer to the industry indices. The Barclay CTA and the CISDM/MAR CTA indices are bounded above by the equally weighted benchmark and bounded bellow by the dollar-weighted benchmark. The Barclay CTA plot looks like an arithmetic average between the CTA-EW and CTA-DW. This can be explained by the fact that the Barclay CTA index, based on a dataset very similar to ours, is an equally weighted average, re-balanced every year. This rebalancing scheme can be thought of as a mix between a monthly equally weighted average (CTA-EW) and a dollar-weighted one (CTA-DW).

Additionally, one can see from the plot that the average CTA manager (represented by CTA-EW) outperforms the industry (CTA-DW).

6.3.2 Trend-followers vs. non-trend-followers

We repeat the same analysis for the trend-followers and non trend-followers. We compare the TFs' benchmarks with the Barclay Systemic Traders Index and CISDM/MAR Trend-Based Subindex, while the non-trend-followers are plotted against the Barclay NTF and CISDM NTF. The results can be seen in Figures 39 and 40 in the Appendix. The conclusions are the same as in the analysis for all CTAs.

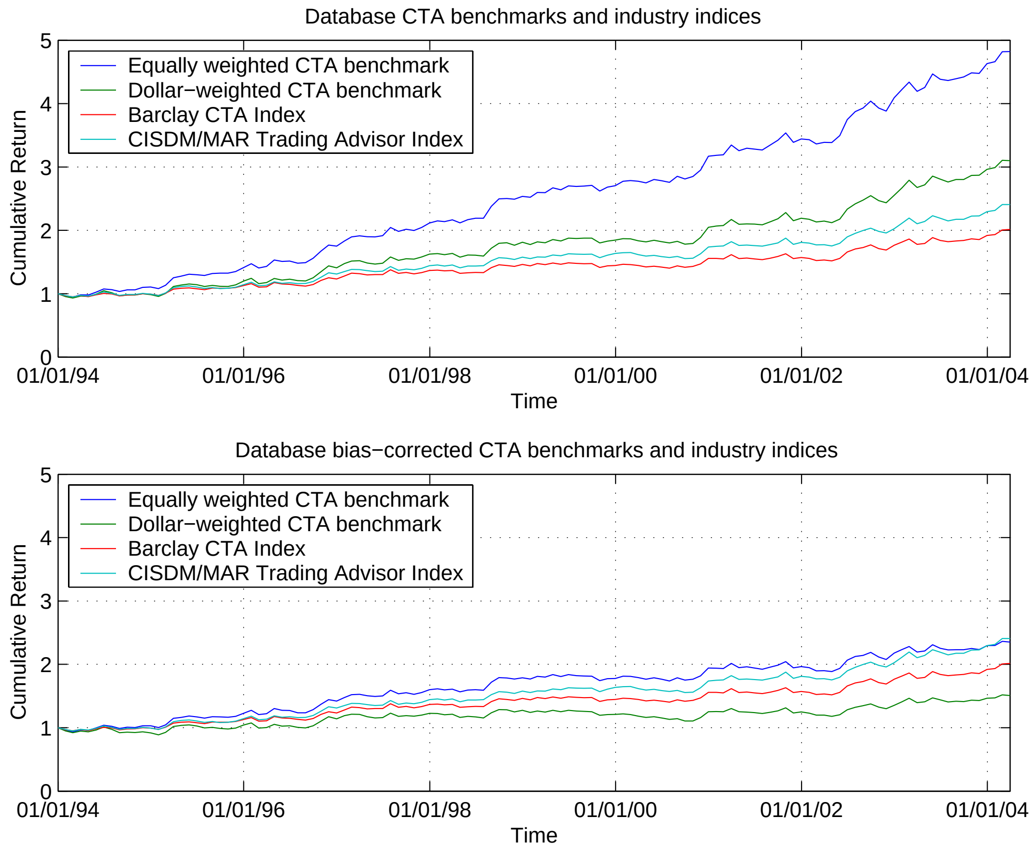

Figure 11: Biased and bias-corrected CTA benchmarks vs. industry indices

6.3.3 Benchmarks' correlations

Table 7 presents the correlation coefficients between the benchmarks built from the database and the industry indices. CTA benchmarks are all strongly correlated. We can also observe that CTA-EW is more correlated to the Barclay CTA index, while CTA-DW has a bigger correlation with the CISDM/MAR Trading Advisor Index. This could be explained by the similarities in the corresponding indices' construction. The p-values for these correlation coefficients are very close to zero. This leads us to conclude that the benchmarks constructed from the database have strong ties with commonly used indices, such as Barclay CTA and CISDM CTA. Except for the biases, our sample seems representative of what has happened in the CTA industry in the past ten years.

Table 7: CTAs benchmarks correlations

	CTA-EW	**CTA-DW**	**Barclay CTA**	**CISDM CTA**
CTA-EW	1	0.970	0.970	0.956
CTA-DW	0.970	1	0.954	0.983
Barclay CTA	0.970	0.954	1	0.957
CISDM CTA	0.956	0.983	0.957	1

6.3.4 Trend-followers vs. non-trend-followers

Trend-followers benchmarks exhibit correlations of a magnitude similar to CTAs (results are disclosed in Table 24 in the Appendix). The same is not true for non-trend-followers: correlations between the database benchmark and the industry indices are about 0.5, while the correlation between Barclay NTF and CISDM NTF is about 0.4. Clearly, the universe of funds used in the construction of each of the 4 benchmarks is very different. These low correlation numbers are also a hint that non-trend-following is a much more heterogeneous investment style than trend-following.

6.4 Homogeneity of the CTA asset class

In order to grasp the extent to which funds are similar to the benchmarks, we proceed with a correlation analysis of 108 individual CTA funds that have been trading for the past 10 years. For each fund, we compute the correlation coefficient with both Barclay CTA and CISDM CTA. We present the histogram of the correlation coefficients in Figure 12. The correlations range from -0.2 to 0.9. Approximately 50% of CTAs have a correlation of 0.5 or more with the Barclay CTA index. The same is true for the CISDM CTA index. Overall, the coefficients are meaningful (94% are significant). Even though for many CTAs the benchmarks turn out to be a relevant proxy, it would be hard to infer each fund's characteristics by looking only at the benchmark. Considering the large number of strategies used by CTAs, it is not very surprising to find evidence of such a high level of heterogeneity. This is why we turn now to our two subsamples, with hope that, by making one further classification, a higher level of homogeneity can be achieved.

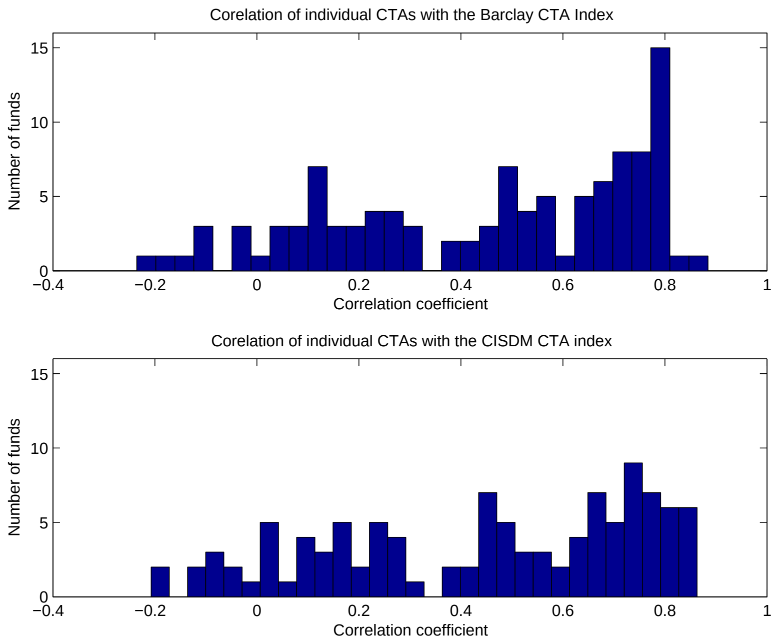

Figure 12: Histogram of individual CTAs correlation with industry benchmarks

6.4.1 Trend-followers vs. non-trend-followers

We run the same analysis as above on 58 trend-followers and 50 non-trend-followers. Trend-followers (Figure 41 in Appendix) seem to form a more homogeneous class than CTAs, with almost 65% having correlations of 0.5 or higher with the benchmarks. Given the variety of markets in which trend-followers operate, as well as the different time horizons used as a basis for trend detection, it is normal for this class not to be fully homogeneous. On top of that, some funds that we classified as trend-followers mentioned in their description the fact that they use both trend-following and contrarian strategies. This would explain the presence of funds negatively correlated with the benchmark. Non-trend-followers (Figure 42 in Appendix) clearly appear as being heterogeneous, a fact that is not surprising at all, since such funds base their investment decisions on indicators ranging from weather forecasts to genetic algorithms.

6.4.2 Diversification possibilities within the CTA class

In this section, we try to determine the magnitude of diversification possibilities in the context of a portfolio of CTAs. Since trend-followers apply homogeneous strategies, we expect them to loose/make money all at the same time. Given the variety of strategies applied by non-trend-followers, we expect less synchronism in non-trend-folllowers' losses/gains, a desirable feature for investors.

Since we have shown that CTAs' returns are characterized by non-normality, we follow an approach relying on drawdowns to quantify diversification, rather than on a Markowitz mean-variance approach. Moreover, we restrict the investment universe to CTAs. For a more general analysis of CTAs in a diversified portfolio context, see Cerrahoglu (2004).

Figures 13,14, and 15 report the frequency of funds currently experiencing a drawdown over the period 1994-2004. While the average frequency of funds in a drawdown is the

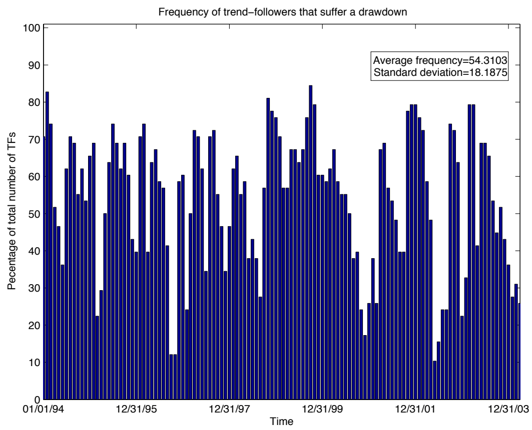

Figure 13: Trend-followers drawdown analysis

same for trend-followers and non-trend-followers, the standard deviation of this frequency is lower (14% versus 18%) for non-trend-followers than for trend-followers, which comforts us in the idea that there is significant diversification gains to be made from investments in non-trend-followers. Figure 15 depicts the situation for the whole CTA dataset. This corresponds roughly to a fund where 50% of the assets are invested in trend-followers, with the remainder

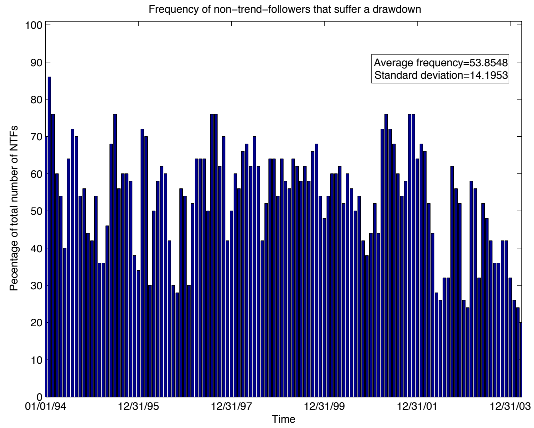

Figure 14: Non-trend-followers drawdown analysis

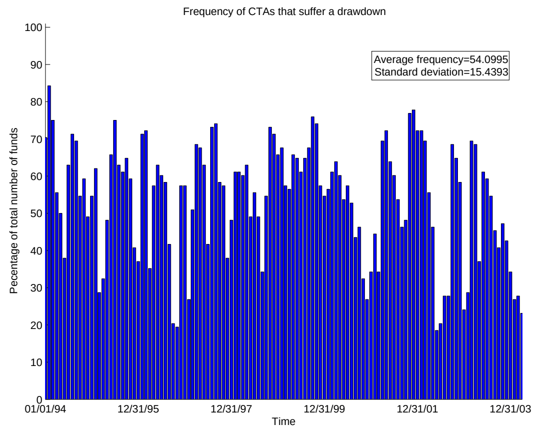

Figure 15: CTAs drawdown analysis

invested in non-trend-followers. Clearly, the diversification effect appears here at the portfolio level. The standard deviation of the drawdown frequency for a diversified portfolio is at the low end of the trend-followers and non-trend-followers frequency range(14%-18%), at around 15%. Thus, it is expected that a diversified portfolio which invests in both categories of trend-followers and non-trend-followers will experience swings of a smaller magnitude than a purely trend-follower portfolio.

7 Study of CTAs' returns in relation with traditional asset classes

In this section we study the relationship between CTAs and traditional asset classes such as equity, bonds, commodities, interest rates, and currencies. We begin with a correlation analysis. We compute correlations between benchmarks for CTAs and indices for different asset classes. We then build an asset-based factor model and proceed to regress CTAs returns, as represented by the benchmarks, on traditional asset classes. The results will help determine if a portfolio of CTAs can be replicated reasonably well using a "buy and hold" portfolio of traditional assets. The same regression will then be repeatedly applied to each CTA fund, in order to see what proportion of funds could be replicated reasonably well. We conclude the section by running a quantile analysis in order to point out non-linear structures in CTAs' returns and correlations with traditional asset classes. Specifically, we look at how CTAs behave in bear, calm, or bull market. The following analysis is conducted on the period from January 1994 to March 2004.

7.1 Introduction

Attempts at using traditional asset classes in a linear factor framework to explain CTAs' returns have been made in previous studies such as Fung and Hsieh (1997a), Agarwal and Naik (2000b), Liang (2003) and Schneeweiss and Spurgin (1998). We consider the following indices representative of traditional asset classes and build our factor model around them:

1. Equity classes
   - NYSE Composite Index (NYSE) - for US equity
   - MSCI EAFE & Canada Index (MSCI EAFE) - for developed countries (ex. USA) equity
   - MSCI EMF Index (MSCI EMF) - for emerging markets equity
2. Bond classes

- Salomon Smith Barney World Government Bond Index (SSBGovt) - for government bonds
- Salomon Smith Barney Corporate Index (SSBCorp.) - for corporate bonds
- Merrill Lynch Global High Yield Index (MLHY)

3. Commodities
   - Goldman Sachs Commodity Index (GSCI)
   - Gold Price on NYMEX (Gold)
   - London Brent Crude Oil Index (Oil)
4. Currencies
   - US Dollar Trade Weighted Index (USDTW)
5. Cash
   - 1 month eurodollar deposit (1MD)

Further information about these indices can be found in Appendix A.

The indices used to represent CTAs, trend-followers and non-trend-followers are the ones used in the previous section.

7.1.1 Asset classes evolution over time

In order to see how the different asset classes performed during the period under review, we plot in Figure 16 the dollar-weighted CTAs benchmark together with NYSE, SSBGovt, GSCI and USDTW. We eliminate some indices from the plot for simplicity. From the graph, it can be seen that the performance of CTAs in the long run is comparable to that of the bond markets, only slightly better. In the short run however, we observe strong divergence for the two asset classes. Another interesting fact is that some of the rallies for CTAs take place when the equity market crashes (Summer to Fall 1997, Asian crisis; and Fall 2001 after 9/11 attacks). This fact, though, does not hold true for commodities, as represented by

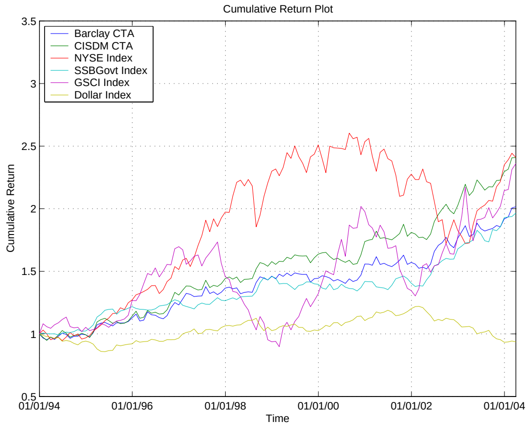

Figure 16: Performance of asset classes over time

the GSCI, emphasizing the uniqueness of CTAs as an effective diversification mean. At this stage, it is difficult to draw a clear picture of the relationship between CTAs and currencies (USDTW).

7.1.2 Trend-followers vs. non-trend-followers

The comments we made for CTAs in the previous section are valid only for trend-followers (see Figure 43 in the Appendix). Trend-followers are very sensitive to what is happening in other markets, while non-trend-followers (see Figure 44 in the Appendix) have a relatively smooth evolution (Barclay NTF), insensitive to the events pertaining to other markets.

7.2 CTAs correlations with traditional assets

The correlation coefficients of Barclay CTA and CISDM CTA versus traditional asset classes are shown in Table 8. CTAs exhibit a low negative correlation with equity, a desirable feature for an alternative investment class, but the coefficients are not significant at generally accepted confidence levels. CTAs have a significant positive correlation with both government and corporate bonds as well as with the GSCI. Commodity trading advisors are positively correlated with gold and oil, but the p-values are higher than 5%, so that they cannot be taken as significant. In all cases, correlations are relatively low.

Table 8: CTAs correlations with traditional asset classes

	Barclay CTA	***p*-value**	**CISDM CTA**	***p*-value**
NYSE	- 0.125	0.168	- 0.111	0.223
MSCI EAFE	- 0.0702	0.441	- 0.118	0.194
MSCI EM	- 0.0882	0.332	- 0.0538	0.554
SSBGovt	0.271	0.00245	0.278	0.00186
SSBCorp.	0.270	0.00256	0.290	0.00113
MLHY	- 0.128	0.158	- 0.0754	0.407
GSCI	0.220	0.0146	0.235	0.00880
Gold	0.162	0.0739	0.131	0.150
Oil	0.106	0.242	0.111	0.220
USDTW	- 0.0994	0.274	- 0.117	0.196
1MD	- 0.00807	0.929	- 0.0122	0.893

7.2.1 Trend-followers vs. non-trend-followers

Trend-followers exhibit a correlation pattern similar to that of CTAs, even though it is more pronounced and more coefficients are significant (see Table 26 in the Appendix). For example, the CISDM TF is negatively correlated with the NYSE index. Both trend-followers indices are positively and significantly correlated with the GSCI index. For non-trend-followers (Table 26 in the Appendix) we can notice a series of inconsistencies between Barclay NTF and CISDM NTF. CISDM NTF has a significant positive correlation with corporate bonds, while Barclay NTF correlation is close to zero and significant. Exactly the opposite happens for oil. These inconsistencies once again point out the difficulty of defining non-trend-followers as an asset class. Heterogeneity among this category of funds makes it difficult for a benchmark to truly represent it, questioning the usefulness of such existing indices.

7.3 Linear factor model analysis

In this section, we use a linear factor model to determine CTA's exposure to various types of systemic risk. We run the following regression of CTAs' returns on major asset classes:

$$R_t = \alpha + \sum_k \beta_k F_{k,t} + \epsilon_t \tag{8}$$

where R_t represents the regressed variable (in our case CTAs' returns), α is the intercept, F_{kt} are the independent variables (the kth asset class returns, with $k = 1...11$), β_k is the factor loading, which represents exposure to a particular sytemic risk, and ϵ_t is the error term, or unexplained residual. Sharpe was referring to $\sum_k \beta_k F_{k,t}$ as investment "style" and to $\alpha + \epsilon_t$ as manager's "skill". Sharpe (1992) runs such an analysis on mutual funds and finds that only a limited number of major asset classes is needed to successfully replicate mutual funds' returns.

We begin our study by regressing the CTA indices, proceed to test the explaining power of the model by comparing the predicted values with the actual values for an out-of sample period and conclude by regressing individual CTA fund's returns using the same model.

7.3.1 Linear factor model of CTA indices

We regress both the Barclay CTA and CISDM CTA indices on the 11 aforementioned assets classes. The results are displayed in Table 9. The regressions are significant, but the adjusted R^2 is very low (around 0.14) for both regressions, with only few factors significant at the generally accepted confidence levels. This is not surprising. Consider the following example of a fund manager trading gold futures contracts exclusively. Assuming no leverage is used, a buy and hold strategy on gold futures would result in the regression showing a factor loading of 1 on the gold index. If leverage is used, the factor loading will be equal to the degree of leverage. Conversely, if the manager shorts gold futures, then β_{Gold} will be smaller than 0. If the manager is alternating his positions on gold, trying to catch both up and down-trends, the regression factor will be close to zero. Is this to say that the manager isn't exposed to gold risk? Clearly no. Thus, we can only infer that CTA returns as an asset class are far from being explained by a "buy and hold" strategy of traditional asset classes. This comforts the idea that, even though the underlying assets traded are relatively common, CTAs may offer unique diversification opportunities through their use of highly dynamic trading strategies.

Table 9: Regression of CTA indices on traditional asset classes

	Barclay CTA β	***p*-value**	**CISDM CTA β**	***p*-value**
α	0.00195	0.740	0.00295	0.624
NYSE	0.0824	0.305	- 0.129	0.117
MSCI EAFE	0.00474	0.923	- 0.0331	0.510
MSCI EM	0.0628	0.207	0.0818	0.109
SSBGovt	0.398	0.183	0.339	0.267
SSBCorp	0.444	0.0810	0.491	0.0601
MLHY	- 0.313	0.0316	- 0.236	0.112
GSCI	0.112	0.0596	0.129	0.0345
Gold	0.0489	0.471	0.00698	0.920
Oil	- 0.0332	0.358	- 0.0417	0.259
USDTW	0.305	0.266	0.213	0.448
1MD	0.0177	0.887	0.0301	0.813
Adj. R^2	**0.141**		**0.140**	
***p*-value**	**0.0028**		**0.0029**	
Sum of squared residuals	**0.0573**		**0.0601**	

Trend-followers vs. non-trend-followers The results for the subsamples can be found in Tables 28 and 29 in the Appendix. Both regressions, for trend-followers and non-trend-followers, are significant. High yield bonds is a factor that is significant for both Barclay TF and CISDM TF, while corporate bonds are significant for CISDM TF only. GSCI is significant for Barclay TF. The factor loadings for TF regressions have the same sign as those obtained in the CTAs regression, but they are bigger in absolute value, suggesting that trend-followers are better explained by traditional assets' returns than non-trend-followers. Indeed, adjusted R^2 is 0.15-0.17 for TFs versus 0.12-0.13 for NTFs. In all cases, the adjusted R^2 is very low compared to what a similar analysis on mutual funds would yield. The non-trend-followers regression results are inconsistent compared with those of trend-followers or CTAs. None of the regressors are significant with the exception of high yield bonds for Barclay NTF and emerging markets equity for CISDM NTF. These inconsistencies are most probably due to a very heterogeneous sample of NTFs.

7.3.2 Out of sample test

In order to test the results obtained in the previous section we look at how well the CTA regression models actual CTA benchmarks' returns for an out of sample period. We apply our model, this time using only 8 years of data, to keep 2 years and 3 months of out of sample data. For the fitting period, the adjusted R^2 and the factor loadings significance are similar to those obtained earlier. The plot of the regression model and the actual CTA indices can be seen in Figure 17. For the in sample period, we notice that the regression line is downward biased with respect to the real CTA indices. The regression line does not intersect the CTA index often, staying upward or downward biased for very long time intervals. This is not surprising, given the low explanatory power of the model. For the out of sample period, the regression line succeeds in following the real index for the first few months, after which it starts to go down, while the value of the index goes up. We find this evidence sufficient to state that a linear model based on buy and hold strategies in traditional asset classes completely fails to replicate a diversified portfolio of CTAs.

Trend-followers vs. non-trend-followers Our regression-based models fails to replicate the portfolios of respectively TFs and NTFs out of sample. The plots can be found in Figures 45 and 46.

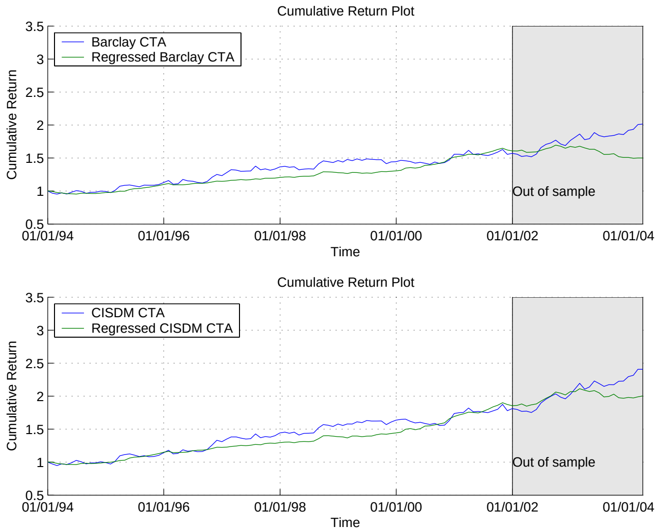

Figure 17: CTA regression out of sample test

7.3.3 Linear factor model for individual CTA funds

We conclude this section by running regressions for 306 individual CTA funds that have been active from June 2000 to March 2004. This restriction was made so as to include more individual CTAs. We follow the approach used in Fung and Hsieh (1997a). Figure 18 summarizes the distribution of the adjusted R^2 for these regressions. Most of the CTAs have an adjusted R^2 between -0.2 and 0.4 with a few outliers. Only 32% of the regressions are significant and the average adjusted R^2 is quite low (0.14). Therefore, not only is a classic linear factor model ill-suited for the analysis of CTAs as an asset class, but it fails to explain anything for an overwhelming majority of CTA funds. Our main task in the rest of this study will be to propose an alternate model to better explain the link between traditional assets and CTAs.

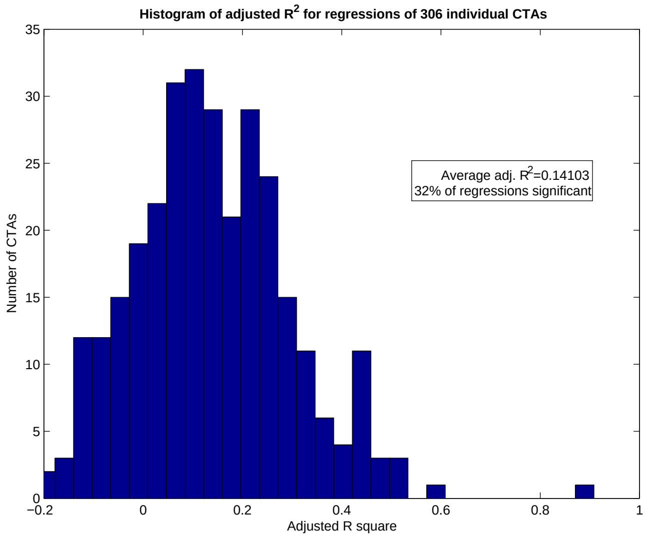

Figure 18: Histogram of adjusted R^2 from 4-year regressions of individual CTAs

Trend-followers vs. non-trend-followers We run the model again for our 2 subsamples of 175 trend-followers and 131 non-trend-followers. The adjusted R^2 distributions are presented in Figures 47 and 48. Average R^2 is approximately the same for trend-followers and non-trend-followers (0.14). Accordingly, the number of regression that are significant is the same for both categories (32%). Perhaps the most interesting aspect to be noticed in these plots is that the adjusted R^2 outliers are non-trend-followers. One non-trend-follower exhibits an adjusted R^2 of 90%. This means that its returns are easily replicable using the traditional assets indices used in our model; this fund should probably be classified as a mutual fund rather than an alternative investment.

7.4 Quantile analysis and conditional correlations

As pointed out by the previous section, the dynamic nature of the strategies employed by CTAs calls for new ways of analyzing their relationship with traditional assets. It is important to realize that all the techniques we have applied so far to investigate such relationships are valid exclusively for linear relationships. As pointed out by Fung and Hsieh (1997a), the high degree of flexibility in the asset choice, together with the relatively high frequency of opening/closing positions and the widespread use of leverage in the CTA industry make for return series that brake the linearity assumption. Building on Agarwal, Daniel, and Naik (2003), Fung and Hsieh (1997a) argue that it is possible to isolate non-linear relationships between CTAs' and traditional assets' return patterns by recognizing the presence of option-like payoffs. In particular, Fung and Hsieh (1997a) and Fung and Hsieh (1999) find that trend-followers exhibit bigger returns when traditional asset markets are at extremes.

In this section, we look for the existence of option-like patterns using simple techniques applied to CTA indices. More specifically, we use quantile analysis to look for option-like patterns in CTAs, using equity, bonds, gold, and a basket of commodities (GSCI) as underlying assets. Our analysis is very similar to the one developed by Fung and Hsieh (1997a). Returns of each asset class are divided into five quantiles that we name respectively: crash, bear, neutral, bull, rally. We compute average returns in each state of nature for CTAs, trend-followers and non-trend-followers as represented by the benchmarks.

The second tool we use in this section is conditional correlation. A correlation coefficient between CTAs and the traditional asset class under review is computed for each of the five previously-defined states of nature. Given the non-normality of CTAs' returns, correlations should be regarded with caution, as pointed out in Kat (2003).

Figure 19 displays the results of the quantile analysis. A common feature of these plots is the high profitabilty of CTAs in both the lowest and the highest quantiles for each asset. No matter if it is bonds, equities, currencies or gold that are experiencing extreme market conditions, CTAs performed well historically. On the other hand, CTAs do not perform well when traditional markets are calm. Overall, we observe a straddle-like payoff pattern, which is consistent with the results of Fung and Hsieh (1997a) . CTAs appear to use market timing strategies in various markets so that they have directional exposure even though they are not correlated with those assets. An interesting fact to notice here is that CTAs' returns in bull markets are different from those in bear markets. This asymmetry in the straddle-like payoff is not present for all markets, but it contradicts some of Fung and Hsieh (1999) results: for instance, CTAs exhibit their best performance in bear equity markets versus bull bond markets. These asymmetries could be the result of investment biases: CTAs seem to catch up-trends more efficiently in bonds or gold. This bias was already captured by the regression analysis: the bonds and gold βs were positive. Overall, we wish to emphasize that the straddle-like payoff observed in our data isn't as clear-cut as the one presented in Fung and Hsieh (1999).

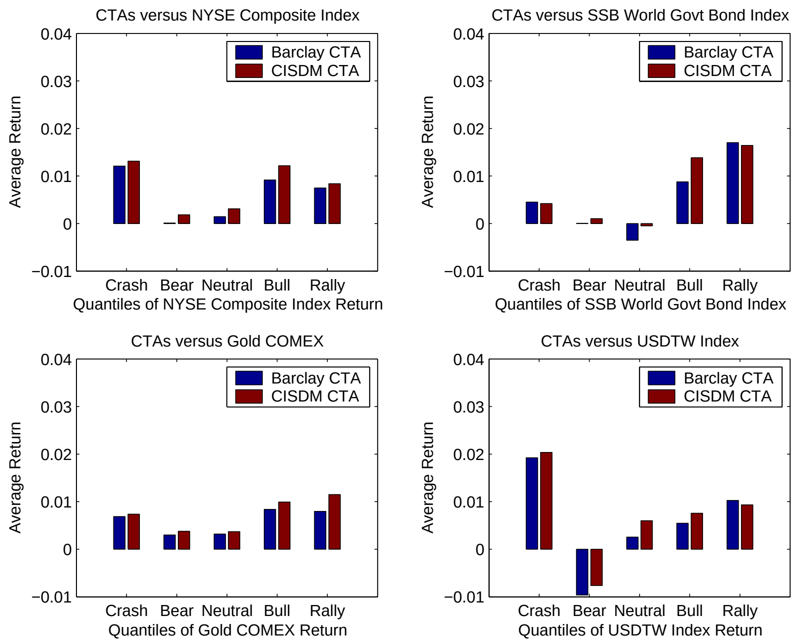

Figure 19: Quantile analysis for CTAs

Figure 20 presents results of the conditional correlation analysis. Correlation is strongly negative in bear markets for all 4 indices. Clearly, CTAs offer an effective insurance for investors with long positions in traditional assets. In bull markets, correlations are positive for bonds and currencies, slightly negative for equities and close to zero for gold. Together with the fact that CTAs have performed relatively well in equity bull markets, the negative correlation means that, in such periods, the main source of returns for CTAs is not to be found in equity markets.

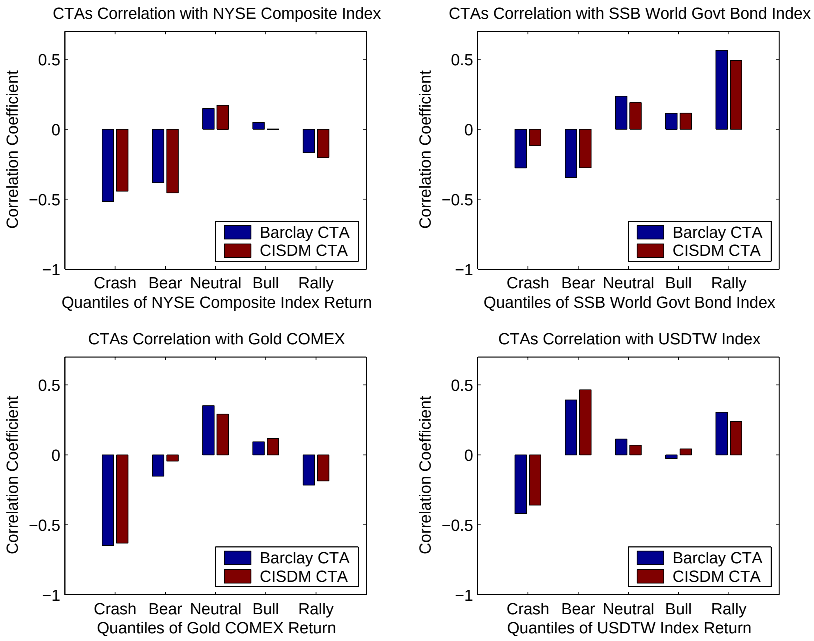

Figure 20: Conditional correlation analysis for CTAs

7.4.1 Trend-followers vs. non-trend-followers

Trend-followers (Figures 49 and 50 in the Appendix) exhibit the same straddle-like payoff as CTAs. Once again, the pattern is not symmetric. Trend-followers seem to catch down-trends more efficiently in equity and currencies , and up-trends in bonds. With the exception of gold, non-trend-followers (Figures 51 and 52 in the Appendix) do not exhibit major sensitivities to traditional assets. Consistent with Fung and Hsieh (1997a), our sample of non-trend-followers exhibits a payoff that looks very much like a call option on gold. This evidence leads us to conclude that non-trend-followers are not the best alternative when it comes to picking CTAs with portfolio insurance in mind.

7.5 Preliminary conclusions

We think it it worth emphasizing here a few broad points our preliminary analysis has led us to.

Firstly, and despite a risk adjusted performance (after correcting for biases) that can hardly compete with other asset classes, CTAs are an interesting investment from the point of view of the asset manager. Including them in a portfolio is an effective way to mitigate the impact of crashes in most financial markets. Additionally, CTAs exhibit positive skewness, both at the individual fund and at the asset class levels. This feature can be used advantageously, especially in the context of a portfolio of hedge funds, as shown in Kat (2004) and Liang (2003). Positive skewness and excess kurtosis also signify that a mean-variance asset allocation framework is not well suited to analyze CTAs, as the oversimplistic assumptions it relies on prevent such a model from capturing some of CTAs most attractive features. With this respect, a portfolio optimisation framework relying on Omega will advantageously be substituted to a classical a la Markowitz approach.

Secondly, CTAs returns are highly unpredictable even when expressed relatively to their peer group's median. While they avoid consistent underperformance, virtually all managers are unable to fulfill a promiss of consistent superior performance. As such, a study focusing solely on CTA funds' returns for the purpose of fund selection seems rather unpromising. On the contrary, we have found that there is undisputable evidence of persistence in the risk profile of CTAs. Thus, any quantitative analysis of CTAs should examine first and foremost a fund's risk profile. As we have shown, this is not an easy task. CTAs' returns are not correlated with that of traditional asset classes, and cannot be explained by a linear model using traditional assets as return generating factors. Nonetheless, it is now clear that CTAs are not immune to movements in traditional markets. CTAs react to movements in traditional assets' value, exhibiting non linear relationships. Even though we have observed option-like payoffs similar to those described in Fung and Hsieh (1999), they were not as obvious. Since, in our opinion, the absence of linear relationships with equities, bonds, currencies, and commodities is due to alternation of long and short positions, as well as rapid movement of funds between markets, the situation calls for a new model. Such a model would ideally capture the dynamic nature of CTA managers' strategies and reflect the systemic risk present in their funds. In the remainder of this study, we design a model with such characteristics.

Lastly, our analysis unraveled one important fact about homogeneity in our two subsamples of trend-followers and non-trend-followers. Whereas trend-followers exhibit similarities among each other, non-trend-followers are characterized by a high degree of heterogeneity. This renders vain any attempt at developing a uniform framework for the analysis of non-trend-followers. Consequently, from now on, our study will focus primarily on trend-followers. Furthermore, let us mention our analysis of non-trend-followers has failed to assertain the existence of a link between risk and return. Given that non-trend-followers seem less effective in providing portfolio insurance or positive skewness, the only rationale for including them in a portfolio of CTAs seems to be driven by potential diversification.

8 Market model for trend-following CTAs

As section 7 has pointed out, models based on static "buy and hold" strategies fail to explain most of CTA funds' returns for a variety of reasons. The model we propose here intends to improve the explanatory power of a market model by building some of trend-following CTA's peculiar features in the regressors so that they represent basic trend-following strategies. Using moving averages, perhaps the most widely used family of momentum strategies, we apply basic trading algorithms to historical futures prices and build trend-following strategies that can be used as return generating factors. Such an approach is not new, Spurgin (1999) and Lequeux and Acar (1998), as well as other researchers have tried to apply similar techniques, some successfully, some not. Using a benchmark constructed in such a way has several advantages over active benchmarks such as the MAR or Barclay indices. Firstly, because it does not rely on CTAs reported returns, it can be updated daily, allowing for closer monitoring of portfolios. Secondly, the benchmark we propose is fully investable and can be perfectly replicated (except for transaction costs and fees). Lastly, the model we propose should allow for a better understanding of systemic risk at the individual CTA level.

The purpose of our model is to benchmark diversified trend-following CTAs, as such, we expect the funds reviewed to employ trend-following strategies across a wide array of commodities, interest rates and currencies. However, by adapting the methodology presented here, it is possible to build benchmarks suitable for specialized CTAs which limit their trading activities to only a small subset of futures contracts.

8.1 Index construction

The first step in building a diversified trend-based trading system is to pick a set of futures contracts that is as representative as possible for the whole universe of available contracts. Once data is retrieved for this subset, prices of expired contracts have to be aggregated in order to obtain a continuous series of prices for which time to maturity is constant. A trend-following algorithm is then applied, and daily returns for a basic trend-following strategy are determined for each contract. Ultimately, simple capital allocation rules are used to combine individual contract strategies into indices, one for each market segment in which CTAs trade. These indices can then be used as return generating factors in the context of a market model of the same form as the one presented in section 7.

8.1.1 Asset selection

We build basic trend-following strategies for the following 27 futures contracts, grouped in 5 market segments. The complete structure of the subset can be found in Table 10. Daily closing prices are gathered for the contracts, starting in year 1990.

Table 10: Asset selection

Currencies	**Equities**	**Short IR**	**Long IR**	**Physicals**
Mexican Peso	Nasdaq 100	3 m. Eurodollar	German Bund	Sugar
U.K. Pound	Nikkei 225	3 m. Euroyen	German Bobl	Corn
Australian Dollar	S&P 500	3 m. Euribor	U.K Long Gilt	Crude Oil
Canadian Dollar	Eurostoxx 50	3 m. Sterling	U.S 2 y. Note	Natural Gas
Swiss Francs	Footsie 100		U.S 10 y. Bond	Copper
Deutsche Mark/Euro				Gold
Japanese Yen				

8.1.2 Building continuous price series

In order to build continuous series of prices for our contracts, we use the technique described in Spurgin (1999). A future's constant maturity price is equal to a linear combination of the prices of the two nearest to maturity contracts. The percentage held in the closest to maturity contract, f_t, is given by equation 9, where DNE stands for the number of days until the first day of the nearby contract's expiration month and DBE for the number of days between the 2 contracts expiration dates. Defining the price of the closest to maturity contract as NX, and the price of the next maturity contract as NB, the price of the continuous, constant-maturity future, P_t, is given by equation 10.

$$f_t = \frac{DNE}{DBE} \tag{9}$$

$$P_t = NB * f_t + NX * (1 - f_t) \tag{10}$$

On each day, a fixed proportion ($\frac{1}{DNE}$) is rolled over from the contract that is closest to maturity to the next contract. This daily rebalancing guarantees that the price series have constant time to expiration, cleaning contango or backwardation evolutions that could otherwise be mistaken as trends.

The daily return from holding such a constant maturity contract can be computed using equation 11, 12 and 13.

$$TotalReturn_t = SpotReturn_t + RollReturn_t \tag{11}$$

$$SpotReturn_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{12}$$

$$RollReturn_t = (NX - NB) * (f_t - f_{t-1}) \tag{13}$$

Let's note that for the Deutsche Mark/Euro contract, we convert all DM prices to Euro equivalents using the historic exchange rate: 1 Euro = DM 1.95583.

8.1.3 Trend-following trading algorithm

As pointed out by Kaufman (1998), moving averages is perhaps the most widely used technique among CTAs to detect trends. There are as many ways to apply moving averages as there are momentum traders. Nevertheless, we believe that when it comes to analyzing CTAs' returns in general, the most simple trading algorithms should provide for maximum explanatory power. We define a h day moving average in equation 14 below.

$$M_t = \frac{\sum_{q=t-h}^{t} P_q}{h} \tag{14}$$

The complete trading-algorithm is made of n moving averages. We define H, which fully characterizes the trading algorithm, as a vector containing the specifications of h for each of the n moving averages. The algorithm generates signals comprised between -100% (fully short) and 100%(fully long) after the close of each day, indicating what position should be taken in a particular contract at the opening of the next day. For each of the n moving averages, the rule is as follows. If the closing price is above the h days moving average, an up-trend is forming and an amount equal to $\frac{1}{n}$ is added to the signal. If the closing price is under the moving average, an amount equal to $\frac{1}{n}$ is substracted from the signal. As an example, if the model relies on 3 moving averages, such that $n = 3$ and H is a vector given by [10 25 135], the corresponding values for the moving averages given by [120 125 122], and that the continuous futures price is $P_t = 121$, the model gives a positive signal for $h = 10$, but two negative signals, for $h = 25$ and $h = 135$. The corresponding signal is thus given by $S_t = \frac{1}{3} - \frac{1}{3} - \frac{1}{3} = -\frac{1}{3}$.

Using multiple moving averages to determine the size of signals is desirable, since it allows for modelling of basic stop-losses, as well as "pyramidal schemes" often used by traders. In an up move, a short term moving average will quickly detect the trend and emit a positive signal, which will lead to an intial position size. If the up trend is sustained, longer term moving averages will detect the trend and require that the original position's size be increased. Such a position piling is often used by traders, and is referred to as a "pyramidal scheme". Stop-loss is also a built-in feature of systems based on multiple moving averages. When the aforementioned uptrend reverses, the short term moving average will quickly emit a negative signal, leading to the downsizing of the position, a fact that can be assimilated to a stop-loss policy.

8.1.4 Single strategy trend-following index

Our strategy implies that every morning the strategy is rebalanced according to the previous day's signal. It is straightforward to compute the returns of a single trend-following strategy using equation 15.

$$R_t = TotalReturn_t * S_{t-1} \tag{15}$$

8.1.5 Risk management & Capital allocation

Traders use various techniques to control the risk of their overall portfolio. One widely used technique is to allocate capital to strategies in a dynamic manner, using risk as the main driver of weights.

Both volatility and liquidity are good candidates for measuring a strategy's risk. For the purpose of this research, standard deviation will be the sole criteria used to determine capital allocation. Specifically, we determine the standard deviation of a strategy on each trading day using an exponentially weighted moving average as described in RiskMetrics (1996). In its recursive form, the formula for computing historical volatility at date t, which will also be our volatility forecast for date $t+1$, is given in equation 16.

$$\sigma_t = \sqrt{\lambda * \sigma_{t-1}^2 + (1-\lambda) * R_t^2} \tag{16}$$

In this formula, λ is the decay factor. It is a measure of the memory of the moving average. A standard value of λ is 0.94, as recommended by RiskMetrics.

Most CTA managers will readily admit that capital allocation and risk management procedures are key to performance. Thus, it represents the crucial element of the model. We chose to allocate capital among single contract strategies using weights determined by relative volatility. The weight for the ith contract of the kth group is specified by formulas 17 and 18, where n is the number of funds in the group of the ith asset.

$$x_{i,t} = \frac{1}{n * \sigma_{i,t}} \tag{17}$$

$$w_{i,t} = \frac{x_{i,t}}{\sum_{i=k(1)}^{k(n)} x_{i,t}} \tag{18}$$

8.1.6 Trend-following indices

The last step in building our 5 trend-following indices (one for each market segment), is to combine the single strategies according to the capital allocation. Since the weights are based on a volatility forecast, there is a 1 day lag between the computation of the weights and the actual implementation of the updated capital allocation. The daily return on index k can therefore be computed using equation 19, which is a weighted average of the returns on the n single trend-following strategies comprised in the kth group.

$$I_{k,t} = \sum_{i=k(1)}^{k(n)} w_{i,t-1} * R_{i,t} \tag{19}$$

8.1.7 Limitations

We wish to mention a few limitations of the model. First, the strategies do not take into account transaction costs nor fees, so that their returns cannot be directly compared to CTAs' returns. Additionally, the strategies do not include interests that would be paid on a margin account deposit. Indeed, taking positions in futures markets does not require a financing equal to 100% of the nominal value of the contracts. Nevertheless, for our strategies, we assume that investments are made on a fully collateralized basis.

8.2 Application to individual fund analysis

The form of the market model and the methodology we use in the remaining of this section are identical to the one used in section 7.3, which will permit comparisons. With the new factors, the model becomes:

$$R_t = \alpha + \sum_k \beta_k I_{k,t} + \epsilon_t \tag{20}$$

8.3 Model Calibration

The first step in the application of our model is to calibrate it. There are several parameters that need to be determined. Most importantly, we need the combination of moving averages to use so as to maximize the explanatory power of the model across CTAs. Then, the optimal λ, on which capital allocation depends, needs to be pinned down. When calibrating, we look at two criteria, namely average adjusted R^2 and the percentage of significant regressions over the period June 2000 to March 2004. We regress only the 175 trend-followers that traded throughout this period.

8.3.1 Moving averages

In order to determine the mix of moving averages that will lead to superior explanatory power, we look at the performance of models based on only 1 moving average. We vary the parameter M_t of equation 14 by varying the number of days used to compute it. Figure 21 presents the results of this operation. Observation of this graph leads us to believe that a model using 3 moving averages, with H given, in vector form, by [2 70 240], should capture most of the trends CTAs focus on, and lead to a superior adjusted R^2.

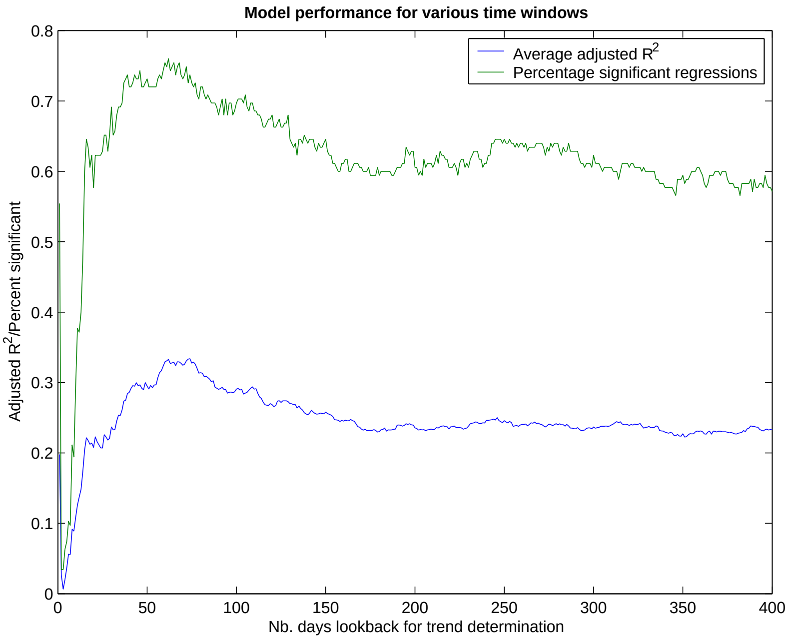

Figure 21: Performance of the model for various moving averages

8.3.2 Volatility

Since capital allocation between single contract trend-following strategies is done according to volatility, the parameter λ from equation 16 needs to be calibrated. The calibration procedure is similar to the one presented above. The results are shown in Figure 22. We can infer that optimum R^2 can be obtained for high values of λ, values that are much higher than the RiskMetrics documents recommended value of 0.94. This means the volatility used to compute the weights is relatively stable, since the exponentially weighted moving average has a long memory. The value we select for the decay factor is 0.99. In all cases, the specification of the decay factor does not appear to influence notably the overall explanatory power of the model, and as such, there seems to be small room for improvement in this area.

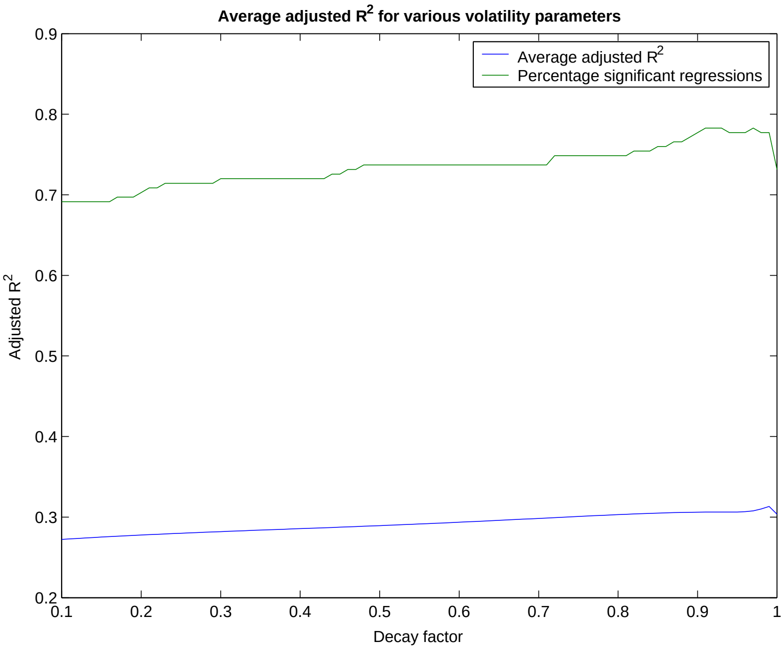

Figure 22: Performance of the model for various decay factors

8.4 The factors

Figure 23 depicts the evolution of the index for each of the 5 market segments. Descriptive statistics can be found in Table 11. Correlation between the indices is given in Table 12.

All strategies led to long term positive returns. Trend-following strategies are characterized by low return and low standard deviation, except for the strategy on stocks. The indices are uncorrelated with each other, with the exception of the 2 interest rates strategies.

8.5 Explanatory power of the model for individual trend-followers

We run the calibrated model on the 175 trend-followers, regressing their returns over the period June 2000 to March 2004. Figure 24 is a histogram of adjusted R^2. We also report

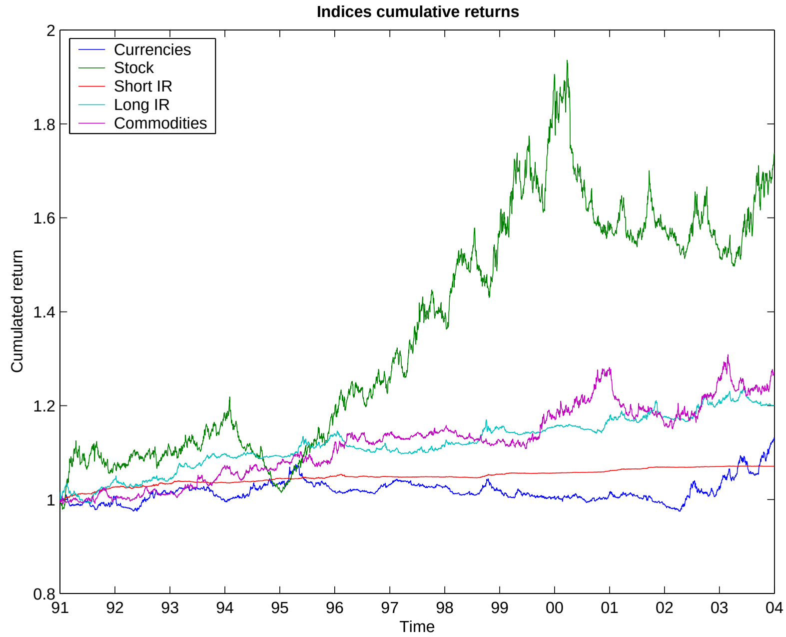

Figure 23: Evolution of indices for 5 market segment over time

Table 11: Trend-following index descriptive statistics

	Currencies	**Stock**	**Short IR**	**Long IR**	**Commodities**
Avg. return	0.96	4.34	0.53	1.41	1.86
Standard dev.	2.77	8.19	0.4	2.18	3.87

average adjusted R^2 as well as the percentage of significant regressions.

Our model offers considerable improvement over the model presented in section 7.3.3. The average adjusted R^2 is more than double the one offered by a model based on "buy and hold" strategies. The percentage of significant regressions also soared from 32% to 76%. Whereas in the classic asset case none of the regressions' R^2 reached 0.5, the model we propose explains about a third of the funds with an R^2 superior to 0.5. Nevertheless, the model's performance is mitigated by the fact that regression fails to explain anything in about 25% of the cases, and that for many funds, the R^2 is very low. We believe that this is due to the diversity of time windows used by funds for the purpose of trend determination, as well as the apparent weakness of our capital allocation procedure, where improvements must be

Table 12: Trend-following index correlation matrix

	Currencies	**Stock**	**Short IR**	**Long IR**	**Commodities**
Currencies	1	0.00	0.07	0.17	0.12
Stock	0.00	1	0.06	0.15	0.03
Short IR	0.07	0.06	1	0.45	-0.02
Long IR	0.17	0.15	0.45	1	-0.01
Commodities	0.13	0.03	-0.02	-0.01	1

made. Additionally, the presence in our data of funds that are not fully diversified lowers the overall explanatory power of the model.

Figure 25 shows the histogram of average adjusted R^2 for the whole sample of CTAs. As expected, the explanatory power of the regressions is lower than for trend-followers only, but still represents a major improvement over the classical model of section 7.3.3.

8.5.1 Fund selection methods

We apply here two selection methods based on our model. In the first method, we select funds with superior α (above the median value for the period) and test if they outperformed out of sample. In the second method, developed in Chen and Passow (2003), we select funds with low overall exposure to the factors. More precisely, our second selection retains only funds that have an exposure to each factor lower than the median exposure to that factor. The argument backing up the method goes as follows: in order to isolate α, "informed traders" are expected to try to be neutral to the risk factors. Such managers may be able to generate returns independent of CTA performance. They are expected to exhibit higher consistency of returns, higher Sharpe ratios, as well as an improved performance persistence.

Both selection methods are applied on trend-following funds exclusively. We first compute α and factor exposures for all funds, running the regression on the period June 2000 to December 2002. We then look at the performance of the selected funds versus that of the funds that weren't selected over the period. Specifically, we look both at absolute average

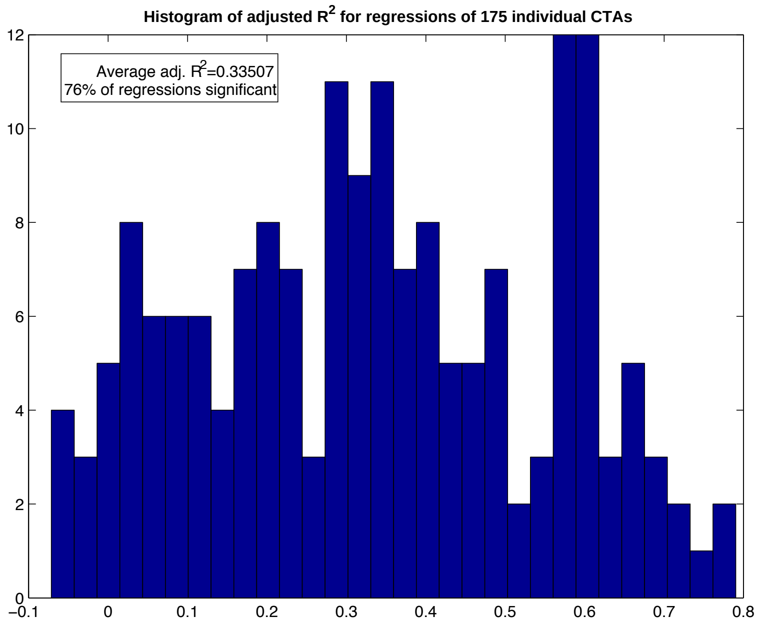

Figure 24: Histogram of adjusted R^2 from 4-year regressions of individual trend-followers

return and Sharpe ratio for both groups during the period January 2003 to March 2004. 2 sample t-tests are performed to test for significance of the two samples difference in mean.

The results of the first selection method, based purely on α, can be seen in Figures 26 and 27. Clearly, this selection method is dissappointing. The selected funds exhibit both lower absolute and risk-adjusted performance than the funds left out in period 2. The differences between the 2 sample means are not statistically different from 0 though. We conclude that α as defined in our model is yet another performance measure that does not persist across time periods.

The results of the second selection method, based on low overall factor exposure, can be seen in Figures 28 and 29. Again, the selection criteria used leads to overall underperformance. The selected funds exhibit much lower absolute return, a fact that was expected, since the out of sample period was bullish for trend-followers in general. Unfortunately, risk-adjusted performance of the selected funds is mediocre, and significantly underperforms that of the funds not selected. This result is driven by the returns of the selected funds which

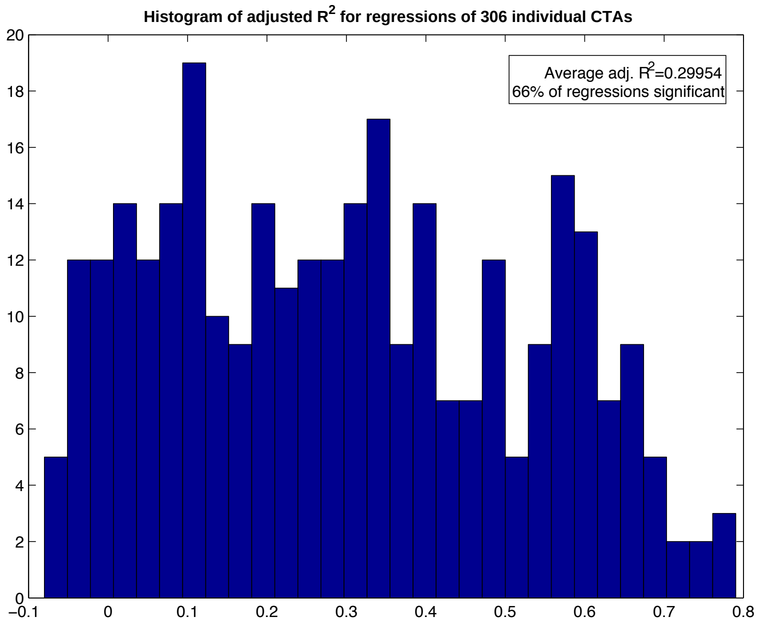

Figure 25: Histogram of adjusted R^2 from 4-year regressions of individual CTAs

tend to be very low, often driving sharpe ratios below 0. Unlike in Chen and Passow (2003), where the authors apply the method successfully to long-short equity funds, it seems that CTAs' purely skill driven returns (α) are very small and fail to compensate for idiosyncratic risk as defined by our model.

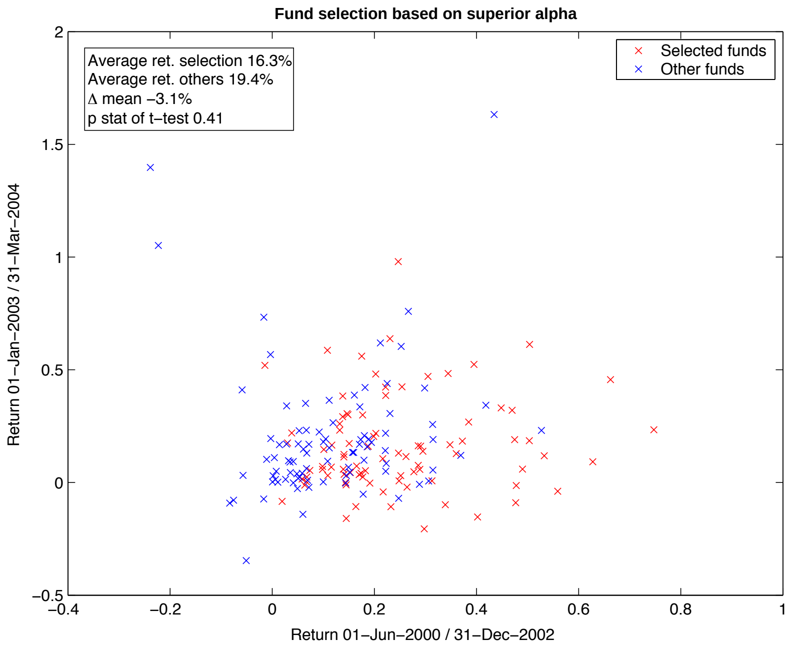

Figure 26: Absolute return comparison for α based selection method

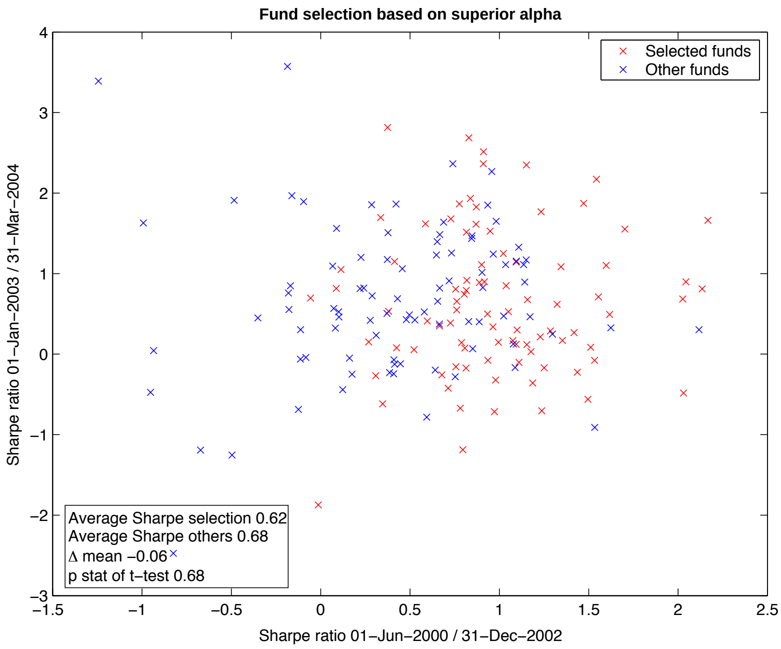

Figure 27: Risk-adjusted performance comparison for α based selection method

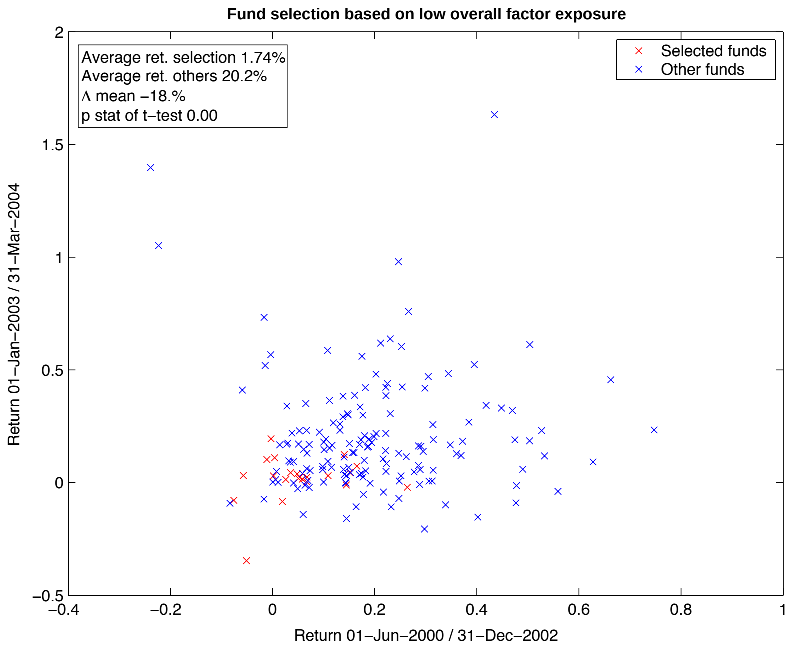

Figure 28: Absolute return comparison for low overall exposure selection method

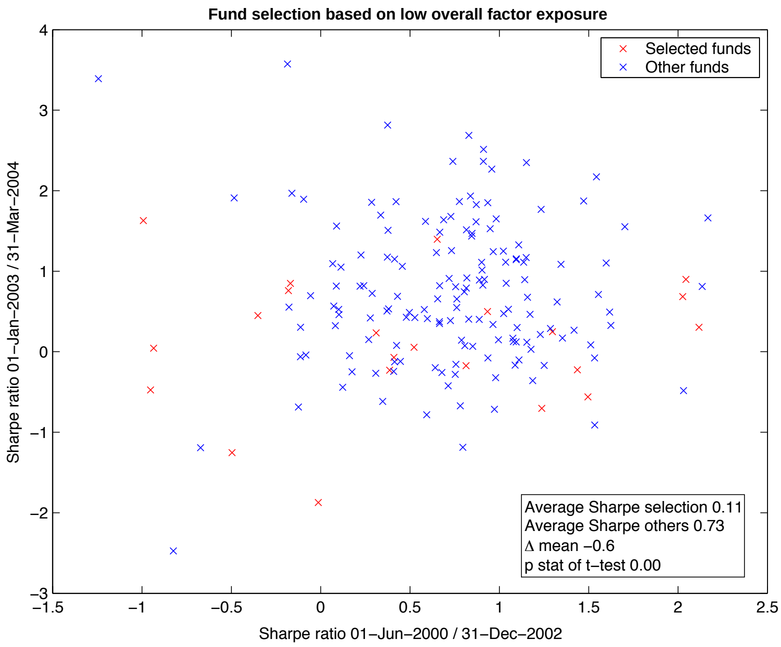

Figure 29: Risk-adjusted performance comparison for low overall exposure selection method

8.6 Explanatory power of the model for CTAs as an asset class

Using the model as calibrated in section 8.3, and in an attempt to explain returns of CTAs as an asset class, we use regression analysis on various CTA indices.

8.6.1 Out of sample test

Using 8 years of data, we compute the exposures of the Barclay and CISDM CTA indices to our trend-following indices and perform an out of sample test similar to the one used in section 7.3.2. We compute the returns of a virtual portfolio of equal exposure for the 2-year period left out. Figure 30 presents the results of the test. The correlation level achieved with the Barclay CTA index out of sample is as high as 0.88, while it is 0.85 with respect to the CIDSM CTA index, indicating only a small amount of overfitting due to the calibration. The figure that we report for tracking error seems quite high when one looks at the impressive goodness of fit of the model. This is partly due to the way yearly standard deviation is computed from monthly data and then scaled. Months when our replicating index underperforms the benchmark are compensated shortly by months of overperformance, which results in the 2 graphs intersecting often.

We find it interesting to push the experience further and try to see how long a virtual trend-following index, based on a calibrated regression, can replicate an active CTA index. Figure 31 plots such a virtual index, together with Barclay and CISDM CTA indices. We choose a 2-year period for calibrating the model and build the return series for our virtual index from 1996 to the end of 2003. The model still performs very well in such a setting. Correlations as high as 0.67 indicate that global risk exposure of CTAs as an asset class stays mostly unchanged through time.

Trend-followers vs. non-trend-followers We run the same out of sample test for both subsamples. The results are shown in Figures 53 and 54. As expected, the model performs very well for trend-followers. The synthetic index exhibits correlations in the 0.8-0.9 range with trend-followers index. Not surprisingly, the performance of the model is rather poor for non-trend-followers, with correlations droping to the 0.2-0.4 range.

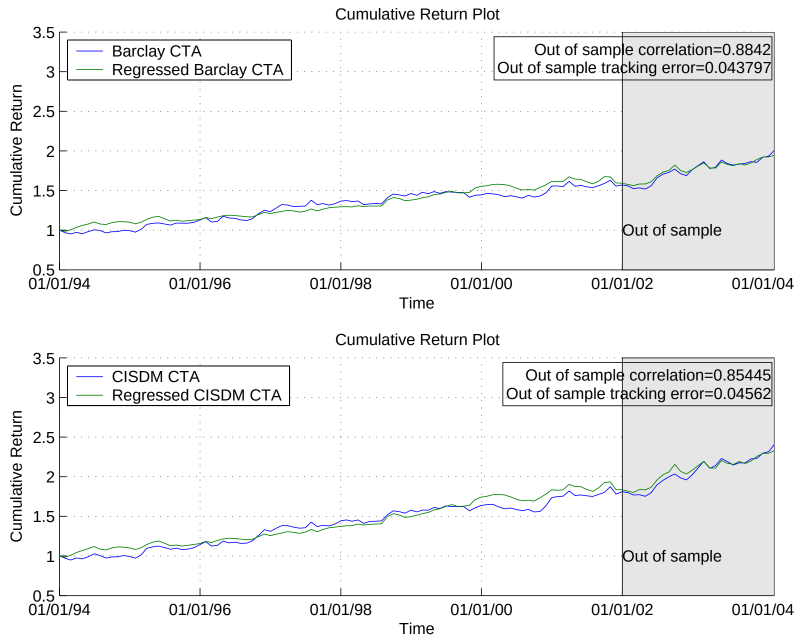

Figure 30: Out of sample performance of model for regression of CTA indices

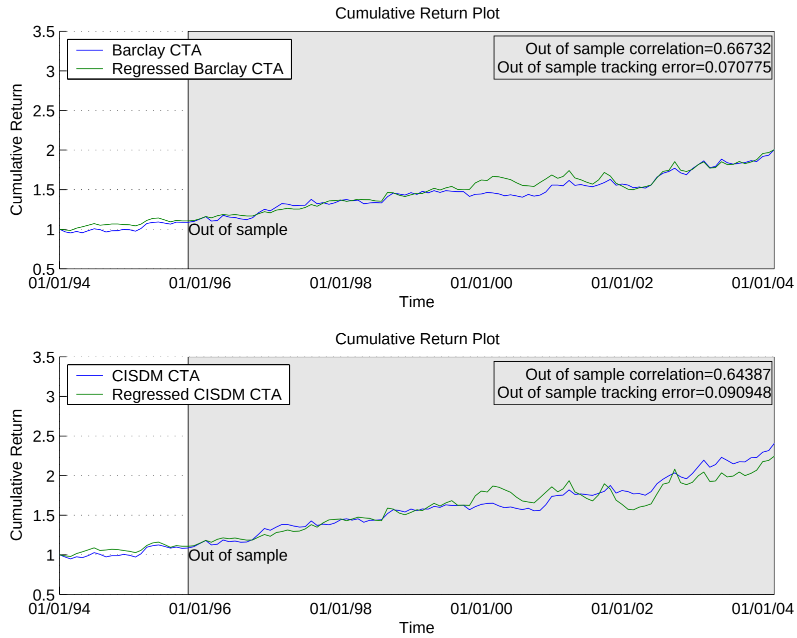

Figure 31: Long out of sample performance of model for regression of CTA indices

8.6.2 Parameter and explanatory power stability

In order to investigate the stability of the factor loadings, as well as the explanatory power of the model, we run 2-year rolling window regressions of the Barclay CTA index. Figure 32 depicts the evolution of the factor loadings for our 5 indices, as well as the explaining power of the model measured by R^2. The model performs relatively well, no matter what time period is used. The major exception to this is the period 1999 to 2001, when R^2 drops to a low of 0.24. Recently though, the model has been performing extremely well, with a maximum R^2 of 0.92.

Table 13 reports descriptive statistics for the factor loadings. In this table, the third row provides the proportion of all regressions in which exposure to the factor was found to be significantly different from 0. This proportion is of only 2% for the short interest rate index's loading. Our model would probably benefit from its elimination. Clearly, currencies and long interest rate exposures are what make up most of CTA returns. Exposure to the stock and commodities indices are somewhat lower and significant in less cases, but they still play a role in replicating CTA returns as measured by the Barclay CTA index. Regarding risk exposure's stability, as can be seen from Figure 32, the loadings are contained roughly between 0 and 2. Exposure to the stock factor seems to have decreased monotonically, while long interest rate exposure exhibits rapid swings, such as the one that occured at the end of 2001. An other interesting figure is that of average leverage. Eliminating the loadings that are not significant, we find an average leverage of approximately 2.5, a figure consistent with Spurgin (1999).

Table 13: Descriptive statistics for factor loadings

	$\beta_{Currencies}$	β_{Stock}	$\beta_{ShortIR}$	β_{LongIR}	$\beta_{Commodities}$
Average	0.88	0.14	2.55	1.16	0.29
Standard dev.	0.28	0.20	2.82	0.56	0.26
Significance	67%	16%	2%	56%	18%

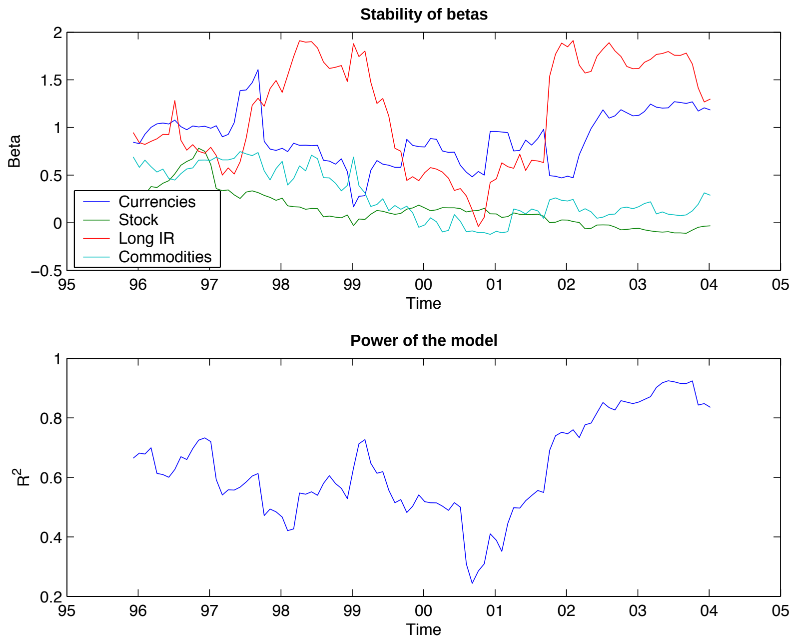

Figure 32: Stability of factor loadings and explanatory power of the model

9 Conclusions

The various analysis conducted throughout this study have helped to better understand the nature of CTAs' returns at various levels.

At the return distribution level, our study has confirmed the presence of both significant positive skewness and excess kurtosis for individual commodity trading advisors. Moreover, skewness was shown to be more pronounced in the case of trend-followers, while non-trend-followers exhibited lower standard deviation together with higher kurtosis. This finding has stressed the need for portfolio allocation techniques that capture some of CTAs most attractive features. In this sense, approaches based on Omega will be preferred over mean-variance optimisation which may penalize CTAs by allocating to them suboptimally.

A significant number of individual funds were shown to exhibit either positive or negative serial correlation for various time lags. Evidence of negative autocorrelation was found for trend-followers as a class for a 2 months lag, creating opportunities for contrarian strategies. With respect to the persistence of distribution parameters, it was clearly demonstrated that CTAs' volatility persists in both the short and long term. This fact is driven by both sound risk management practice and funds' desire to target investors characterized by a specific risk aversion profile. Despite hints of persistence for other measures, such as return, simplified sharpe ratio, skewness and Omega, the evidence from various significance tests was mixed. Kurtosis didn't seem to persist in our dataset.

In a slightly different setting, we were unable to ascertain the persistence of performance among CTA funds, whether measured in absolute or risk adjusted terms. Applying two different selection methods, the set of funds we selected underperformed its peers. In one case, the selection method led to significant risk-adjusted underperformance, suggesting the existence of an outperforming contrarian strategy.

With respect to risk premia, we have shown that CTAs have paid in recent history an average of 25 basis points for every percentage point of additional risk taken (as measured by standard deviation). Whereas a risk premium of 34 basis points was found for trend-followers, no significant risk premium was found for non-trend-followers despite the relatively long period under review (1994-2002).

Our analysis of various CTA benchmarks pointed out the relative homogeneity of trend-followers. Indeed, trend-followers benchmarks exhibited strong correlation with each other. General CTA benchmarks exhibited similar features. On the other hand, non-trend-follower benchmarks were found to be mildly correlated with each other at most, pointing out the heterogeneity of a class in which managers employ strategies that can vary widely. This fact was confirmed by a correlation analysis of individual non-trend-followers. None of the funds were found to have a correlation coefficient superior to 0.5 with either of the Barclay or CISDM non-trend-followers indices.

The second part of the study has focused on linking CTAs returns with that of traditional assets. We showed that a "buy and hold" strategy of classical assets was ill-suited to explain such relationships. We concluded that the dynamic nature of CTA strategies affects the ability of linear models to isolate CTA funds' systemic risks. We pointed out the existence of option-like payoffs in CTA return patterns. Indeed, CTAs were shown to react in a non-linear fashion to movements in traditional markets. Historically, CTAs, driven by trend-followers, have performed well when stock, bond, gold or currency markets were experiencing extreme movements. In relatively calm markets though, CTAs haven't performed well. Overall, trend-followers' payoff patterns resemble that of an option straddle, providing effective portfolio insurance, while keeping a significant upside exposure to traditional assets markets.

In the last part of our research, we proposed a linear model in which factors capture the dynamic nature of CTA managers' strategies. Using trading algorithms based on moving averages, we created daily historical returns for trend-following strategies on 27 futures contracts. In a second step, the single-contract strategies were aggregated into 5 market segment factors, in an attempt to replicate traders' typical capital allocation practices. The model was shown to be clearly superior to a model based on "buy and hold" strategies. Applied at the individual trend-follower level, the average explanatory power reached an adjusted R^2 of 0.33. Applied at the asset class level, the model performed very well, out of sample correlations with various CTA benchmarks reaching the 0.8-0.9 range.

We feel our research was successful in clarifying the type of risks CTAs are exposed to. More work needs to be done, especially with respect to more relevant capital allocation techniques. Also, trend-following indices needs to be built for specialised CTAs. However, we believe the insights provided by our model should allow for more efficient risk management and asset allocation among the CTA class. Purely quantitative approaches failed to provide techniques for selection of overperforming CTAs. Approaches based on operational variables, such as assets under management growth or fee structure should be explored. Specifically, the approach developed in Diz (1999) may be more successful at picking future outperformers.

References

Agarwal, Vikas, Naveen D. Daniel, and Narayan Y. Naik, 2003, Flow, performance and managerial incentives in the hedge fund industry, *EFA Annual Conference Paper* 501.

Agarwal, Vikas, and Narayan Y. Naik, 2000a, Multi-period performance persistence analysis of hedge funds, *Journal of Financial and Quantitative Analysis* 35, 327–342.

——— , 2000b, On taking the 'alternative' route: Risks, rewards style and performance persistence of hedge funds, *Journal of Alternative Investments* 2, 6–23.

Amenc, Noel, and Lionel Martellini, 2002, The brave new world of hedge fund indexes, *Working paper, Edhec.*

Brown, Stephen J., and William N. Goetzmann, 1995, Performance persistence, *Journal of Finance* 50, 679–698.

——— , Roger G. Ibbotson, and S. A. Ross, 1992, Survivorship bias in performance studies, *Review of Financial Studies* 5, 553–580.

Brown, Stephen J., William N. Goetzmann, and James M. Park, 1997, Conditions for survival: Changing risk and the performance of hedge fund managers and ctas, *Yale International Center for Finance Working Paper* F-59.

Carpenter, Jennifer N., and Anthony W. Lynch, 1999, Survivorship bias and attrition effects in measures of performance persistence, *Journal of Financial Economics* 54, 337–374.

Cerrahoglu, Burak, 2004, The benefits of managed futures, *Center for International Securities and Derivatives Markets, Isenberg School of Management, Working Paper.*

Chen, K., and A. Passow, 2003, Quantitative selection of long-short hedge funds, *International Center for Financial Asset Management and Engineering, University of Lausanne.*

CISDM, 1999, The benefits of managed futures, *Alternative Investment Management Association, working paper.*

Diz, Fernando, 1999, How do ctas' return distribution characteristics affect their likelihood of survival?, *Journal of Alternative Investments* 2, 37–41.

Fienberg, Stephen E., 1980, The analysis of cross-classified categorical data, *MIT Press, Cambridge, Mass.*

Fung, William, and David A. Hsieh, 1997a, Empirical characteristics of dynamic trading strategies: The case of hedge funds, *The Review of Financial Studies* 10, 275–302.

——— , 1997b, Survivorship bias and investment style in the returns of ctas, *Journal of Portfolio Management* 24, 30–41.

——— , 1999, A primer on hedge funds, *Journal of Empirical Finance* 6, 309–331.

Goetzmann, William N., and Roger G. Ibbotson, 1994, Do winners repeat? patterns in mutual fund return behavior, *Journal of Portfolio Management* pp. 9–17.

Habib, Rami, 2004, A look inside the black box, *Working paper.*

Kat, Harry M., 2003, The dangers of using correlation to measure dependence, *Journal of Alternative Investments* 6, 54–58.

——— , 2004, Managed futures and hedge funds: A match made in heaven, *Journal of Investment Management* 2.

——— , and Faye Menexe, 2003, Persistence in hedge fund performance: The true value of a track record, *Journal of Alternative Investments* 5, 66–72.

Kaufman, Perry J., 1998, Trading systems and methods, *John Wiley & Sons.*

Keating, Con, and William F. Shadwick, 2002a, The omega function, *The Finance Development Centre Limited.*

——— , 2002b, A universal performance measure, *The Finance Development Centre Limited.*

Lequeux, P., and E. Acar, 1998, A dynamic index for managed currencies funds using cme currency contracts, *The European Journal of Finance* 4, 311–330.

Liang, Bing, 2003, On the performance of alternative investments: Ctas, hedge funds, and funds-of-funds, *Isenberg School of Management, Working Papers Series*.

Malkiel, Burton G., 1995, Returns from investing in equity mutual funds 1971 to 1991, *Journal of Finance* 16, 549–558.

Park, James M., 1995, Managed futures as an investment set, *Ph.D Dissertation, Columbia University*.

RiskMetrics, 1996, Technical document, *Fourth Edition*.

Schneeweiss, Thomas, and Richard Spurgin, 1997, Comparisons of commodity and managed futures benchmark indices, *Journal of Derivatives* 4, 33–50.

——— , 1998, Multifactor analysis of hedge fund, managed futures, and mutual fund return and risk characteristics, *Journal of Alternative Investments* 1, 1–24.

——— , and David McCarthy, 1996, Survivorship bias in commodity trading advisor performance, *Journal of Futures Markets* 16, 757–772.

Schwager, Jack D., 1996, Schwager on futures: Managed trading myths & truths, *John Wiley & Sons*.

Sharpe, William F., 1992, Asset allocation: Management style and performance measurement, *Journal of Portfolio Management* 18, 7–19.

Spurgin, Richard, 1999, A benchmark for commodity trading advisor performance, *Journal of Alternative Investments* 2, 11–21.

A Asset Classes Indices

Goldman Sachs Commodity Index: The GSCI, officially launched in 1992, is a measure of the performance of highly liquid, dollar-denominated commodity futures contracts. There are currently 24 commodities in the index. The weights assigned to individual commodities are based on a five-year moving average of world production. Weights are determined each July and made effective the following January. All contracts are rolled on the fifth business day of the month prior to the expiration month of the contract. Two versions of the index are available: a total return version, which assumes that capital necessary to purchase the basket of contracts is invested at the risk-free rate, and a spot version, which only tracks movements in the futures prices. This study uses the total return measure.

Merrill Lynch Global High Yield Index: The MLHY index tracks the performance of bonds considered to be below investment-grade, as defined by Moodys' ratings.

MSCI EAFE & Canada Index: The EAFE index contains stocks of companies listed on the major stock exchanges of the following countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Hong Kong, Ireland, Italy, Japan, Malaysia, Netherlands, New Zealand, Norway, Singapore, Spain, Sweden, Switzerland, and the UK. The U.S. is excluded. The index represents a sampling of large, medium, and small capitalization companies from each market, taking into account the stocks' liquidity in determining each components' weights.

Morgan Stanley Capital Emerging Market Index The MSCI EM index measures performance of a universe of 28 emerging markets. The index is constructed using market capitalization weights and a consistent 60% target market coverage. A number of factors determine wheter a country may be recognized as an emerging market, the most important being a country's per capita gross domestic product.

The NYSE Composite Index: The NYSE index measures performance of all common stocks listed on the New York Stock Exchange. The index tracks the change in aggregate market values of NYSE common stocks, adjusted to eliminate the effects of capitalization changes, new listings and de-listings.

Salomon Smith Barney Corporate Bond Index:: The SBB Corp. index includes both US and non-US corporate bonds, as well as non-US sovereign and provincial securities.

Salomon Smith Barney World Government Bond Index:: The SBB Govt tracks the performance of the 19 following government bond markets: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States. The components of the index are weighted according to their market capitalization.

B Figures

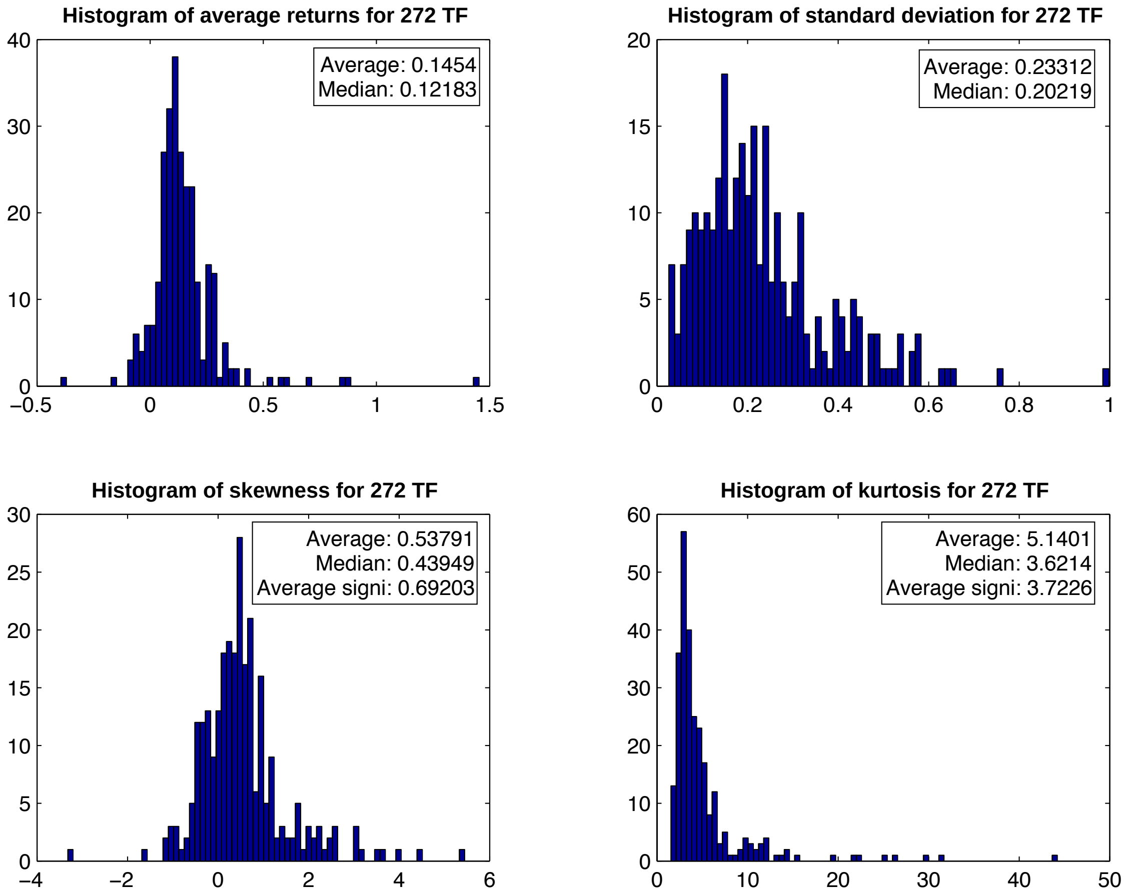

Figure 33: Histograms of first 4 moments of trend-followers' return distributions

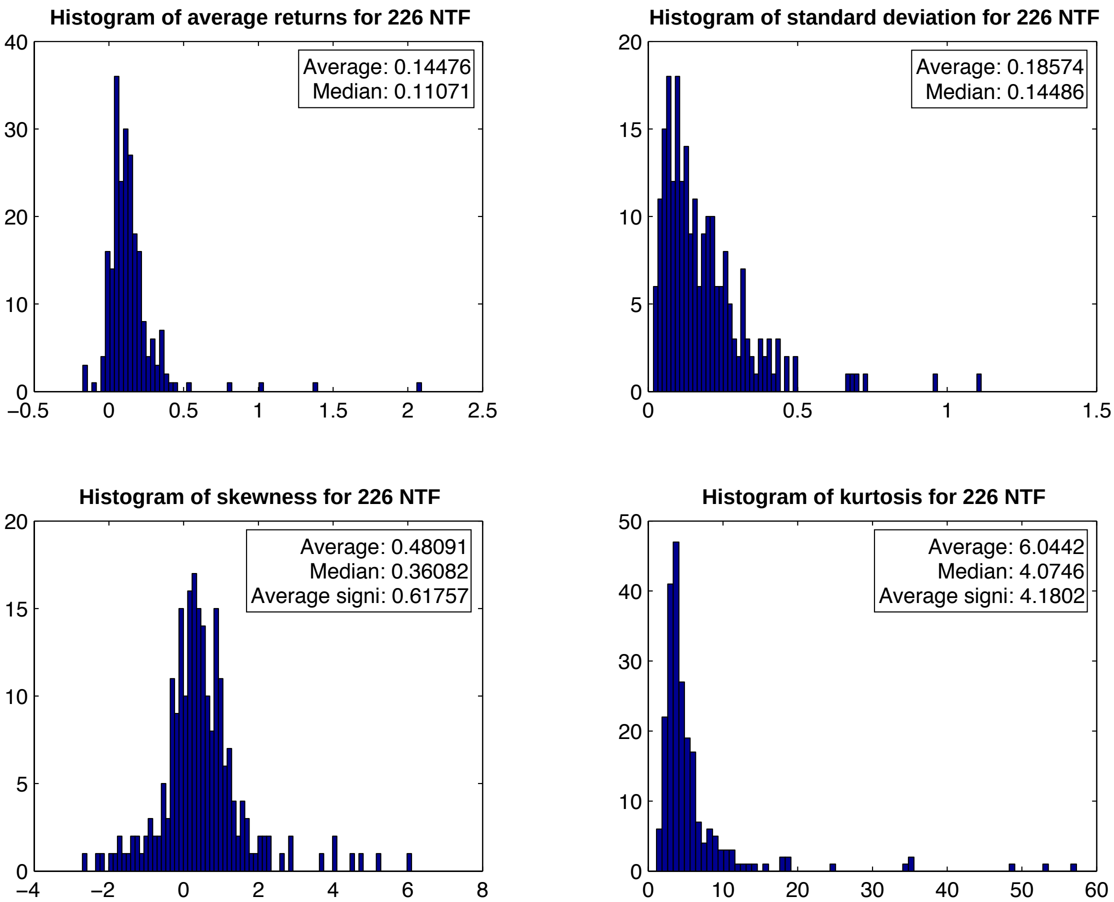

Figure 34: Histograms of first 4 moments of non-trend-followers' return distributions

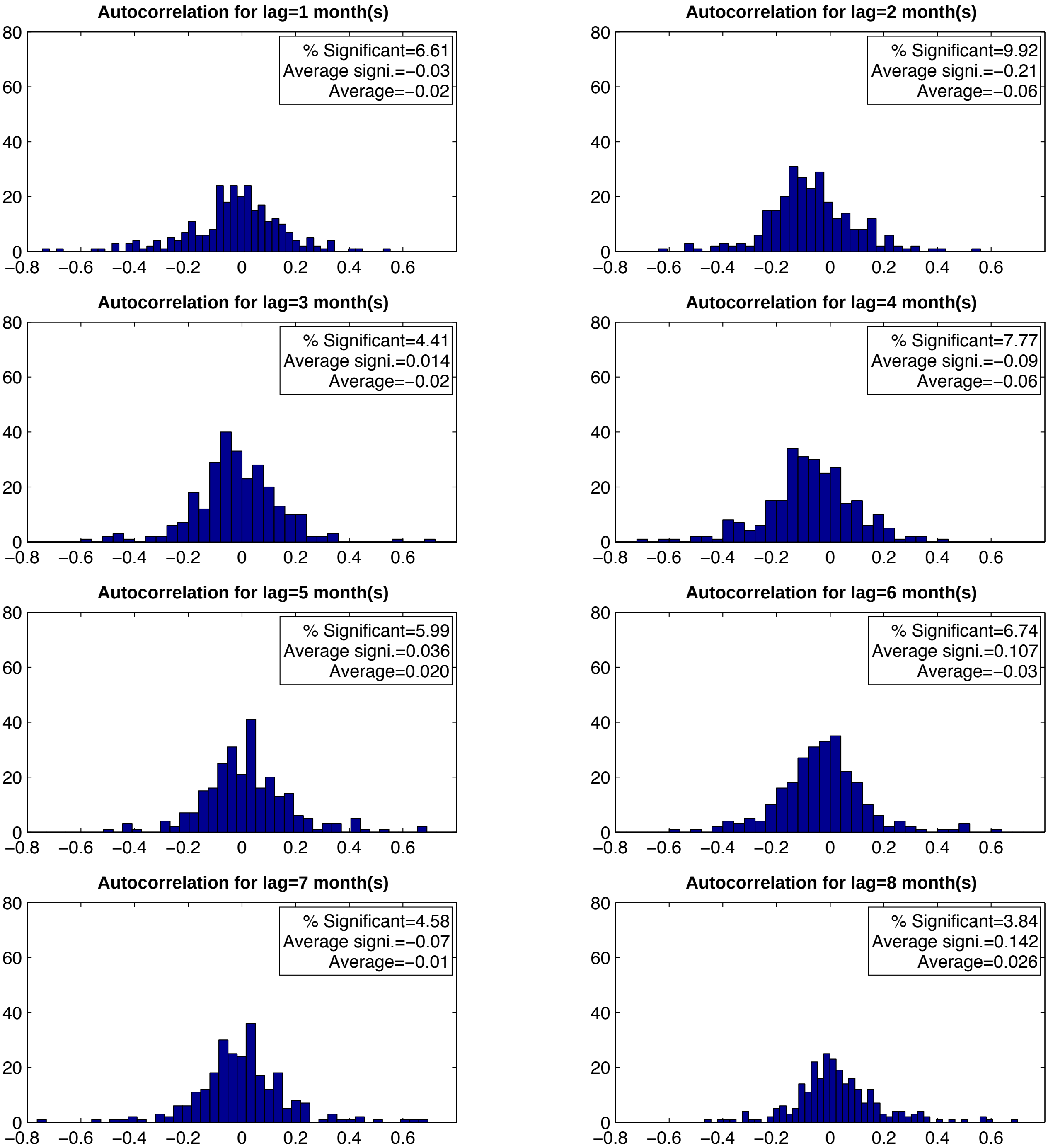

Figure 35: Histograms for trend-followers' autocorrelation for lags between 1 and 8

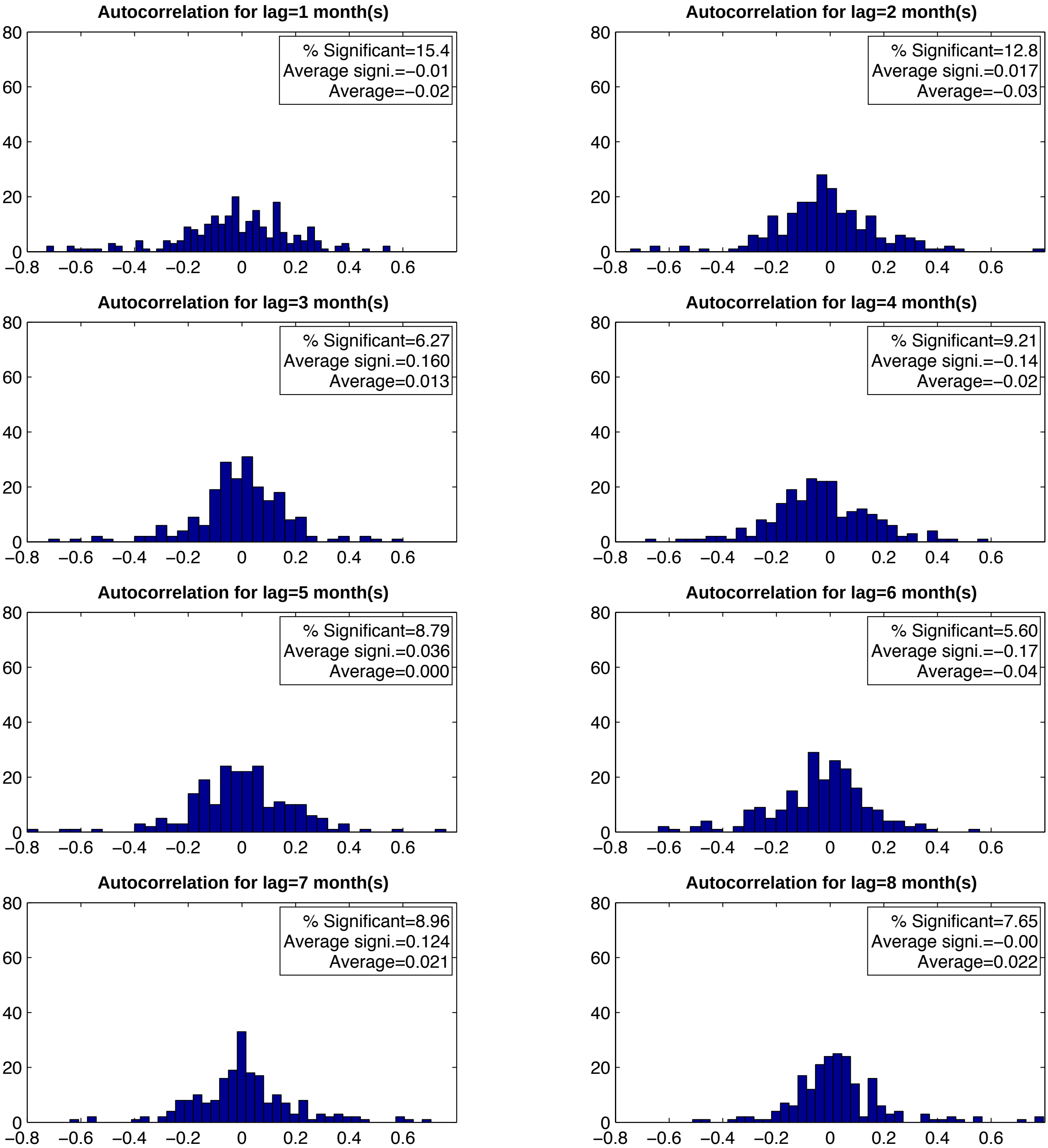

Figure 36: Histograms for non-trend-followers' autocorrelation for lags between 1 and 8

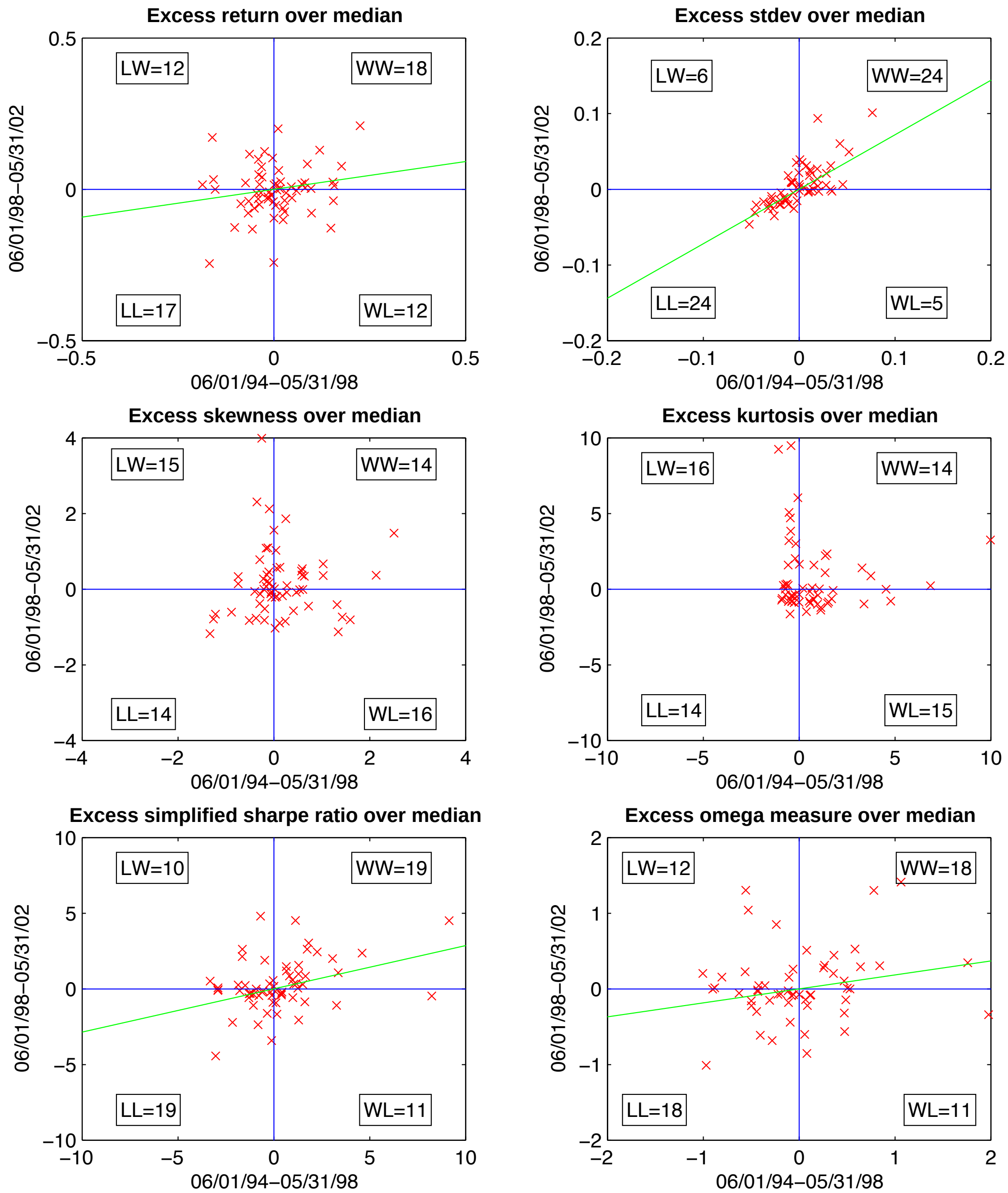

Figure 37: Persistence in trend-followers' return distributions parameters

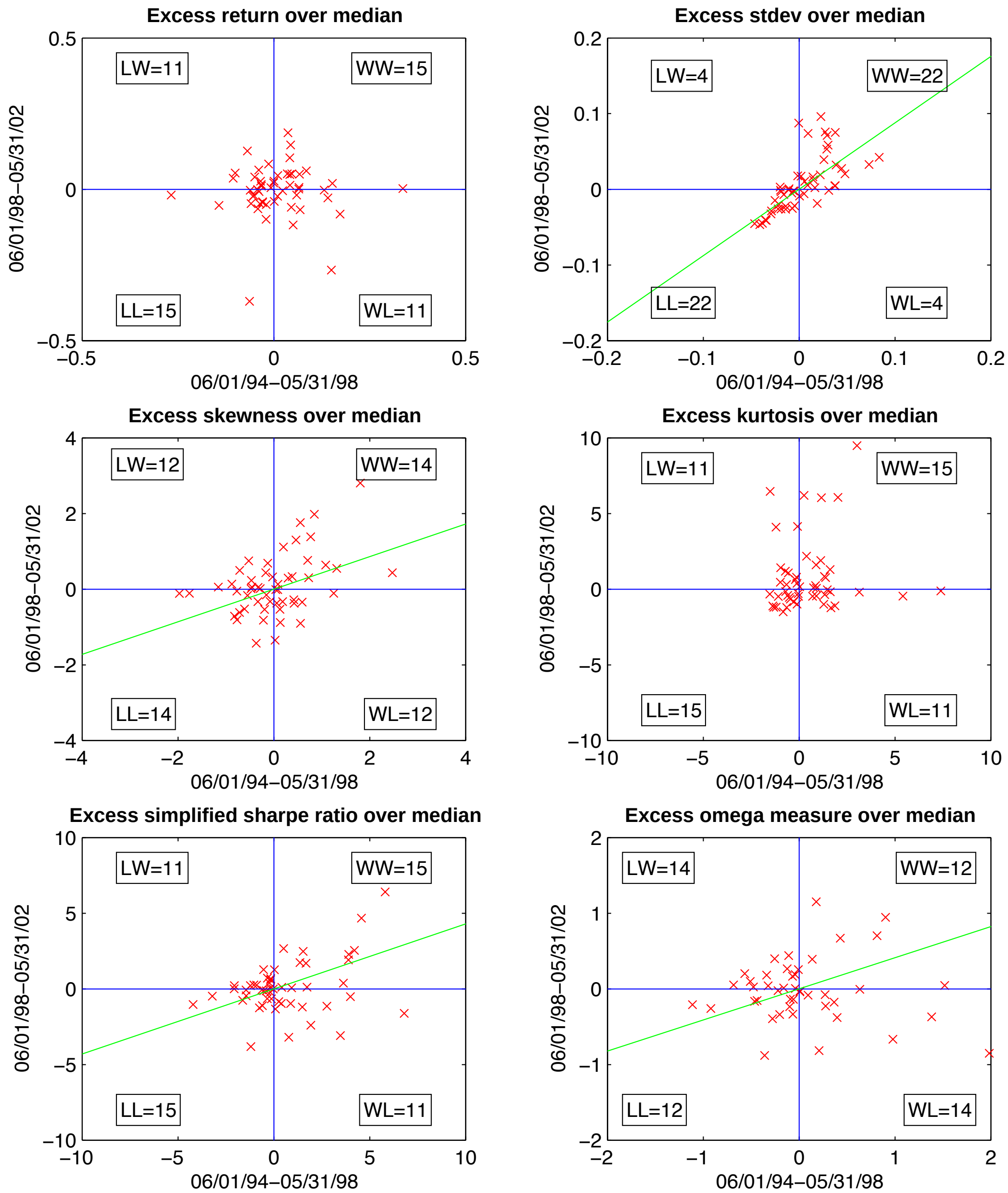

Figure 38: Persistence in non-trend-followers' return distributions parameters

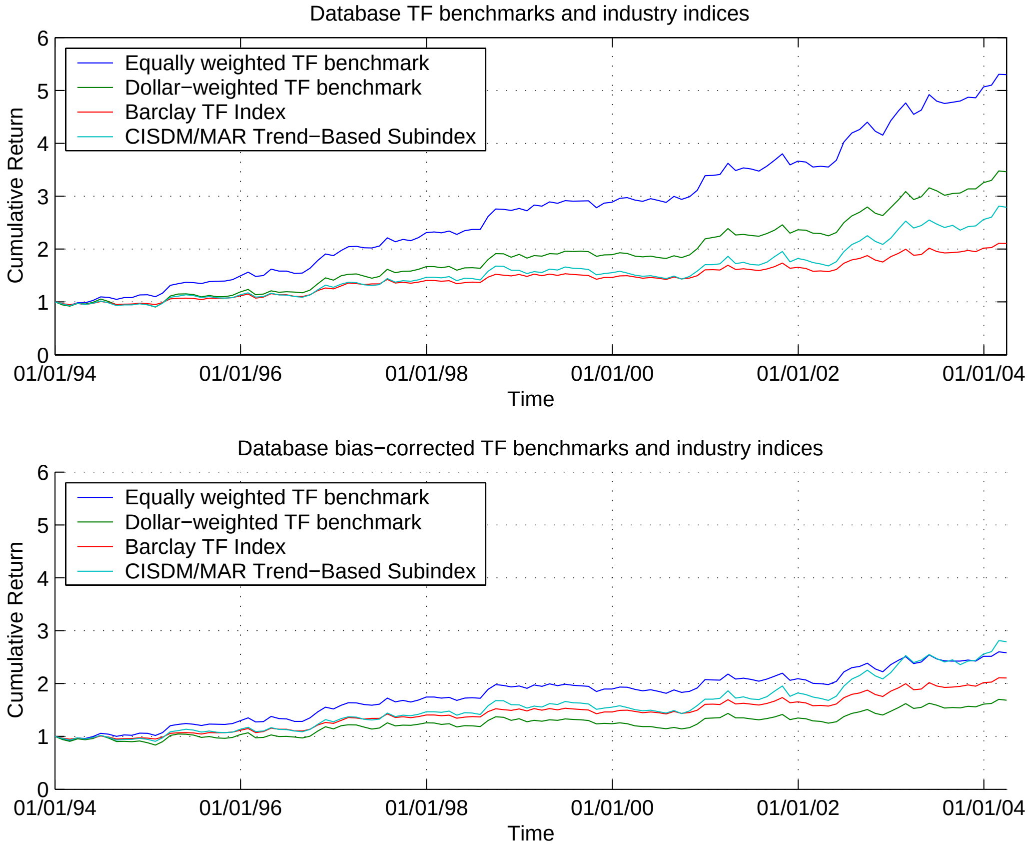

Figure 39: Biased and bias-corrected TFs benchmarks vs. industry indices

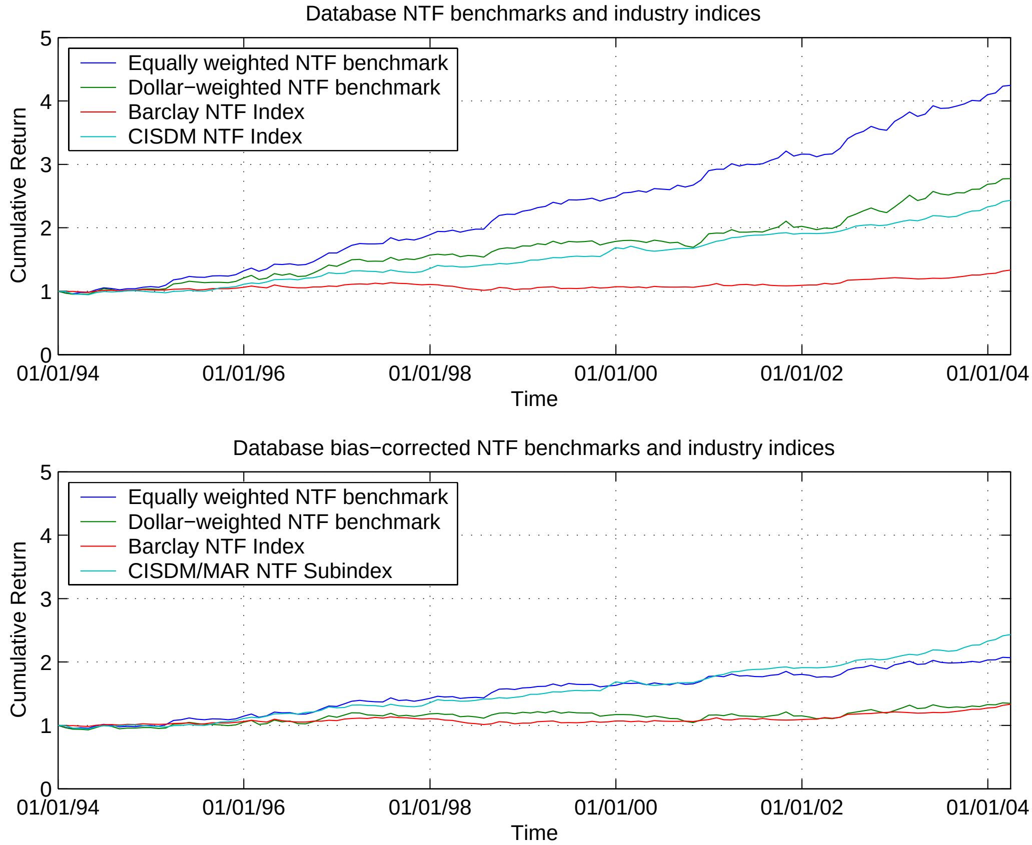

Figure 40: Biased and bias-corrected NTFs benchmarks vs. industry indices

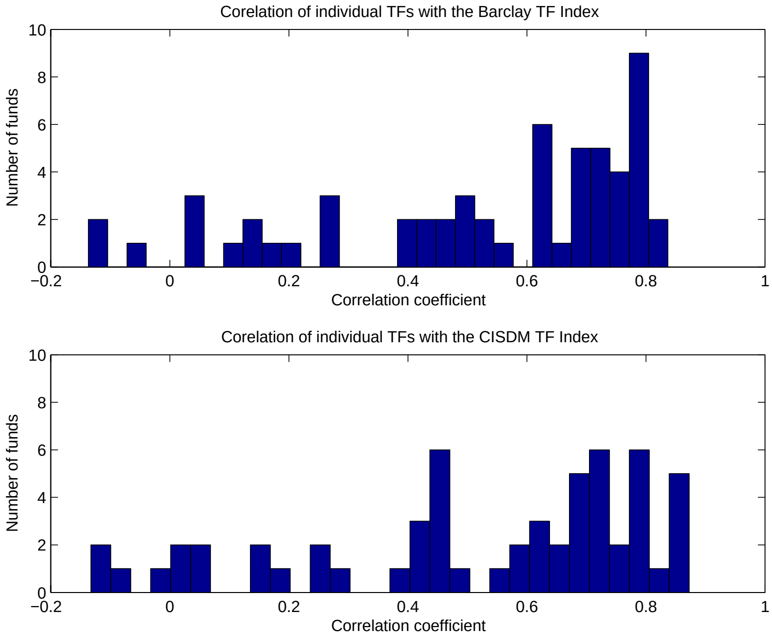

Figure 41: Histogram of individual TF correlations with benchmarks

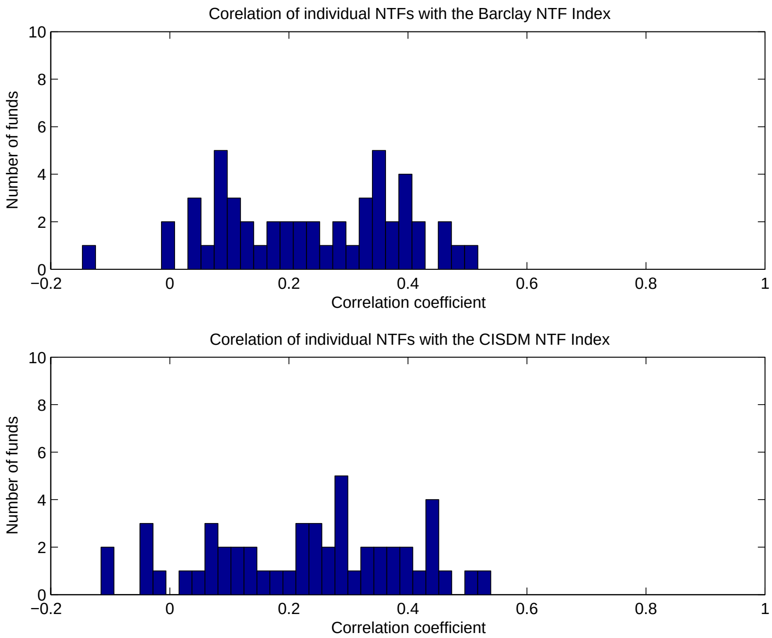

Figure 42: Histogram of individual NTF correlations with benchmarks

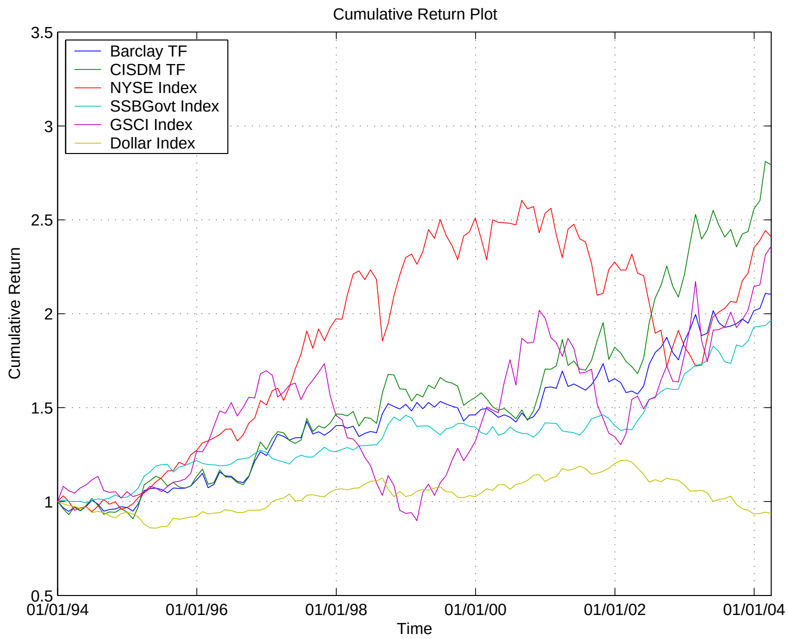

Figure 43: Evolution of asset classes over time

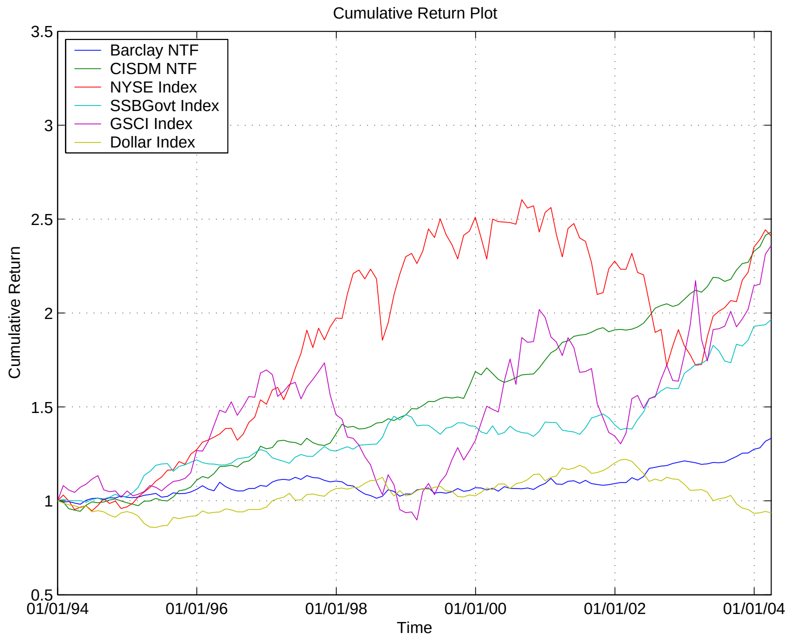

Figure 44: Evolution of asset classes over time

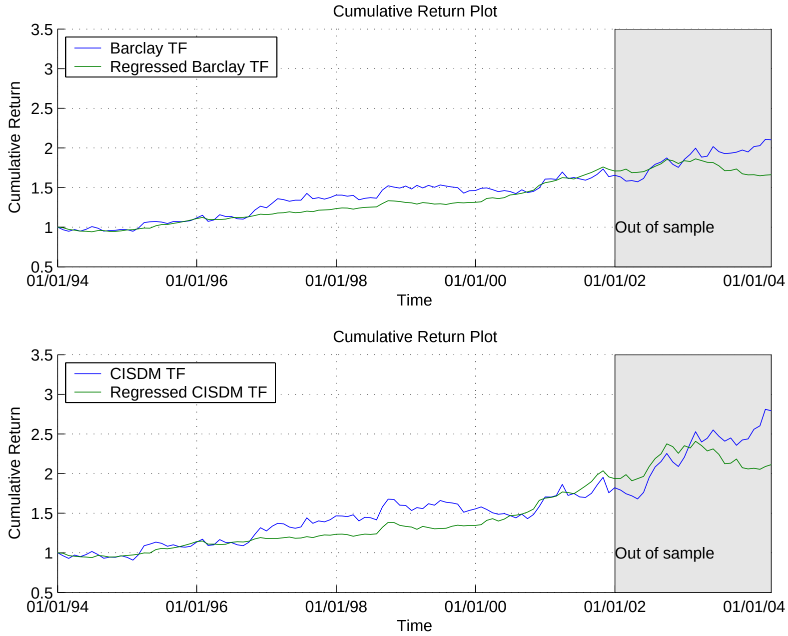

Figure 45: TF regression out of sample test

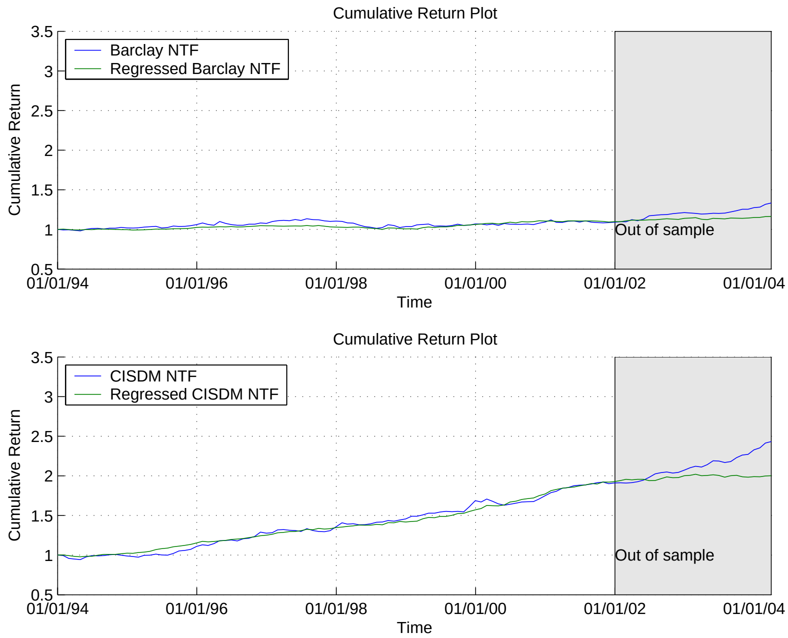

Figure 46: NTF regression out of sample test

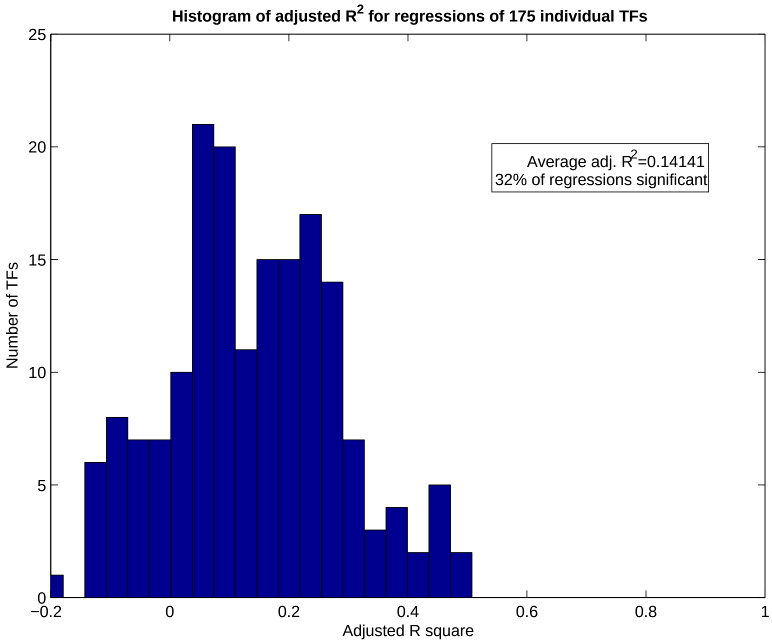

Figure 47: Histogram of adjusted R^2 from 4-year regressions of individual TFs

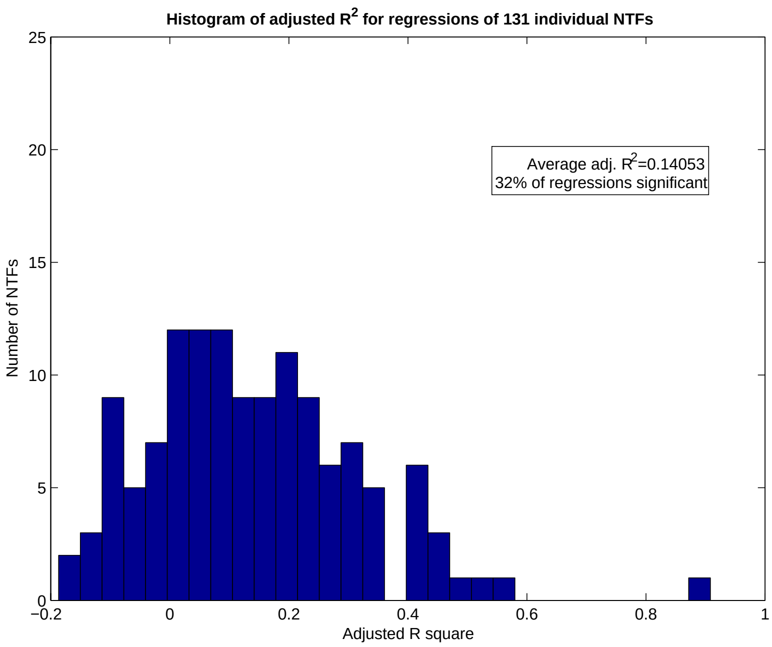

Figure 48: Histogram of adjusted R^2 from 4-year regressions of individual NTFs

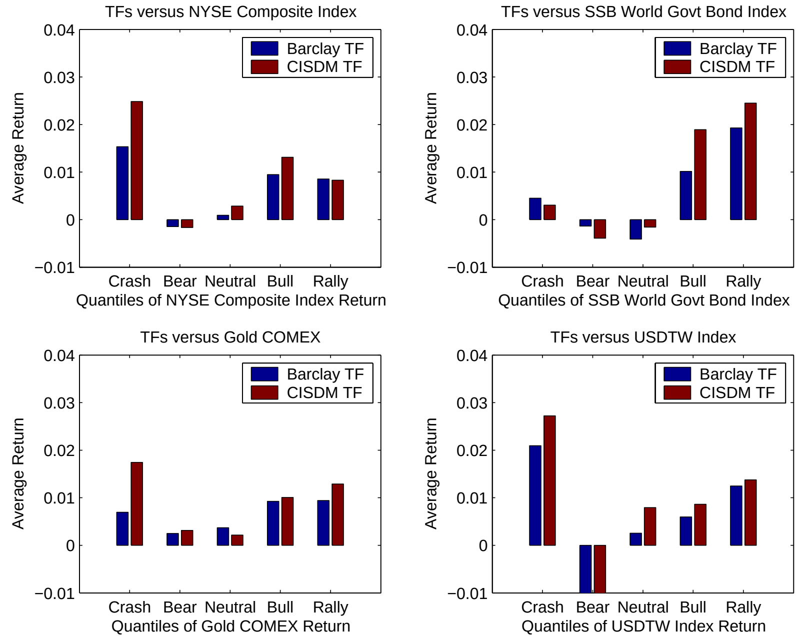

Figure 49: Quantile analysis for TFs

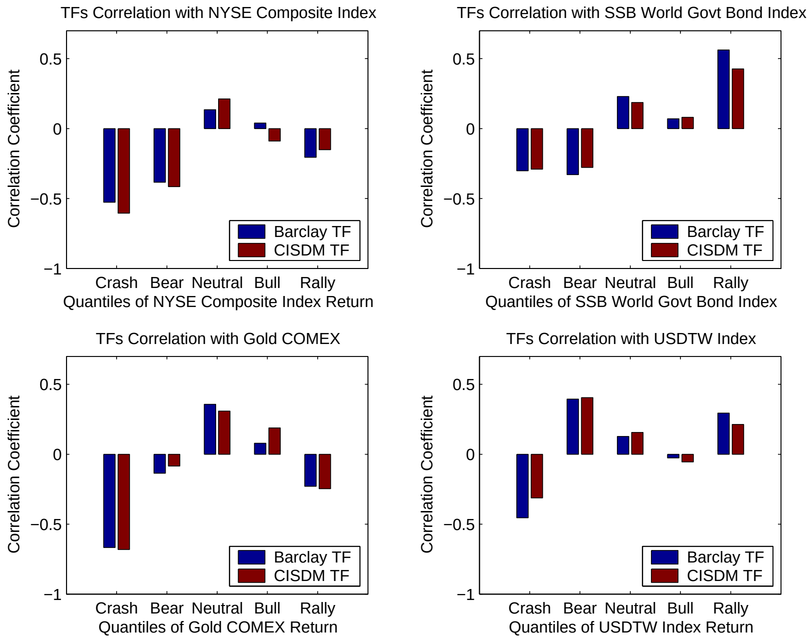

Figure 50: Conditional correlation analysis for TFs

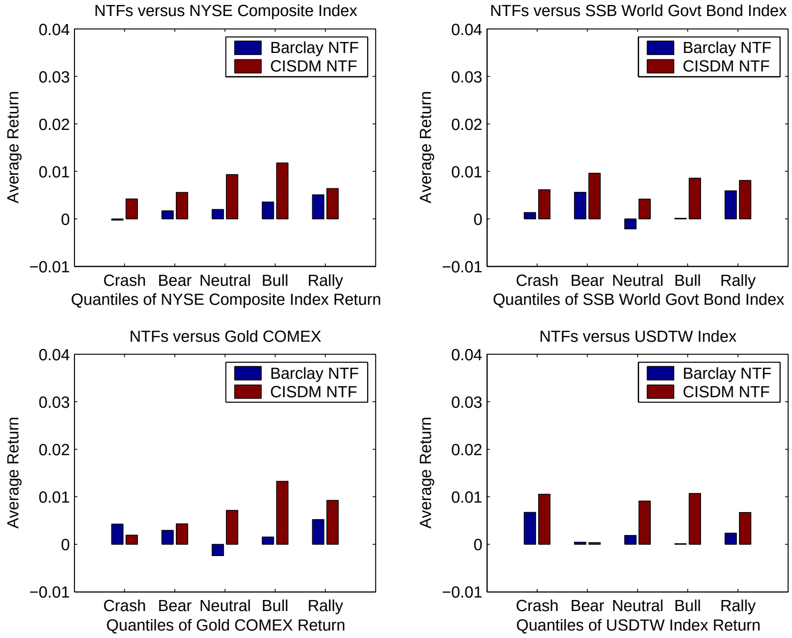

Figure 51: Quantile analysis for NTFs

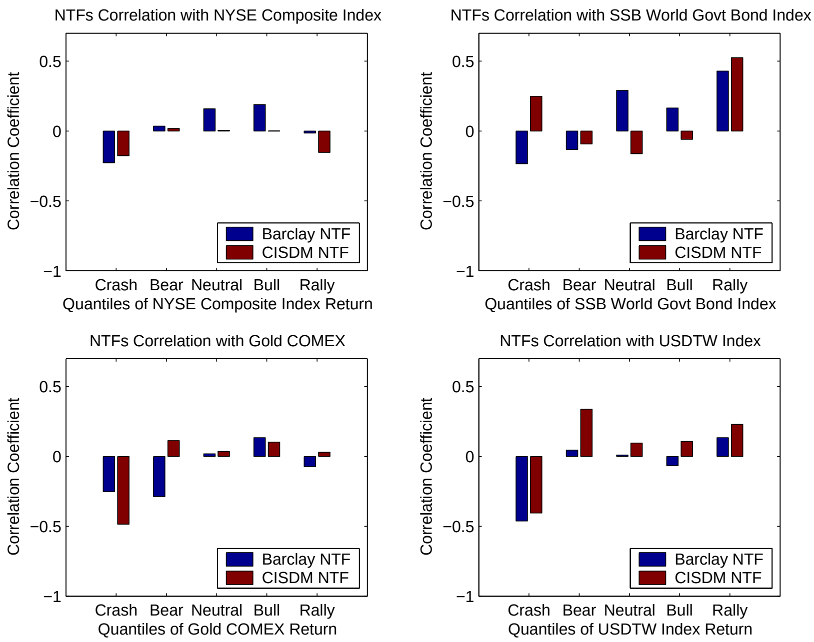

Figure 52: Conditional correlation analysis for NTFs

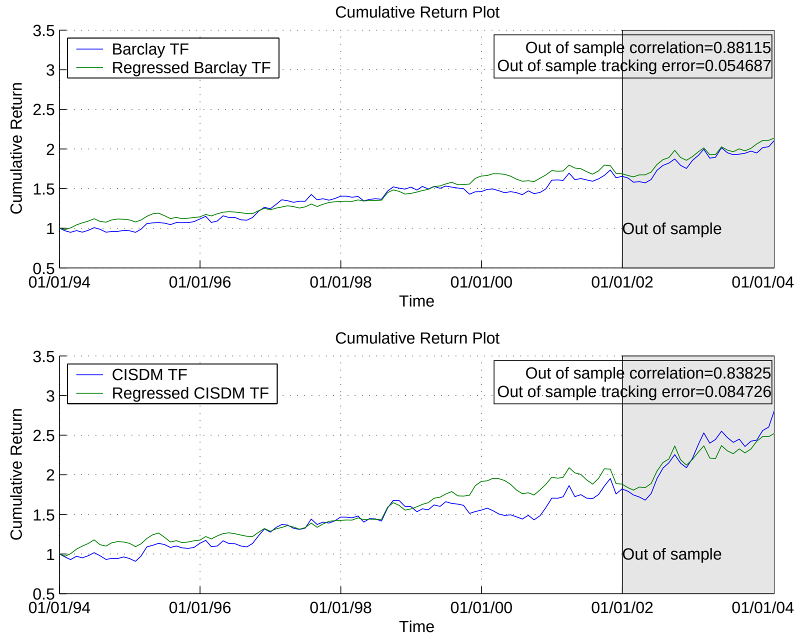

Figure 53: Out of sample performance of model for regression of trend-followers indices

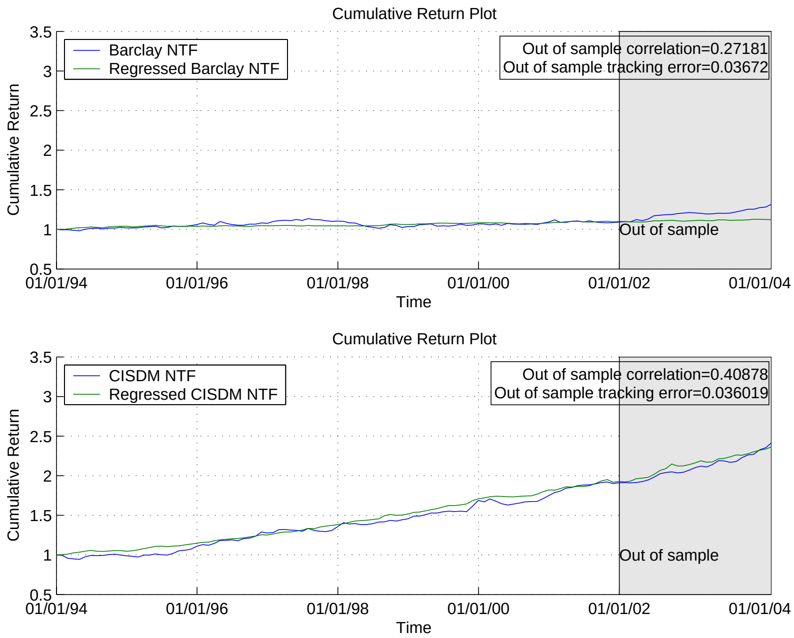

Figure 54: Out of sample performance of model for regression of non-trend-followers indices

C Tables

Table 14: Regression statistics for persistence of trend-followers' distribution parameters

	β	**Significant (5%)**	R^2
Avg. return	0.18	Yes	0.16
Stdev	0.72	Yes	0.58
Skewness	0.12	No	0.03
Kurtosis	0.05	No	0.10
$\frac{\mu}{\sigma}$	0.28	Yes	0.15
Ω	0.18	Yes	0.19

Table 15: Malkiel winner persistence test statistics for trend-followers' distribution parameters

	% Repeat winners	**Z-stat**	p**-value**
Avg. return	0.60	1.06	0.27
Stdev	0.83	3.53	0.00
Skewness	0.47	-0.37	0.71
Kurtosis	0.48	-0.19	0.85
$\frac{\mu}{\sigma}$	0.60	1.48	0.14
Ω	0.62	1.3	0.19

Table 16: CPR statistics for persistence of trend-followers' distribution parameters

	CPR	**Z-stat**	***p*-value**
Avg. return	2.12	1.42	0.16
Stdev	19.2	4.4	0.00
Skewness	0.81	-0.4	1
Kurtosis	0.81	-0.4	1
$\frac{\mu}{\sigma}$	3.3	2.18	0.03
Ω	2.46	1.68	0.09

Table 17: χ^2 statistics for persistence of trend-followers' distribution parameters

	χ^2	***p*-stat**
Avg. return	0.52	0.47
Stdev	5.64	0.02
Skewness	0.06	0.80
Kurtosis	0.06	0.80
$\frac{\mu}{\sigma}$	1.2	0.27
Ω	0.72	0.39

Table 18: Regression statistics for persistence of non-trend-followers' distribution parameters

	β	**Significant (5%)**	R^2
Avg. return	-0.02	No	0.00
Stdev	0.87	Yes	0.50
Skewness	0.43	Yes	0.20
Kurtosis	0.13	No	0.11
$\frac{\mu}{\sigma}$	0.43	Yes	0.30
Ω	0.41	Yes	0.71

Table 19: Malkiel test statistics for non-trend-followers' distribution parameters

	% Repeat winners	**Z-stat**	***p*-value**
Avg. return	0.58	0.78	0.43
Stdev	0.85	3.53	0.00
Skewness	0.54	0.40	0.70
Kurtosis	0.58	0.78	0.43
$\frac{\mu}{\sigma}$	0.58	0.78	0.43
Ω	0.46	-0.39	0.69

Table 20: CPR statistics for persistence of non-trend-followers' distribution parameters

	CPR	**Z-stat**	***p*-value**
Avg. return	1.86	1.1	0.27
Stdev	30.2	4.44	0.00
Skewness	1.36	0.55	0.58
Kurtosis	1.86	1.10	0.27
$\frac{\mu}{\sigma}$	1.86	1.10	0.27
Ω	0.73	-0.55	1

Table 21: χ^2 statistics for persistence of non-trend-followers' distribution parameters

	χ^2	***p*-value**
Avg. return	0.3	0.58
Stdev	6.23	0.01
Skewness	0.08	0.78
Kurtosis	0.30	0.58
$\frac{\mu}{\sigma}$	0.30	0.58
Ω	0.08	0.78

Table 22: TFs benchmarks autocorrelation

Lag	**TF-EW**	***p*-value**	**TF-DW**	***p*-value**
1	0.0232	0.800	0.0847	0.353
2	- 0.182	0.0456	- 0.176	0.0531
3	- 0.0354	0.701	- 0.0601	0.515
4	- 0.115	0.214	- 0.100	0.278
5	0.0153	0.870	- 0.0529	0.570
6	0.0261	0.780	- 0.0630	0.500
7	- 0.119	0.202	- 0.137	0.142
8	- 0.0107	0.910	- 0.00671	0.943

Table 23: NTFs benchmarks autocorrelation

Lag	**NTF-EW**	***p*-value**	**NTF-DW**	***p*-value**
1	- 0.00647	0.944	- 0.00868	0.924
2	- 0.118	0.197	- 0.132	0.149
3	- 0.0845	0.359	- 0.0955	0.300
4	- 0.0820	0.375	- 0.0905	0.328
5	0.0168	0.857	- 0.00635	0.946
6	- 0.0753	0.420	- 0.0361	0.699
7	- 0.0834	0.373	- 0.0918	0.327
8	0.0150	0.874	0.0336	0.722

Table 24: TFs benchmarks correlations

	TF-EW	**TF-DW**	**Barclay TF**	**CISDM TF**
TF-EW	1	0.953	0.963	0.935
TF-DW	0.953	1	0.934	0.957
Barclay TF	0.963	0.934	1	0.935
CISDM TF	0.935	0.957	0.935	1

Table 25: NTFs benchmarks correlations

	NTF-EW	**NTF-DW**	**Barclay NTF**	**CISDM NTF**
NTF-EW	1	0.936	0.548	0.582
NTF-DW	0.936	1	0.539	0.596
Barclay NTF	0.548	0.539	1	0.383
CISDM NTF	0.582	0.596	0.383	1

Table 26: TFs correlations with traditional asset classes

	Barclay TF	***p*-value**	**CISDM TF**	***p*-value**
NYSE	- 0.143	0.114	- 0.204	0.0235
MSCI EAFE	- 0.0928	0.307	- 0.147	0.105
MSCI EM	- 0.115	0.204	- 0.168	0.0633
SSBGovt	0.266	0.00297	0.262	0.00339
SSBCorp.	0.286	0.00135	0.247	0.00598
MLHY	- 0.126	0.166	- 0.201	0.0260
GSCI	0.211	0.0191	0.217	0.0160
Gold	0.140	0.122	0.102	0.262
Oil	0.0908	0.318	0.109	0.231
USDTW	- 0.0798	0.380	- 0.0791	0.385
1MD	0.00789	0.931	- 0.0450	0.621

Table 27: NTFs correlations with traditional asset classes

	Barclay NTF	***p*-value**	**CISDM NTF**	***p*-value**
NYSE	0.112	0.217	0.0629	0.489
MSCI EAFE	- 0.0532	0.559	- 0.0311	0.733
MSCI EM	0.197	0.0287	0.199	0.0276
SSBGovt	0.136	0.134	0.129	0.156
SSBCorp.	0.0891	0.327	0.304	0.000621
MLHY	- 0.0578	0.525	0.137	0.131
GSCI	0.335	0.000149	0.179	0.0473
Gold	0.171	0.0584	0.152	0.0941
Oil	0.283	0.00151	0.104	0.254
USDTW	- 0.159	0.0793	- 0.0263	0.773
1MD	- 0.163	0.0717	0.00107	0.991

Table 28: Regression of TFs indices' returns on traditional asset classes

	Barclay TF β	**p-value**	**CISDM TF β**	**p-value**
$\boldsymbol{\alpha}$	0.00101	0.886	0.0101	0.301
NYSE	- 0.117	0.222	- 0.181	0.174
MSCI EAFE	- 0.0117	0.842	- 0.0622	0.443
MSCI EM	0.0688	0.248	0.0784	0.340
SSBGovt	0.450	0.208	0.563	0.254
SSBCorp.	0.580	0.0575	0.844	0.0455
MLHY	- 0.349	0.0453	- 0.593	0.0142
GSCI	0.141	0.0478	0.189	0.0550
Gold	0.0421	0.604	- 0.0155	0.890
Oil	- 0.0463	0.284	- 0.0484	0.417
USDTW	0.359	0.274	0.436	0.336
1MD	0.0351	0.814	- 0.116	0.572
Adj. R^2	**0.150**		**0.178**	
p-value	**0.0018**		**0.00044**	
Sum of squared residuals	**0.0822**		**0.1567**	

Table 29: Regression of NTFs indices' returns on traditional asset classes

	Barclay NTF β	***p*-value**	**CISDM NTF β**	***p*-value**
α	0.00577	0.0639	0.00278	0.427
NYSE	0.0326	0.440	- 0.0835	0.0815
MSCI EAFE	0.00172	0.947	- 0.00972	0.739
MSCI EM	0.0405	0.122	0.0916	0.00233
SSBGovt	0.0284	0.856	0.112	0.526
SSBCorp.	0.144	0.281	0.295	0.0512
MLHY	- 0.189	0.0140	- 0.0549	0.523
GSCI	0.0489	0.118	0.0471	0.182
Gold	0.0173	0.628	0.0329	0.415
Oil	0.0112	0.554	- 0.0167	0.435
USDTW	0.0210	0.884	0.146	0.370
1MD	- 0.0959	0.145	0.0598	0.419
Adj. R^2	**0.132**		**0.118**	
***p*-value**	**0.0042**		**0.0079**	
Sum of squared residuals	**0.0158**		**0.0203**	