When it pays to follow the crowd: Strategy conformity and CTA performance

forthcoming, Journal of Futures Markets

Nicolas P.B. Bollen†
Frank K. Houston Professor of Finance
Vanderbilt University
Nashville, TN USA

Mark C. Hutchinson
Professor of Finance
University College Cork
Cork, Ireland

John O'Brien
College Lecturer
University College Cork
Cork, Ireland

February 8, 2021

Abstract

Prior research in hedge fund and mutual fund management finds a positive relation between portfolio distinctiveness and subsequent performance, suggesting that strategy differentiation is associated with superior skill. We find that CTAs with returns that correlate more strongly with those of peers feature higher performance and are more highly exposed to a time series momentum factor. Strategy conformity appears to be a signal of managerial skill in CTAs, in contrast to hedge funds and mutual funds. These results indicate that a common trend following strategy drives CTA returns and that CTAs offer investors an opportunity to invest in momentum.

---

†contact author: Nicolas Bollen, Vanderbilt University, 401 21st Avenue South, Nashville, TN 37203. email: nick.bollen@vanderbilt.edu phone: 615 397 6711 fax 615 343 7177. Mark Hutchinson email: m.hutchinson@ucc.ie John O'Brien email: j.obrien@ucc.ie. We are grateful to Juha Joenväärä, Olga Kolokolova, Linus Nilson, Adam Reed, Cristian Tui and the participants at the 12th Annual Paris Hedge Fund conference on January 13th and 14th 2020, and seminar participants at the Owen Graduate School of Management, Vanderbilt University for helpful comments. Mark Hutchinson would like to acknowledge the support of Bill Christie and the Owen Graduate School of Management, Vanderbilt University, for hosting him as a Visiting Scholar, while he completed the core empirical research for this paper. Mark Hutchinson and John O'Brien are grateful for financial support from the Science Foundation Ireland under Grant Number 18/SPP/3459.

When it pays to follow the crowd: Strategy conformity and CTA performance

Alternative assets constitute a core component of many institutional portfolios. According to NACUBO (2018), for example, U.S. university endowments with at least $1 billion in assets allocate on average 57% to alternatives, including hedge funds, real estate, and commodities, as of June 2017. Commodity Trading Advisors (CTAs), also referred to as Managed Futures, are the typical vehicle for gaining exposure to commodities. According to BarclayHedge, the aggregate assets managed by CTAs totaled $355 billion as of December 2018, reflecting 44% growth over the prior ten years.

Given the importance of CTAs in institutional portfolios, recent evidence of their generally poor performance is surprising. Bhardwaj et al. (2014) report that the average net-of-fee excess returns in CTAs from 1994 – 2012 was indistinguishable from zero, and that CTAs were not able to generate any incremental return above that of simple futures trading strategies. The poor performance documented by Bhardwaj et al. (2014) is a puzzle considering the properties of commodity futures returns. Moskowitz et al. (2012) find that futures markets feature strong serial correlation resulting in robust momentum in a wide array of contracts. The authors demonstrate that momentum effects are highly correlated across markets and that a diversified momentum trading strategy generates a Sharpe ratio more than double that of the equity market. Profitable momentum-based trading strategies in commodity futures markets are also found by Shen et al. (2007) and Narayan et al. (2014), among others. Given the robust evidence of profitability in commodity futures markets, it seems likely that at least some CTAs can offer investors a benefit.

This paper studies whether investors can choose a subset of available CTAs with a reliably higher than average likelihood of generating attractive future returns, motivated by the aforementioned momentum profits in futures markets. We draw from the literature on hedge fund performance prediction to select CTAs for investment. A widely accepted explanation for the success of some hedge fund managers attributes their abnormal returns to the ability to identify unique profit opportunities. Titman and Tiu (2010) and Sun et al. (2012) provide empirical support for this view, and develop two measures of a hedge fund's distinctiveness used to predict future performance.

Titman and Tiu (2010) measure the distinctiveness of a hedge fund by the ability of the Fung and Hsieh (2004) seven-factor model to capture the variation in the fund's returns. Hedge funds with low regression R^2 tend to outperform in the future, which can be interpreted as evidence that managers who hedge systematic risk in their pursuit of abnormal returns are skillful and that this skill persists. Sun et al. (2012) show that hedge funds with past returns which feature lower correlation with those of their peers also tend to outperform in the future. Consistent with economic theory, which suggests that only unique ideas would generate superior investment performance, Sun et al. (2012) demonstrate how strategy distinctiveness is a cross sectional determinant of hedge fund performance. Specifically, Sun et al. (2012) measure portfolio distinctiveness by one minus the correlation between a hedge fund's returns and those of other funds in its style grouping. When a fund has low correlation with its style grouping, it has a high Strategy Distinctiveness Index (hereafter *SDI*). They show that the top *SDI* funds outperform the bottom *SDI* funds by 3.5% in the following year. Similar results appear in the mutual fund literature. Kacperczyk et al. (2005) show that mutual funds which overweight particular industries tend to perform better. Similarly, Cremers and Petajisto (2009) find that mutual funds which deviate further from their indices perform better.

Titman and Tiu (2010) and Sun et al. (2012) show that hedge funds with differentiated trading strategies outperform. Price pressure from their actions and those of copycat traders eventually eliminates the potential for abnormal returns, requiring managers to continually search for new and different trades. We conjecture that for CTAs the opposite may be true. To understand why, note that Sun et al. (2012) identify two mechanisms which might weaken the positive association between strategy distinctiveness and hedge fund performance. First, as described by Goetzmann et al. (2003), unskilled managers may take excessive idiosyncratic bets in the hopes of achieving extreme levels of performance and compensation, given the typical hedge fund performance contract. Consistent with this notion, Bollen (2013) shows that hedge funds with high idiosyncratic risk (as measured by low factor model R^2) fail at a higher rate than other funds. Second, and more relevant for our study, Shleifer and Vishny (1997) argue that funds face limits to arbitrage when their investors are sensitive to short-term losses. Consequently, skilled managers may choose not to attempt to correct a mispricing in the market, but rather to profit from its continuation. Brunnermeier and Nagel (2004) provide a prominent example: many hedge fund managers loaded up on technology stocks as the internet bubble was inflating but were able to pare exposure before the crash. Since commodities feature robust time series momentum, as shown by Moskowitz et al. (2012), CTA managers must decide whether it is more profitable to follow the trend or pursue more distinct trading strategies. In this paper, we resolve this empirical question by measuring the relation between *SDI* and subsequent CTA performance.

To preview our results, we find that, in sharp contrast to hedge funds, CTAs with more differentiated strategies underperform those that conform, and explain this counterintuitive finding as a failure to pursue profitable momentum opportunities that characterize successful CTA management. Using return data on a large sample of CTAs from January 1994 to July 2015, we compute each CTA's *SDI* against its style every month using a 24-month rolling window. We control for survivorship, backfill bias and potential censoring of poor returns by failing CTAs. We follow Sun et al. (2012) in defining CTA investment styles using clustering.[1] The literature on mutual funds and hedge funds has demonstrated that clustering is generally superior to self-classified styles for characterizing cross-sectional past and future performance.[2] We document wide cross sectional variation in *SDI*, and find it is highly persistent, indicating that it is a viable performance predictor for CTAs.

Our empirical analysis contains three parts. First, we replicate the main tests in Sun et al. (2012) except now executed using CTAs instead of hedge funds. The most important result concerns the relationship between cross sectional differences in strategy distinctiveness and cross sectional differences in future CTA performance. We form portfolios of CTAs based on their levels of *SDI* and then examine the subsequent performance of the portfolios. We find that higher strategy distinctiveness is associated with lower future performance, exactly opposite to the finding for hedge funds in Sun et al. (2012). When we sort funds into portfolios based on *SDI* and hold for a year, CTAs in the low *SDI* quintile generate average annual returns an economically and statistically significant 5% higher than those of the high *SDI* quintile. Funds with less distinctive strategies perform consistently better even after controlling for risk and style differences.

Second, we determine the predominant trading strategy followed by CTAs which are highly correlated with their peers. We use the performance of a momentum-based futures contract trading rule, which relies entirely on analysis of past price movements, as a proxy for a standard trend following strategy of the type historically associated with CTAs. We first regress CTA

---

[1] Results are robust to an alternative measure of the *SDI*, based upon absolute correlation and BarclayHedge style classifications.
[2] See, for example, Brown and Goetzmann (1997) and Brown and Goetzmann (2003).

returns on the returns of our hypothetical momentum strategy. CTAs in the low *SDI* quintile feature an average regression coefficient on momentum returns of 0.67, significant 68% of the time, compared to an average coefficient of -0.01 for CTAs in the high *SDI* quintile. Thus it is clear that the conforming CTAs are jointly following a momentum strategy, whereas distinctive CTAs are not. We also show that the relationship between *SDI* and subsequent CTA performance is dependent on the profitability of momentum. When the returns to momentum are positive, there is an inverse relationship between *SDI* and CTA performance, i.e., the CTAs that conform outperform. When the returns to momentum are negative, for example when a rising market crashes or a falling market rallies, then there is a positive relationship between *SDI* and CTA performance. This result suggests that investors can use CTAs as a vehicle for investing in momentum.

Third, we assess whether investors can use *SDI* to select a realistic subset of CTAs that reliably offers a realizable incremental benefit. Following a simulation procedure developed in Bollen et al. (2019), we draw five CTAs at random from each *SDI* quintile once per year throughout our sample period and hold for one year. The process yields a time series of returns for a portfolio of five CTAs drawn from each quintile. We repeat 1,000 times to generate a distribution of outcomes. The portfolios selected from the low *SDI* quintile generate an average annual return of 6.30% versus just 0.66% for the portfolios selected from the high *SDI* quintile. Differences between the risk-adjusted performance of the portfolios drawn from the top and bottom quintiles are all highly significant. This result indicates that *SDI* conveys information precise enough for investors to feasibly realize significant incremental performance when selecting CTAs.

Our paper makes two contributions. First, we help resolve recent conflicting evidence about CTA performance. As mentioned, Bhardwaj et al. (2014) report that CTAs in aggregate generated an average return essentially equal to the risk-free rate over the 1994 – 2012 period. This result is consistent with the view of many practitioners, including Warren Buffett, that investors should avoid commodities as an asset class.[3] Trading on momentum in commodities, however, is found to be highly profitable by Moskowitz et al. (2012), among others. We show in a variety of analyses that CTAs which adhere to momentum-based strategies reliably outperform those that do not. Hence even if CTAs as a whole perform poorly, a subset that exploits momentum generally outperforms others.

Second, and more broadly, we contribute to the literature on strategy distinctiveness and the performance of actively managed portfolios. Abnormal returns measured by factor models are by definition the result of unspecified actions of a manager, motivating a set of markers for differentiated managerial activity, including the *SDI* of Sun et al. (2012) and R^2 of Titman and Tiu (2010). A natural tension exists. On the one hand, the ability of a manager in possession of profitable positions and strategies to exploit them is inversely related to the number of other market participants pursuing them as well. Hence we might expect that the more idiosyncratic a fund's returns are the better. On the other hand, when we know little about how a manager is generating returns, there is a chance an investor will be unknowingly exposed to unrewarded risk, leading in some cases to fund failure (Bollen (2013)). In the case of CTAs, the set of available securities is limited and well-defined. With no room to differentiate on security selection, it may be the case that successful CTA managers generally rely on the profitability of momentum noted above. In this context, then, following the crowd as opposed to pursuing idiosyncratic bets may be the optimal strategy.

---

[3] In the 2018 Berkshire Hathaway investor letter, Buffett compares the performance of gold to that of U.S. business over his investing lifetime: "the magical metal was no match for the American mettle." See also "Commodities in your portfolio? It's all hogwash says Wall Street dissenter," *Wall Street Journal* 14 April 2017.

I. Literature Review

The literature on identifying hedge fund managerial skill and the cross sectional determinants of hedge fund performance, including incentives (Agarwal et al. (2009)) and illiquidity (Aragon (2007)), is large and growing.[4] Researchers have also focused on the risk exposure of hedge funds to predict future returns, with high exposure to macroeconomic factors, higher systematic risk, and exposure to macro-economic uncertainty all being associated with higher future returns (Bali et al. (2011, 2012, 2014)).

Titman and Tiu (2010) show that hedge funds with relatively low factor exposures in the past tend to outperform on a risk-adjusted basis in the future. While a low factor model R^2 indicates that a manager is pursuing strategies different from those proxied for by the factor model, it does not explicitly assess whether funds with low R^2 are trading in similar or differentiated ways. In contrast, Sun et al. (2012) directly measure the correlation between a fund's returns and those of its peers and so can identify those managers who are pursuing differentiated strategies. Sun et al. show that strategy distinctiveness predicts higher future hedge fund returns. Hedge funds pursuing more unique strategies outperform those that follow less differentiated approaches.

Several studies focus exclusively on CTAs. Using principal components analysis, Fung and Hsieh (1997) show that there is one dominant investment style of CTAs, which they identify as trend following. Fung and Hsieh (2001, 2004) use the returns on option strategies to mimic the payoff from a successful trend following strategy. The lookback straddle used in their analysis generates the returns of a successful market timer, e.g., increasing exposure prior to price increases generates a long call payoff and decreasing exposure prior to declines generates a long put payoff. In our study, we focus instead on trend following in the form of a standard momentum strategy.

[4] See Bollen et al. (2020) for a comprehensive review.

Trend following will fail when there is a sudden shift in trend, as occurred for the banking sector when the U.S. Federal Reserve organized lending facilities in the depths of the 2008 – 2009 crisis, leading to large losses for momentum traders.[5] While the Fung and Hsieh factors have subsequently been adopted widely for modelling the returns of hedge funds, they were originally specified to model trend following CTA returns. In a sense, then, our paper is returning to the roots of contemporary research in hedge funds by studying CTAs and the relation between strategy distinctiveness and performance.

Naturally, the trend following strategy studied in Fung and Hsieh (2001, 2004) is likely to perform well when markets feature robust momentum. Trend following relies entirely upon past price movements to generate forecasts of expected futures prices: it involves no analysis of fundamental value, with all trading decisions based upon "trends" in asset prices, which are related to serial correlation in asset prices (Fung and Hsieh (2001)). Consistent with this view, Fung and Hsieh (1997) show that trend following returns are largest during the best and worst months for equity markets. More recently, Moskowitz et al. (2012) form time series momentum portfolios to capture serial correlation in futures markets by establishing long positions in assets with positive past 12-month returns and shorting assets with negative past-12 month returns. Moskowitz et al. document significant profitability in their hypothetical trading strategies.

Evidence suggests that the returns of CTAs are highly correlated with time series momentum, which can be considered a form of trend following (Baltas and Kosowski (2013)). Our paper contributes to this line of research by identifying a subset of CTAs that are trading most similarly and linking their commonality to momentum in futures markets. We use both a single momentum factor based on the 12-month lookback and 1-month holding periods in Moskowitz et

[5] See Daniel and Moskowitz (2016) for an analysis of momentum crashes.

al. (2012) as well as a three-factor model from Baltas and Kosowski (2013) that includes periods measured in days, weeks, and months to assess the specific momentum strategies CTAs employ.

II. Methodology

II.A. SDI

A central aim of this study is to test whether the relationship between strategy distinctiveness and subsequent hedge fund performance documented by Sun et al. (2012) holds for CTAs as well. Consequently, some of our analysis replicates the main elements of their methodology. The strategy distinctiveness indicator (*SDI*) of Sun et al. (2012) is defined as:

$$SDI_{i,t} = 1 - corr_t(r_i, r_c) \tag{1}$$

where $SDI_{i,t}$ is the strategy distinctiveness measure of fund i at time t, r_i is the return of fund i, and r_c is the return of the cluster of which fund i is a member. The correlation at time t is calculated over the prior 24 months, with a minimum of 12 months of returns. All available funds over this window are partitioned into clusters such that all funds within a cluster have the highest group correlation. The purpose is to define style categories in an objective fashion without relying on self-reported styles.[6] A fund with high *SDI* features returns that have low correlation with its peers, presumably because the manager is pursuing unique trading opportunities, and hence Sun et al. (2012) interpret it as a measure of hedge fund manager skill.

II.B. Clustering to determine styles and identify peers

Measuring the *SDI* requires that each fund be assigned to a style category. BarclayHedge classifies CTAs into eight main styles, with the largest styles being split into further sub-styles, giving 19 styles in total. While it is possible to use these as the basis of the *SDI* calculation, there

[6] We test the robustness of our results to alternative ways of grouping funds as described later in the paper.

are a number of issues with this approach, most notably the reliability of self-assigned styles and the potentially dynamic nature of the strategies pursued by a given manager. The literature on mutual funds and hedge funds demonstrates that statistical clustering is generally superior to self-classified styles for predicting cross-sectional past and future performance (Brown and Goetzmann (1997) and Brown and Goetzmann (2003)). Self-classification provides great latitude for funds to conduct widely divergent behavior when operating in the same classification. Even highly regulated mutual funds are prone to misclassification when relying on self-reported styles (Brown and Goetzmann (1997), Chen et al. (2019)). We follow Sun et al. (2012) and use statistical methods to define our clusters to objectively place funds into groups as a function of their realized return dynamics.

We use an iterative process, following Brown and Goetzmann (2003), to generate our clusters. The aim of the clustering is to find an optimized grouping of funds that minimizes the distance of all funds to their cluster. We measure the time series of each cluster's returns as the average cross sectional return of its constituents. Each quarter the distance between each fund's return and each cluster's return is calculated, using a minimum of twelve months of returns over the preceding twenty-four months. Using these estimates, funds are reassigned to the cluster with which they have the highest correlation. The return series of the new clusters are updated and the process is repeated.

We specify eight clusters to match the number of main styles in BarclayHedge. Figure 1 plots the histogram of the *SDI*s of the CTAs in our sample based on the statistical clusters alongside the histogram of the *SDI*s based on the self-reported style classifications from BarclayHedge. The mean for the cluster style-based *SDI* is 0.47 versus 0.58 for the *SDI* based on BarclayHedge styles, indicating that statistical clustering results in higher correlations across CTAs in the same style as compared to relying on styles reported to the database. A Kolmogorov-Smirnoff test applied to the two cumulative distributions of *SDI* yields a test statistic ten times larger than the critical value at a 1% significance level, hence we can easily reject the hypothesis that the two distributions are equivalent. Another way to compare the two is to measure the percentage of CTAs with *SDI* greater than one, i.e., with a negative correlation to its style. Roughly 10% of funds have a BarclayHedge style *SDI* greater than 1, compared to just 2.5% when using style clusters. These results illustrate the ability of the clustering methodology to more effectively assign CTAs to peer groups pursuing similar strategies.

II.C. Performance Measures

We test the relationship between *SDI* and four performance measures: abnormal return, appraisal ratio, Sharpe ratio and the manipulation proof performance measure.

The abnormal return (alpha) is measured as the return in excess of the expected return generated by exposure to risk factors. We use the Fung and Hsieh (2004) seven factor risk model to measure a CTA's risk exposures and alpha. The appraisal ratio (AR), developed by Treynor and Black (1973), is defined as the estimated abnormal return of the fund scaled by its standard error. The Sharpe ratio is measured as a CTA's average return in excess of the risk-free rate divided by the standard deviation of its returns.

Goetzmann et al. (2007) identify a range of ways managers can game standard measures to appear skillful. Their solution is the manipulation proof performance measure (MPPM):

$$MPPM = \frac{1}{(1-\rho)\Delta t} ln\left(\frac{1}{T}\sum_{t=1}^{T}((1+r_t)/(1+rf_t))^{1-\rho}\right) \qquad (2)$$

where T is the number of observations, Δt is the time between observations, r_t and rf_t are the CTA return and risk free rate respectively, and ρ is a risk aversion parameter. The MPPM can be interpreted as the minimum incremental return that a risk-averse investor would accept to exchange a risky investment, e.g., a portfolio of CTAs, for the risk-free asset. In other words it is the difference between the risk-free return and the certainty-equivalent of the risk investment. Goetzmann et al. (2007) argue ρ typically varies between two and four. We follow Sun et al. (2012) and use a value of three. The statistical power of the manipulation proof performance measure is estimated using a bootstrap method.[7] This allows us to determine the *p*-value, and hence the statistical significance of differences in MPPM values between portfolios.

III. Data

Our paper uses two main datasets: a sample of futures market data to construct a momentum factor and a sample of CTAs for analyzing strategy distinctiveness and fund performance.

III.A. Futures Data

Table 1 lists the futures contracts used in the study: 19 commodity futures, 12 equity index futures, eight 10-year bond futures, and nine foreign exchange futures. We require data from January 1993 to form our momentum portfolios.

We require continuous cumulative excess returns of positions in the futures contracts to construct momentum factors. Two methods are used to create these futures contract return series. The first takes the price series for individual futures contracts trading on an exchange and combines these to produce a continuous excess return series. The second approach creates a synthetic return series by combining the underlying spot price, yield and risk free rate. All futures and forward contract data are from Thomson Reuters.

---

[7] See Sun et al. (2012) for a description of the bootstrapping methodology, which involves simulating the distribution of MPPMs under the null hypothesis of no statistically significant relationship between *SDI* and fund performance.

The continuous return series created from futures is derived from the monthly price series. For each month, the return for that month is the return of the nearest to deliver contract, which trades for the full month. In effect, this means we roll contracts on the last day of the month, prior to the delivery month.

In the case of synthetic forwards, the excess return is defined as a function of spot price, yield and risk free rate. The excess return from a long position in a forward contract i in month t, er_{it} is given by:

$$er_{it} = (1 + r_{it})\left(\frac{1+q_{it}}{1+r_{ft}}\right)^{(1/12)} - 1 \tag{3}$$

where r_{it} is the spot price return for asset i in month t, r_{ft} is the one-month risk free rate, and q_{it} is the annualized yield on underlying asset i at month t.

As listed in Table 1, the vast majority of futures contract return series have positive monthly serial correlation, and the correlation coefficient is significantly different from zero in many cases.

III.B. CTA Data

We use the BarclayHedge CTA database as our source of returns data due to its depth of coverage. The raw data include 5,199 funds over the 1994 – 2015 period.

To ensure the data is representative of investors' experience of investing in CTAs a number of filters are applied. We first remove funds of funds and any indices, leaving 4,971 funds. Next, we remove all non-US dollar denominated funds, funds that do not report monthly, and funds that only report returns gross of fees. Our next step is to identify and remove duplicate funds. Following Jorion and Schwarz (2014), we use a return based filter to identify duplicate share classes. If two funds from the same management company have a correlation of 0.99 or higher, then the fund with the latest start date is removed.

CTA returns are voluntarily reported to the BarclayHedge database and so are subject to the biases studied extensively in the hedge fund literature. We address three explicitly. First, to address survivorship bias, the BarclayHedge database includes both live funds and a graveyard file, which contains the returns of funds that have ceased reporting. Second, as studied by Aiken et al. (2013), funds that cease reporting due to poor performance often continue to exist, hence the database suffers from a censorship bias. We follow Liang et al. (2010) and identify a CTA failure when a fund has a negative six-month average return and a negative 12-month change in AUM prior to the cessation of reporting. In unreported analysis we reverse the censorship bias by appending to a failing fund's return history one ad-hoc additional loss of 50%, as in Titman and Tiu (2002). These results are stronger, with a larger performance difference between low *SDI* and high *SDI* funds.

Third, we address backfill bias.[8] Since 2002, BarclayHedge have reported a "date added" variable for each fund, indicating when the fund was added to the database. This easily allows us to remove all backfilled returns for these funds, prior to the date they were added to the database. For pre-2002 data, BarclayHedge provided us with the constituents of the BarclayHedge CTA Index. To qualify for inclusion in the index an advisor must be reporting to BarclayHedge at the beginning of the year. Additional programs introduced by qualified advisors are not added to the Index until after their second year. As constituents are added at the beginning of each year, in order to qualify for inclusion in the index, a fund would have to be already reporting to BarclayHedge.

[8] Recent academic attention has focused on how to address this bias. For example, Bhardwaj et al. (2014) and Getmansky et al. (2015) remove any observations prior to the "date added" field in the TASS database. However, subsequent evidence by Jorion and Schwarz (2014) indicates that this approach is likely to overstate backfill bias, as many funds in TASS in fact report to HFR at an earlier date. Using TASS is further complicated by the merger of TASS with Tremont in the late 1990s. The "date added" field in the TASS database for the Tremont funds is not the date the funds were added to Tremont but the date they were merged with TASS, which is much later (Fung and Hsieh (2009)). Whereas for HFR the earliest add date is May 1996, which is most likely to represent the date HFR started to collate this variable (Jorion and Schwarz (2014)).

Hence, pre-2002 we only include a fund's returns in our sample if they have, or have in the past, been constituents of the BarclayHedge CTA Index.[9] This leaves us with a sample of 3,461 funds. Removing funds which do not report AUM data leaves us with 3,419 funds. Estimation of clusters and strategy distinctiveness requires a minimum of twelve months of returns. Removing funds with fewer than twelve months eliminates a further 950 funds. Finally, we filter out fund months with assets under management (AUM) of less than 5 million dollars, resulting in a final sample with 966 unique CTAs.

Figure 2 shows the number of CTAs and aggregate assets under management each year across the sample period. There is a notable increase in funds from 2002 onward when BarclayHedge introduce the date added field.

Table 2 lists summary statistics of the returns of CTAs in our sample. All statistics except *SDI* are measured once for each CTA over its life. The cross-sectional average of these, as well as the three relevant percentiles, are listed. Panel A shows results for the full sample. The average annual return is a modest 3.14% per year with an excess return above that of the 3-month T-bill of just 1.33%. The latter is similar to the 1.81% average excess return in Bhardwaj et al. (2014). Note, however, the wide cross-sectional variation, with the 20th and 80th percentiles feature average returns of -1.78% and 7.78%, respectively, motivating our goal of predicting CTA performance. The average factor model R^2 is 16%, indicating that for many funds the Fung and Hsieh (2004) seven-factor model is not able to capture much time series variation in returns. Note though that the 80th percentile is 32%, so that for a subset of funds the trend following factors developed by

[9] There are 19 funds that are listed in the BarclayHedge CTA Index constituent list which are not in the database. BarclayHedge clarified that there are a number of funds who provide them with data solely for index calculation purposes, with whom they have an agreement not to redistribute their data. A number of funds cease reporting to the database for a period before resuming reporting, leaving a gap in their return history. In this instance we remove all returns prior to the fund resuming reporting.

Fung and Hsieh (2004) for the express purpose of describing trend following factors work as advertised.

The focus of our paper is *SDI*, the last statistic listed. We measure *SDI* for each CTA monthly using a 24-month rolling window and report the average and three percentiles of the pooled set of *SDI* observations. The average *SDI* in the full CTA sample is 0.42, reflecting a substantial degree of commonality in CTA returns. The bottom 20% has *SDI* of 0.17 or less, meaning for these CTAs the correlation between their returns and their peers is over 80%, whereas the top 20% feature an *SDI* of at least 0.66. The purpose of our study is to investigate the implication of this large cross-sectional variation.

Panels B and C show results for the live and dead sub-samples, respectively. As one might expect, the live CTAs generate far superior returns than dead CTAs, since underperformers are more likely to liquidate. The average excess returns of the two groups are 4.91% and -0.10%, respectively, for example. Pronounced differences in risk-adjusted returns are also evident. Live CTAs feature average Sharpe ratios of 0.49, for example, versus -0.21 for dead CTAs. Live CTAs are on average much larger, $795 million versus $158 million, as better performing funds attract more capital. Importantly, the live CTAs feature lower *SDI* than dead CTAs, and the 0.12 difference is statistically significant. This is preliminary evidence that there is an inverse relation between *SDI* and performance in the CTA space.

Table 3 presents more direct evidence of an inverse relationship between *SDI* and CTA performance. CTAs are sorted into quintiles based on the *SDI* computed on the last month of their observation history. Summary statistics of their returns are computed over their lives, and then averaged across the CTAs in each quintile. There is wide variation in the average *SDI* across the quintiles with the low *SDI* quintile (Q1) featuring an average *SDI* of 0.12 versus 0.89 for the high

SDI quintile (Q5). Low *SDI* funds have histories 1.57 years longer on average relative to high *SDI* funds. This result suggests that more distinct CTAs are more likely to fail, consistent with the findings in Bollen (2013) for hedge funds. More important, the inverse relation between *SDI* and performance is strong and economically meaningful. The average return of the low *SDI* quintile is 5.08% per year, for example, versus just 1.40% for the high *SDI* quintile. Differences in risk-adjusted performance are also large and significant; the alpha of the low *SDI* quintile is 3.23% per year versus 0.00% for the high *SDI* quintile. The differences in risk-adjusted performance is remarkable given that low *SDI* CTAs have annual volatility roughly double that of high SDI CTAs. This result suggests that the common trading strategy of low *SDI* CTAs is itself quite risky.

The CTAs in the low *SDI* quintile feature the highest R^2 from the Fung and Hsieh (2004) factor model, suggesting that the similarity in CTA returns in the low *SDI* quintile is at least partly attributable to a common pursuit of trend following strategies. We pursue this conjecture more thoroughly in the Section V.

IV. Results

This section presents our analysis of the persistence of *SDI*, which is necessary for *SDI* to be a reliable predictor of performance, and of the relationship between *SDI* and subsequent CTA performance. To preview, we find strong persistence in *SDI* and an inverse relationship between *SDI* and subsequent CTA performance, indicating that in this asset class a predominant common strategy drives success.

IV.A. Persistence of SDI

If *SDI* measures skill in CTA management then it should persist since skill by definition is the consistent ability to generate superior returns. Following the test for persistence in Sun et al.

(2012), we measure the persistence of *SDI* by dividing CTAs into five portfolios each month based on their *SDI* and then measuring the average *SDI* of each CTA at various dates in the future ranging from three months up to three years. Table 4 lists the average *SDI* for CTAs in each quintile, both at the ranking date (time 0) as well as the subsequent future dates. The results indicate a high level of persistence, with the spread between the average *SDI* in the low and high quintiles narrowing only slightly as the horizon extends. At a three month horizon, *SDI*s are 0.13 for CTAs in the low *SDI* quintile and 0.80 for those in the high *SDI* quintile. These correspond to average correlations between CTAs of 0.87 and 0.20, respectively. The difference persists at all future horizons. Two years post ranking, for example, the average *SDI*s are 0.16 for CTAs in the low *SDI* quintile and 0.72 for those in the high *SDI* quintile. At all future dates the difference between the two extreme quintiles is large and statistically significant.

Persistence of *SDI* in CTAs suggests it may be a reliable performance predictor; thus, we turn next to a test of the relation between *SDI* and subsequent CTA performance.

IV.B. SDI and CTA performance

In order to understand the link between strategy distinctiveness and performance we implement a series of trading strategies that selects CTAs based on their *SDI*. Here again we follow the procedure in Sun et al. (2012). Each quarter, we sort CTAs into quintiles according to their prior 24-month *SDI*.[10] The equally-weighted quintile portfolios are then held for periods ranging from three months to three years.[11] For holding periods greater than three months, we form overlapping portfolios from the relevant prior quarterly sorts.

---

[10] In unreported robustness tests we repeat the main analysis of CTA distinctiveness and performance using the BarclayHedge style classifications to define peers rather than the clustering procedure. Results are qualitatively identical.

[11] We consider overlapping portfolios to minimize *SDI* estimation error. However, we acknowledge that our approach could not be replicated by an investor in hedge funds, due to frequent rebalancing. To ensure our results are robust, we also create portfolios in January of each year, based on prior month *SDI*, and hold for one year. These results, which are unreported, are very similar, with less distinctive funds outperforming.

Table 5 lists the average annualized return of each quintile over the different holding periods. In each case, the return is monotonically decreasing across the *SDI* quintiles. At the one-year horizon, for example, CTAs in the low *SDI* quintile (Q1) feature an average return of 7.05% versus just 1.90% for those in the high *SDI* quintile (Q5). The difference between the average returns of Q1 and Q5 is between 3% and 5% at each horizon. The differences are statistically significant for all except the quarterly holding period. Clearly the more that CTAs conform to others in their style cluster, and the lower is their *SDI*, the higher are their subsequent returns on average.

As shown in Table 6, risk-adjusted performance measures also decrease monotonically with past *SDI*. The difference between the performance of the low and high *SDI* quintiles is statistically significant in almost all cases. The CTAs in the low *SDI* quintile conform the most to their peers, possessing the least distinctive strategies, and yet they perform the best. This robust result stands in complete contrast to the findings for hedge funds in Sun et al. (2012). For a trading strategy with a one-year holding period, for example, funds in the low *SDI* decile, which is the least distinctive category, earn an annualized alpha of 4.59%, whereas those in the high *SDI* quintile, which are most unique, earn an annualized alpha of -0.43%. The performance difference is notable, showing an economically significant difference between the abnormal returns of CTAs in the low and high *SDI* quintiles of 5%.

To take into account idiosyncratic risk, we also present results for the appraisal ratio. As with alpha, the appraisal ratio decreases with *SDI*. For a one-year holding period, the average appraisal ratio of CTAs in the low *SDI* quintile exceeds that of the CTAs in the high *SDI* quintile by a statistically significant 0.46. As with the other performance metrics, the average Sharpe ratios decrease across the *SDI* quintiles. Furthermore, the differences in average Sharpe ratios between the low and high *SDI* quintiles are much larger than those reported in Sun et al. (2012) for hedge funds, indicating that conformity in CTA strategies is a stronger indicator of subsequent CTA performance than distinctiveness in hedge fund strategies is for subsequent hedge fund performance.

We report results for manipulation proof performance measures in Table 7. We use a bootstrap approach to test for statistical significance of MPPMs following Titman and Tiu (2010) and Sun et al. (2012). The results for MPPM quintiles confirm the earlier results. MPPMs decrease with *SDI*. The difference between the MPPM of the low and high *SDI* portfolios are statistically significant for all holding periods.

Taken together the results in this subsection are striking. Unlike hedge funds, for which strategy uniqueness is positively related to future fund performance, we find exactly the opposite for a group of sophisticated CTAs operating exclusively in futures markets. Economic theory suggests that unique investment strategies are important for performance. In a competitive industry, if enough managers are pursuing a similar strategy, then they should compete away the returns to the strategy. As explained first by Shleifer and Vishny (1997), however, in some investment contexts managers are limited in their ability to trade in this fashion, and may choose to invest with the crowd, as in Brunnermeier and Nagel (2004). The results here indicate that the futures markets in which CTAs operate conform to this characterization.

We turn next to an exploration of the economic underpinning of the inverse relation between *SDI* and subsequent CTA performance. The analysis is motivated by the widely held view that CTA managers are predominantly trend followers, presumably because the most reliable, and perhaps only, exploitable time series property of futures markets is momentum.

V. Momentum and *SDI*

As described by Fung and Hsieh (2001), the predominant trading strategy in CTAs is trend following. Since the success of a trend following strategy is naturally associated with markets featuring momentum, we explore next the role of momentum in the relationship between *SDI* and subsequent CTA performance.

V.A The role of momentum in CTA performance

Moskowitz et al. (2012) report that time series momentum, defined as the tendency of prices of individual securities to continue to rise or fall over successive periods, is a common phenomenon in a wide array of futures contracts. A time series momentum portfolio is formed entirely based upon univariate analysis of past price movements: going long assets whose past returns are positive and shorting assets with negative returns. Moskowitz et al. (2012) naturally find that the returns of following a time series momentum strategy are high when asset returns possess persistent serial correlation.

We follow Moskowitz et al. (2012) and form a time series momentum portfolio for each futures contract using a 12-month look-back window and a one-month holding period. In Subsection V.D we vary these parameters to determine whether CTAs as a group focus on momentum at other horizons. We then combine returns of each contract type using equal volatility weighting into an aggregate time series momentum portfolio. The time series momentum portfolio, *TSMOM*, has positive average monthly returns and an extremely high annualized Sharpe ratio of 0.71. The returns of *TSMOM* must be interpreted with caution since they do not include transaction costs and so constitute an unobtainable series in practice. For comparison, the annual Sharpe ratio of the 12-month holding period Q1 *SDI* portfolio is 0.23 as listed in Table 6. Transaction costs and other real-world frictions help explain why this realized performance is inferior to that of *TSMOM*.

We expect low *SDI* funds to perform better relative to high *SDI* funds when momentum returns are higher, since low *SDI* funds are likely following a common trend following strategy. To support this conjecture, we regress CTA returns on the *TSMOM* factor. Table 8 shows the average coefficient on *TSMOM* for CTAs split into quintiles by their *SDI*, as well as the percentage of CTAs within each quintile for which the regression coefficient is statistically significant at the 10% level. CTAs in the low *SDI* quintile have an average coefficient of 0.67 versus -0.01 for CTAs in the high *SDI* quintile. Furthermore, 68% of the CTAs in the low *SDI* quintile feature significant coefficients on *TSMOM* versus just 13% of the CTAs in the high *SDI* quintile. The momentum factor's explanatory power in terms of average regression R^2 is 0.25 for the low *SDI* CTAs and 0.01 for the high *SDI* CTAs. In all cases the difference between the regression results for the two extreme quintiles is highly significant. These results indicate that low *SDI* CTAs are pursuing momentum strategies – and it is reasonable to expect that their success is largely dependent on the degree to which futures markets feature momentum.

We also report the percentage of CTAs in each quintile that fail in the ensuing 12 months in Table 8. As described in Section II, we use the Liang and Park (2010) method of identifying the subset of instances of cessation of reporting that can be interpreted as failures. There is a large difference in the failure rates of low and high *SDI* funds, with less distinctive funds failing at a rate of 2% per annum, whereas more distinctive funds fail at a rate of 8% per annum. To the extent that failed funds continue to exist and generate poor returns, the performance difference between the extreme quintiles reported in Tables 5 – 7 likely understate the superiority of the conforming CTAs.

What explains the relatively weak performance of the highly distinctive CTAs? We have shown that conforming CTAs are following a momentum strategy, but what about the high *SDI*

CTAs? Bollen (2013) shows that hedge funds with high idiosyncratic risk (as measured by low factor model explanatory power) fail at a higher rate than other funds and our evidence for high *SDI* funds is consistent. However it might be the case that the high *SDI* CTAs are to some extent following an unspecified alternative strategy that underperforms. To assess the degree to which CTAs in the different *SDI* quintiles conform to a common yet unspecified strategy, we perform Principal Component Analysis (PCA) to extract the first principal component for each quintile. For each quintile we then estimate the cross sectional variance explained by the first principal component and the correlation between that component and the return on the *TSMOM* factor. As the sample of funds have returns over different time periods, we follow Fung and Hsieh (2001) and conduct the analysis on the maximum number of funds that had returns within a common sub-sample period, sorting funds into quintiles based on their average *SDI* for the sub-sample period.

The results of this analysis, reported in Table 9, show that for the low *SDI* funds the first principal component explains between 69% and 73% of the total variance across funds in the quintile, depending on the length of the sub-sample period. In contrast, for the high *SDI* funds the first principal component explains just 26% to 32% of the variance, depending upon the time period length. Note that the first principal component for low *SDI* funds is highly correlated to *TSMOM* (coefficient ranging from 0.77 to 0.86), whereas the first principal component for high *SDI* funds is only weakly correlated with *TSMOM*. The difference between the correlation of *TSMOM* and the low and high *SDI* principal components is over 0.61.

The results indicate that the high *SDI* funds are in general idiosyncratic traders, as opposed to a set of managers that are following some other unidentified common strategy. This result reinforces the idea that for CTAs momentum trading is in fact the only game in town, so that those who pursue some other strategy are essentially noise traders and generate on average inferior performance.

To emphasize the importance of momentum for successful and conforming CTAs, we repeat the analysis that compares the ex-post performance of *SDI* sorted quintiles of CTAs but here split the sample into two subsets, depending on the sign of the return of the *TSMOM* trading strategy. The results, presented in Table 10, are clear. During positive *TSMOM* return months (more than 68% of the sample period), the average performance of quintiles is decreasing in *SDI*, as before. During positive momentum return months, for example, the alpha of the low *SDI* quintile is a significant 1.87% per month versus -0.06% per month for the high *SDI* quintile. During negative momentum return months, in contrast, the relationship between *SDI* and performance flips. These results demonstrate the importance of momentum for explaining the association between the relative uniqueness of a CTA strategy and relatively lower returns. To illustrate the relationship, we also plot in Figure 3 the low *SDI* portfolio cumulative returns along with vertical grey bars that indicate the ten years in which *TSMOM* returns are above their median. It is apparent that the low *SDI* CTAs generate much of their cumulative return during times when momentum strategies are successful, again consistent with the idea that the conforming CTAs are following a common momentum-based strategy.

V.B. Robustness of SDI as a Performance Predictor using a Single Strategy

The results above suggest that trading on momentum is the predominant strategy in CTAs, and that those managers who conform to the strategy reliably generate attractive returns relative to other managers. Recall that the approach taken by Sun et al. (2012) when computing *SDI* is to first form clusters of hedge funds as a means of sorting on strategy, and then assess the similarity of each fund to the others within-strategy. Our analyses above follows this procedure using eight clusters to match the number of main styles in the BarclayHedge database. If it's the case for CTAs that momentum trading is the only game in town, then all CTAs could be considered to belong to the same cluster.

To test the robustness of our main results to this view, we repeat the analysis measuring the relation between *SDI* and subsequent CTA performance except now compute *SDI* by using the correlation between a CTA's returns and the average return of all other CTAs. We measure the average return of *SDI* quintiles over subsequent horizons ranging from 3 months to 3 years, as well as risk-adjusted performance metrics. The results, unreported for the sake of brevity, are qualitatively identical to those in Table 5 and Table 6. Thus, our prior results are robust to the notion that all CTA managers can be clustered together as momentum traders.

V.C. Implementation of SDI as a Selection Criterion

In this subsection we study whether an investor could realize a benefit of using *SDI* to select a subset of CTAs for investment as opposed to all CTAs in a given quintile. Selecting a small subset likely will reduce the performance of an investor, as superior performers might be omitted and less diversification will be achieved, nonetheless it is useful to assess the feasibility of a selection strategy as a means of measuring the economic significance of *SDI* as a selection variable.

On January 1st of each year CTAs are sorted into quintiles using *SDI* computed from the previous 24 return months. Five funds are randomly selected within each quintile, with an equal initial weighting to each fund and monthly rebalancing. Returns are calculated for each five-CTA portfolio over the twelve month holding period. The process is repeated each year and the returns are concatenated into a time series for the full sample period. The simulation is then repeated 1,000 times.

Table 11 reports the average performance metrics of the 1,000 simulated strategies drawn from each quintile. As with our other analysis of the relation between *SDI* and subsequent performance, there is a monotonic decline in returns as the CTAs become more distinct. On a risk-adjusted basis, the average Sharpe ratio of the CTA portfolios drawn from the low *SDI* quintile is 0.23 versus -0.26 for the portfolios drawn from the high *SDI* quintile. The difference between the top and bottom quintiles is statistically significant for all performance measures.[12] These results indicate that investors could use the *SDI* to select a realistic subset of CTAs that reliably outperforms the others.

V.D. Alternative Momentum Portfolios

Baltas and Kosowski (2013) construct a number of alternative time series momentum factors by varying the window lengths and sorting frequencies to provide a richer characterization of the strategies investors might employ. Baltas and Kosowski (2013) show that momentum portfolios alternately formed with daily, weekly and monthly formation and holding periods have relatively low correlation, indicating that these details matter. To better assess the nature of momentum pursued by the CTA *SDI* quintiles in our study, we form nine *TSMOM* portfolios, mirroring the formation and holding periods in Baltas and Kosowski (2013). To verify that these portfolios are related to CTA performance similarly to our original *TSMOM* factor, we regress the returns of each CTA on each of the momentum factors. The results are all consistent with those reported in Table 8. CTAs in the low *SDI* quintile have average coefficients on the momentum factors that are much higher than CTAs in the high *SDI* quintile, irrespective of lookback or

[12] As pointed out by the referee, while the full-quintile portfolios generally feature positive MPPMs. the MPPMs of the five-CTA portfolios are on average negative, indicating that a risk-averse investor would prefer to hold the risk-free asset instead. The reason for the difference is that the smaller portfolios have higher volatility. However, in the context of a multi-asset portfolio, an investor might still benefit from an allocation to CTAs as a diversification vehicle.

holding period. Furthermore, the percentage of CTAs with statistically significant coefficients is itself significantly higher in the low *SDI* quintile compared to the high *SDI* quintile.

Baltas and Kosowski (2013) also show that a three factor combination of portfolios with lookback periods of 15 days, eight weeks and 12 months combined with holding periods of one day, one week and one month, respectively, are the optimal combination for explaining the returns of aggregate CTAs. In order to test the robustness of our results to this three factor combination we repeat the analysis in Table 8 using all three factors together. The results of this analysis are reported in Table 12. The explanatory power of the models is much higher, confirming the efficacy of the models in explaining the returns of CTAs. For the low *SDI* quintile, for example, the three-factor average R^2 is 46%, almost double that of the single *TSMOM* regression in Table 8. Consistent with our earlier findings, the average regression coefficients are highest for the low *SDI* quintile, and decrease monotonically, showing that the CTAs which conform the most are pursuing momentum strategies more heavily than other CTAs.

We also report in Table 12 the percentage of CTAs with significant regression coefficients as well as the percentage of explained variation that is attributable to each of the three factors. For the low *SDI* quintile, the first two factors, which have the longer lookback and holding periods, are significant about half the time, whereas the third factor with a 15 day lookback and one day holding period is significant only about a quarter of the time. More important, all three factors contribute meaningfully to the model's explanatory power, indicating that the successful low *SDI* quintile CTAs pursue momentum at multiple horizons simultaneously. We leave additional analysis of variation in formation and holding periods of momentum strategies, and their impact on CTA performance, for future research.

VI. Conclusions

Prior research has shown that strategy distinctiveness is a key determinant of cross sectional differences in hedge fund and mutual fund performance. It is intuitive that funds with more unique strategies might outperform, as the returns to more well-known strategies are competed away. However, CTA managers are limited to investing in a well-defined set of securities: futures contracts. In the classic decomposition of managerial skill into security selection and market timing, there are no opportunities to differentiate along the selection dimension. Futures markets are characterized by a high level of momentum, leading to the prevalence of trend following strategies. If momentum trading is the only game in town, then offering a conforming, as opposed to a distinctive, series of returns may be optimal.

We measure the distinctiveness of a CTA's investment strategy using the strategy distinctiveness index (*SDI*) of Sun et al. (2012). We estimate the correlation of a CTA's return with that of its peers and classify funds with low correlation as high *SDI* funds. We find that *SDI* is highly persistent at horizons up to three years, indicating that this property is a stable characteristic of individual CTAs. Our central result is that, in complete contrast to prior literature on *SDI* and hedge funds, *SDI* is negatively associated with future CTA performance. CTAs which are more unique tend to underperform, after controlling for risks and styles, irrespective of holding period.

Following the original insight of Fung and Hsieh (1997), we posit that the conforming CTAs are pursuing a trend-following strategy, and so investigate the relation between subsequent CTA performance and the returns of a time-series momentum strategy, *TSMOM*. We find that the returns of conforming CTAs are highly correlated with *TSMOM*. Furthermore, we split the sample into months in which the return on *TSMOM* is positive and negative. Our evidence indicates *SDI* is an informative measure for predicting CTA performance only during times when *TSMOM* yields positive returns. The best performing CTAs trade largely on momentum, therefore, and offer investors exposure to this strategy.

Our results shed light on a tradeoff that exists for investors in investment vehicles that are not subject to the disclosure requirements of the Investment Company Act. Skilled managers can better exploit their advantage by keeping their trading technology secret; in the model of Glode and Green (2011), a degree of opacity is a necessary component for the manager to share the benefit with their investors. Yet with incomplete information, investors have a more difficult time monitoring their managers. Gorovyy et al. (2020) show that a proprietary measure of fund secrecy is negatively related to performance in down markets, which they associate with the presence of unrewarded and hidden risks. In the context of CTAs, the well-defined and profitable nature of momentum-based futures contract trading strategies appears to render idiosyncratic bets especially perilous for investors.

Data Availability Statement

The futures data that support the findings of this study are available from Thomson Reuters at https://eikon.thomsonreuters.com/index.html. The CTA data are available from BarclayHedge at https://www.barclayhedge.com/databases/cta-database. The Fung and Hsieh factors are available from David Hsieh at https://faculty.fuqua.duke.edu/~dah7/HFData.htm.

References

Agarwal, V., N. D. Daniel, and N. Y. Naik, 2009. Role of managerial incentives and discretion in hedge fund performance, Journal of Finance 64, 2221-2256.

Aiken, A., C. Clifford, and J. Ellis, 2013. Out of the dark: Hedge fund reporting biases and commercial databases, Review of Financial Studies 26, 208-243.

Aragon, G. O., 2007. Share restrictions and asset pricing: Evidence from the hedge fund industry, Journal of Financial Economics 83, 33-58.

Bali, T. G., S. J. Brown, and M. O. Caglayan, 2011. Do hedge funds' exposures to risk factors predict their future returns? Journal of Financial Economics 101, 36-68.

Bali, T. G., S. J. Brown, and M. O. Caglayan, 2012. Systematic risk and the cross section of hedge fund returns, Journal of Financial Economics 106, 114-131.

Bali, T. G., S. J. Brown, and M. O. Caglayan, 2014. Macroeconomic risk and hedge fund returns, Journal of Financial Economics 114, 1-19.

Baltas, N., and R. Kosowski, 2013. Momentum strategies in futures markets and trend-following funds, Working paper.

Bhardwaj, G., G. B. Gorton, and K. G. Rouwenhorst, 2014. Fooling some of the people all of the time: The inefficient performance and persistence of commodity trading advisors, Review of Financial Studies 27, 3099-3132.

Bollen, N. P., 2013. Zero-R^2 hedge funds and market neutrality, Journal of Financial and Quantitative Analysis 48, 519-547.

Bollen, N.P., J. Joenväärä, and M. Kauppila, 2020. Hedge fund performance: End of an era?, Vanderbilt University Working Paper.

Brown, S. J., and W. N. Goetzmann, 1997. Mutual fund styles, Journal of Financial Economics 43, 373-399.

Brown, S. J., and W. N. Goetzmann, 2003. Hedge funds with style, Journal of Portfolio Management 29, 101-112.

Brunnermeier, M., and S. Nagel, 2004. Hedge funds and the technology bubble, Journal of Finance 59, 2013-2040.

Chen, H., L. Cohen, and U. Gurun, 2019. Don't take their word for it: The misclassification of bond mutual funds. U. of Notre Dame Working Paper.

Cremers, K. M., and A. Petajisto, 2009. How active is your fund manager? A new measure that predicts performance, The Review of Financial Studies 22, 3329-3365.

Daniel, K., and T. Moskowitz, 2016. Momentum crashes, Journal of Financial Economics 122, 221-247.

Fung, W., and D. A. Hsieh, 1997. The information content of performance track records: Investment style and survivorship bias in the historical returns of commodity trading advisors, Journal of Portfolio Management 24, 30-41.

Fung, W., and D. A. Hsieh, 2001. The risk in hedge fund strategies: Theory and evidence from trend followers, The Review of Financial Studies 14, 313-341.

Fung, W., and D. A. Hsieh, 2004. Hedge fund benchmarks: A risk-based approach, Financial Analysts Journal, 65-80.

Getmansky, M., A. W. Lo, and I. Makarov, 2004. An econometric model of serial correlation and illiquidity in hedge fund returns, Journal of Financial Economics 74, 529-609.

Glode, V., and Green, R. (2011). Information spillovers and performance persistence for hedge funds. Journal of Financial Economics 101, 1-17.

Goetzmann, W., J. Ingersoll, M. Spiegel, and I. Welch, 2007. Portfolio performance manipulation and manipulation-proof performance measures, The Review of Financial Studies 20, 1503-1546.

Goetzmann, W. N., J. E. Ingersoll Jr, and S. A. Ross, 2003. High-water marks and hedge fund management contracts, Journal of Finance 58, 1685-1718.

Gorovyy, S., O. Kuzmina, and P. Kelley, 2020. Does secrecy signal skill? Characteristics and performance of secretive hedge funds. U. of Melbourne Working Paper.

Jorion, P., and C. Schwarz, 2014. Are hedge fund managers systematically misreporting? Or not? Journal of Financial Economics 111, 311-327.

Kacperczyk, M., C. Sialm, and L. Zheng, 2005. On the industry concentration of actively managed equity mutual funds, Journal of Finance 60, 1983-2011.

Liang, B., and H. Park, 2010. Predicting hedge fund failure: A comparison of risk measures. Journal of Financial and Quantitative Analysis 45, 199-222.

Moskowitz, T. J., Y. H. Ooi, and L. H. Pedersen, 2012. Time series momentum, Journal of Financial Economics 104, 228-250.

Narayan, P., H. Ahmed, and S. Narayan, 2014. Do momentum-based trading strategies work in the commodity futures markets? Journal of Futures Markets 35, 868-891.

National Association of College and University Business Officers, 2018. NACUBO-TIAA Study of endowments.

Sharpe, W., 1992. Asset allocation: Management style and performance measurement, Journal of Portfolio Management 18, 7-19.

Shen, Q., A. Szakmary, and S. Sharma, 2007. An examination of momentum strategies in commodity futures markets, Journal of Futures Markets 27, 227-256.

Shleifer, A., and R. W. Vishny, 1997. The limits of arbitrage, The Journal of Finance 52, 35-55.

Sun, Z., A. Wang, and L. Zheng, 2012. The road less traveled: Strategy distinctiveness and hedge fund performance, Review of Financial Studies 25, 96-143.

Titman, S., and C. Tiu, 2010. Do the best hedge funds hedge? Review of Financial Studies 24, 123-168.

Treynor, J. L., and F. Black, 1973. How to use security analysis to improve portfolio selection, Journal of Business 46, 66-86.

[dataset] BarclayHedge. (2020). *Managed Futures (CTA) Database*. Available from BarclayHedge web site: https://www.barclayhedge.com/databases/cta-database

[dataset] Hsieh, D. (2020). *David A. Hsieh's Hedge Fund Data Library*. Available from David Hsieh's web site: https://faculty.fuqua.duke.edu/~dah7/HFData.htm

[dataset] Thomson Reuters. (2020). *Commodity Futures Data*. Available from Thomson Reuters web site: https://eikon.thomsonreuters.com/index.html

Figure 1. Histograms of *SDI*

Depicted in black bars is the histogram of *SDI* when correlations are measured between a fund's returns and the returns of its style cluster using the Brown and Goetzmann (2003) clustering algorithm. Depicted in grey bars is the histogram of *SDI* when correlations are measured between a fund's returns and the returns of a portfolio of all funds in its style as self-reported to the BarclayHedge database.

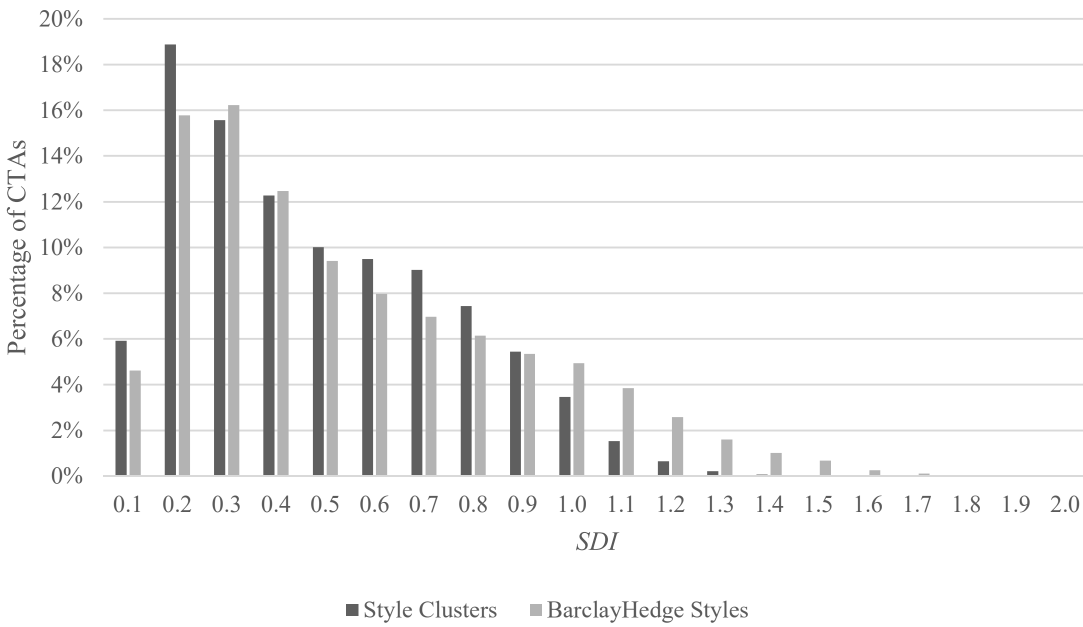

Figure 2. CTA sample size

Depicted are the number of funds reporting returns to the BarclayHedge database that survive our data filters each year from 1994 to 2015 as well as the total assets under management in USD billion of the funds in our sample.

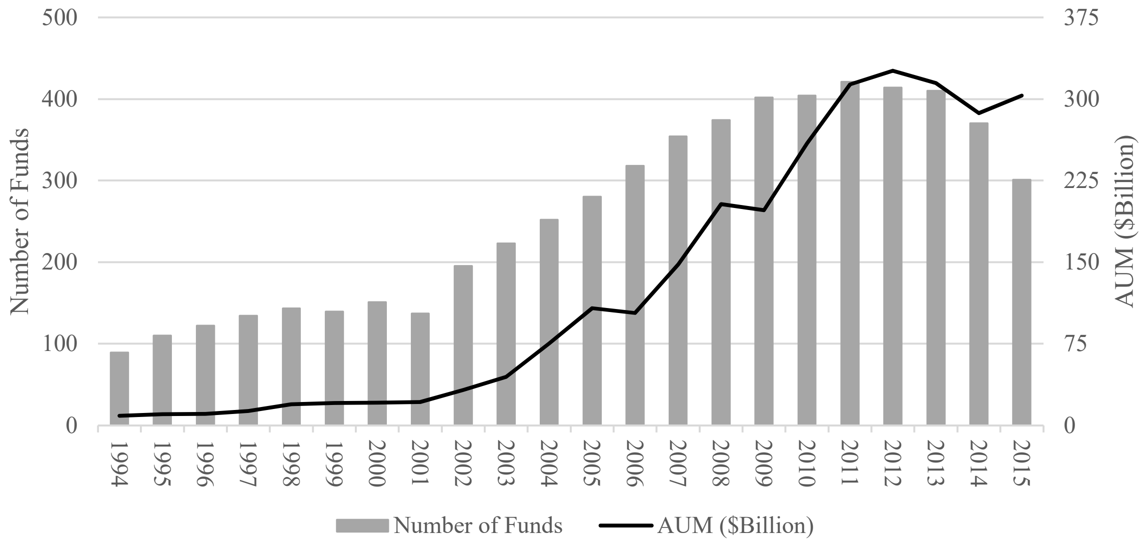

Figure 3. Cumulative return of the low *SDI* quintile portfolio

Depicted is the cumulative return of the low *SDI* quintile portfolio formed by annual sorts. Vertical grey bars indicate the ten years out of 20 in the sample in which the time series momentum factor *TSMOM* delivers an above-median return.

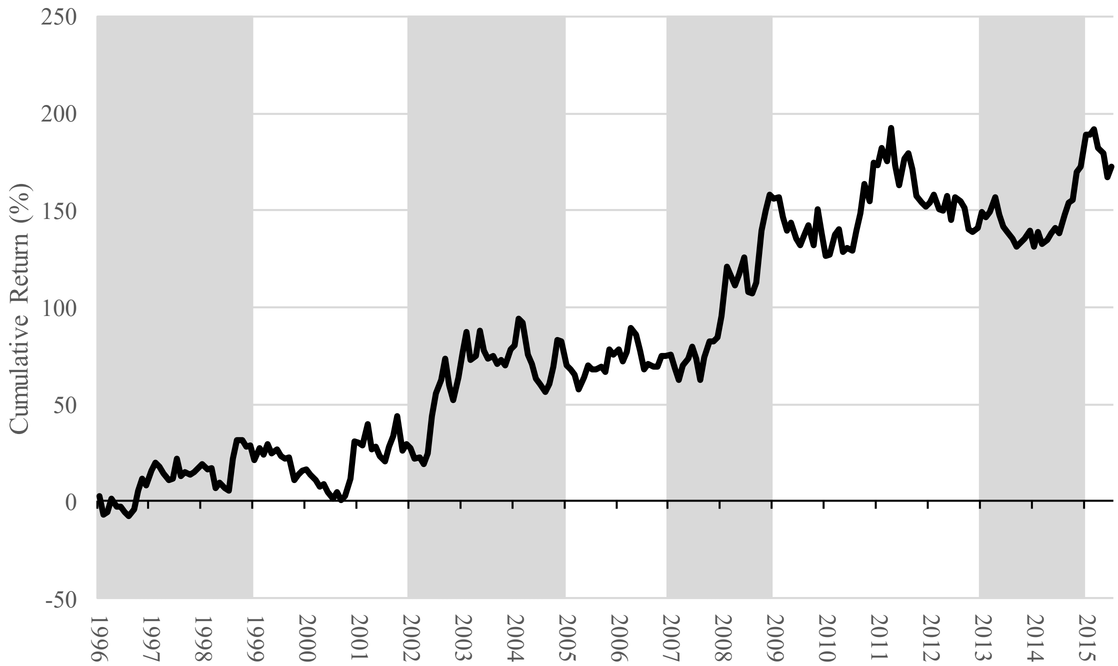

Table 1. Futures data

Listed are the futures contracts used to construct the time series momentum factor *TSMOM* and their serial correlations, ρ, with significance levels from the Ljung-Box test statistic. The futures contracts are divided into four classes: Commodities, Government Bonds, Equity Indices, and Currencies. All data series run from January 1993 to July 2015. DEM/USD is used in place of EUR/USD prior to the introduction of the Euro in January 1999. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

Commodities	ρ	*Equity Indices*	ρ
COCOA	-5.93**	SPI 200	4.55
COFFEE	7.47*	S&P TSX60	8.43*
COPPER	11.07**	SMI	7.81**
CORN	4.06	DAX	2.66
COTTON	5.53	IBEX 35	2.82
GOLD	-4.05	CAC 40	3.45
LEAN HOGS	-5.87	FTSE 100	-0.82
LIGHT CRUDE	11.20*	HANG SENG	0.60
LIVE CATTLE	-0.83	MIB	2.09
NYHEATINGOIL	6.96	NIKKEI 225	11.84***
PALLADIUM	4.95	AEX	5.79
PLATINUM	8.24	S&P 500	4.99
RBOB GASOLINE	3.68*		
SILVER	-3.54	*Currencies*	ρ
SOYABEAB MEAL	-4.77	AUD/USD	6.39*
SOYABEAB OIL	-1.29	CAD/USD	-0.43
SOYABEANS	-4.28*	CHF/USD	0.73
SUGAR	3.83*	EUR/USD	4.01
WHEAT	1.38	GBP/USD	5.54
		JPY/USD	7.41
Government Bonds	ρ	NOK/USD	4.60
Australia-10Y	3.00**	NZD/USD	8.24***
Canada-10Y	2.67	SEK/USD	3.78
Switzerland-10Y	8.33		
Germany-10Y	4.66		
Denmark-10Y	8.90***		
UK-10Y	3.21		
Japan-10Y	-3.01**		
US-10Y	0.40**		

Table 2. Summary statistics

Listed are annualized summary statistics of monthly CTA returns. Listed are: average return, average return in excess of the 3-month T-Bill rate, volatility, assets under management ($m), Sharpe ratio, Fung and Hsieh (2004) 7-factor alpha, appraisal ratio, Fung and Hsieh (2004) 7-factor R^2, and the Sun et al. (2012) strategy distinctiveness index (*SDI*). Listed are the cross-sectional mean and three percentiles of each statistic. A test for significant differences between live and dead *SDI* is conducted. *** denotes significance at a 1% level.

		Percentile		
	Mean	20^{th}	Median	80^{th}
Panel A: All Funds				
Average Return	3.14	-1.78	2.57	7.78
Excess Return	1.33	-3.94	0.86	6.23
Volatility	14.05	6.40	11.64	20.41
AUM	338.95	14.06	40.62	186.38
Sharpe Ratio	-0.01	-0.50	0.08	0.47
Alpha	1.33	-4.72	0.72	6.80
Appraisal Ratio	-0.01	-0.63	0.06	0.63
R^2	0.16	0.01	0.13	0.32
SDI	0.42	0.17	0.36	0.66
Panel B: Live Funds				
Average Return	5.48	0.92	4.73	9.72
Excess Return	4.91	0.60	4.30	8.66
Volatility	12.98	7.17	11.12	18.10
AUM	794.80	22.44	69.49	381.56
Sharpe Ratio	0.49	0.08	0.36	0.74
Alpha	4.21	-0.65	4.16	9.12
Appraisal Ratio	0.48	-0.05	0.48	0.85
R^2	0.18	0.03	0.14	0.34
SDI	0.33	0.14	0.26	0.51
Panel C: Dead Funds				
Average Return	2.21	-2.58	1.30	6.58
Excess Return	-0.10	-5.19	-0.72	4.42
Volatility	14.47	6.19	11.76	21.07
AUM	158.46	13.06	32.14	131.75
Sharpe Ratio	-0.21	-0.64	-0.05	0.35
Alpha	0.20	-6.03	-0.60	5.30
Appraisal Ratio	-0.20	-0.77	-0.07	0.45
R^2	0.15	0.00	0.12	0.31
SDI	0.45	0.20	0.40	0.69
$SDI_{Live} - SDI_{Dead}$	-0.12***			

Table 3. Summary statistics of *SDI* quintiles

Quintiles are formed by sorting CTAs by *SDI* measured over their entire reported history. Listed are averages of individual CTA annualized summary statistics of monthly returns: the Sun et al. (2012) strategy distinctiveness index (*SDI*), the Fung and Hsieh (2004) 7-factor R^2, fund age (years), AUM in $millions, average return, average return in excess of the 3-month T-Bill rate, volatility, Sharpe ratio, the Fung and Hsieh (2004) 7-factor alpha, and appraisal ratio. Tests for significant differences between quintiles Q1 and Q5 are conducted. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

SDI Quintile		*SDI*	R^2	Age	AUM	Avg. Return	Excess Return	Volatility	Sharpe Ratio	Alpha	Appraisal Ratio
Low	Q1	0.12	0.27	6.90	584.95	5.08	2.99	19.56	0.15	3.23	0.10
	Q2	0.25	0.20	6.19	162.67	4.16	2.22	16.19	-0.02	2.23	0.02
	Q3	0.41	0.15	6.15	555.20	3.92	2.15	14.08	0.06	1.98	0.09
	Q4	0.59	0.12	5.63	232.59	1.12	-0.43	10.87	-0.18	-0.77	-0.14
High	Q5	0.89	0.06	5.33	152.85	1.40	-0.32	9.54	-0.08	0.00	-0.11
	Q1-Q5	-0.77***	0.21***	1.57 ***	432.10**	3.68 ***	3.31***	10.02***	0.23**	3.23**	0.21*

Table 4. Persistence of *SDI*

Quintiles are formed monthly by sorting CTAs by *SDI* measured over the prior 24 months. The *SDI* of each CTA is then measured at future dates ranging from three months to three years as listed. Averages of the *SDI*s are reported by quintile at each future date. Tests for significant differences between quintiles Q1 and Q5 are conducted. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

SDI Quintile		Subsequent *SDI* 3M	6M	1Y	2Y	3Y
Low	Q1	0.13	0.14	0.14	0.16	0.17
	Q2	0.23	0.24	0.24	0.24	0.24
	Q3	0.37	0.37	0.37	0.37	0.37
	Q4	0.55	0.55	0.55	0.54	0.53
High	Q5	0.80	0.78	0.76	0.72	0.70
	Q1-Q5	-0.67***	-0.64***	-0.62***	-0.56***	-0.53***

Table 5. Portfolio returns of *SDI* quintiles

Quintiles are formed quarterly by sorting CTAs by *SDI* measured over the prior 24 months. Quintile portfolios are then held for periods ranging from three months to three years as listed. Annualized time series means of the post-formation return of each holding period is reported by quintile. Tests for significant averages and significant differences between quintiles Q1 and Q5, adjusted for heteroscedasticity and autocorrelation, are conducted. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

		Return				
SDI Quintile		3M	6M	1Y	2Y	3Y
Low	Q1	6.28**	6.68**	7.05***	6.24***	6.27***
	Q2	5.12**	5.14**	5.72***	5.69***	5.93***
	Q3	3.92**	4.96***	5.38***	5.29***	6.11***
	Q4	3.64***	3.36***	3.64***	4.11***	4.63***
High	Q5	1.72***	1.76***	1.90***	2.53***	3.09***
	Q1-Q5	4.56	4.92*	5.15**	3.71*	3.18*

Table 6. Portfolio performance of *SDI* quintiles

Quintiles are formed quarterly by sorting CTAs by *SDI* measured over the prior 24 months. Quintile portfolios are then held for periods ranging from three months to three years as listed. Time series means of the post-formation performance statistics for each holding period are reported by quintile. Tests for significant performance and significant differences between quintiles Q1 and Q5, adjusted for heteroscedasticity and autocorrelation, are conducted. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

		Alpha					Appraisal Ratio					Sharpe Ratio				
SDI Quintile		3M	6M	1Y	2Y	3Y	3M	6M	1Y	2Y	3Y	3M	6M	1Y	2Y	3Y
Low	Q1	5.24***	5.02***	4.59***	5.50***	6.47***	0.25***	0.21***	0.17***	0.19***	0.21***	0.12**	0.15***	0.17***	0.16***	0.18***
	Q2	2.36***	2.72***	3.13***	3.96***	5.47***	0.10*	0.14***	0.14***	0.16***	0.21***	0.12**	0.15***	0.17***	0.18***	0.22***
	Q3	1.20*	0.64	0.87	1.51**	2.52***	0.02	-0.01	0.00	0.06	0.12***	0.01	0.07	0.09**	0.14***	0.19***
	Q4	-0.16	0.64	1.19**	0.86	1.06*	-0.07	-0.09	-0.03	-0.05	-0.01	-0.05	-0.05	0.00	0.00	0.04
High	Q5	-0.96*	-0.60	-0.43	-0.10	-0.14	-0.24***	-0.26***	-0.28***	-0.29***	-0.31***	-0.27***	-0.29***	-0.29***	-0.31***	-0.36***
	Q1-Q5	6.20***	5.60***	5.02***	5.60***	6.61***	0.48***	0.47***	0.46***	0.48***	0.52***	0.39***	0.44***	0.46***	0.48***	0.54***

Table 7. Manipulation Proof Performance Measure of *SDI* quintiles

Quintiles are formed quarterly by sorting CTAs by *SDI* measured over the prior 24 months. Quintile portfolios are then held for periods ranging from three months to three years as listed. The Manipulation Proof Performance Measure (MPPM) of each holding period measured over the full sample is reported by quintile. Tests for significant MPPM and significant differences between quintiles Q1 and Q5 are conducted. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

		MPPM				
SDI Quintile		3M	6M	1Y	2Y	3Y
Low	Q1	0.56**	0.25*	0.50**	0.23*	0.30**
	Q2	-0.09	0.08*	-0.04	0.23*	0.30**
	Q3	0.71**	1.08***	1.33***	1.27***	1.16***
	Q4	0.60**	0.35**	0.24**	0.19*	0.16*
High	Q5	-1.31	-1.46	-1.77	-1.62	-1.63
	Q1-Q5	1.87**	1.71*	2.28**	1.85**	1.94**

Table 8. Momentum exposure of *SDI* quintiles

Quintiles are formed annually by sorting CTAs by *SDI* measured over the prior 24 months. Each CTA's returns over the prior 24 months are then regressed on the contemporaneous time series momentum factor *TSMOM*. Listed are average regression coefficient β on *TSMOM*, the percentage of CTAs with β statistically significant at the 10% level, and the average adjusted R^2. Also reported for each quintile is the failure rate, defined as the average percentage of CTAs that fail in the ensuing 12 months with failures identified following the Liang and Park (2010) procedure. Tests for significant differences between quintiles Q1 and Q5 are conducted. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

SDI quintile		Average β	% Sig.	R^2	Failure Rate
Low	Q1	0.67	0.68	0.25	0.02
	Q2	0.51	0.58	0.17	0.05
	Q3	0.29	0.42	0.09	0.05
	Q4	0.10	0.25	0.03	0.05
High	Q5	-0.01	0.13	0.01	0.08
	Q1-Q5	0.68***	0.55***	0.24***	-0.06***

Table 9. Principle Component Analysis of *SDI* quintiles

Listed are the results of a Principal Component Analysis of the funds within *SDI* quintiles. As the sample of funds have returns over different time periods, the analysis was conducted on the maximum number of funds that had returns over a common sub-sample period. Sub-sample periods are 72 months, 60 months and 48 months in Panel A, B and C, respectively. Quintiles are formed by sorting CTAs by *SDI* measured over each sub-sample period. Tests for significant correlations and significant differences between quintiles Q1 and Q5 are conducted. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

SDI quintile		Variance Explained	*TSMOM* Correlation	Num. Funds	Start Date	End Date
Panel A: 72 Months Overlapping Returns						
Low	Q1	0.70	0.77***	27	07-2005	06-2011
	Q2	0.58	0.55***	26	07-2005	06-2011
	Q3	0.42	0.20*	27	07-2005	06-2011
	Q4	0.33	-0.01	26	07-2005	06-2011
High	Q5	0.32	-0.02	27	07-2005	06-2011
	Q1-Q5	0.38	0.79***			
Panel B: 60 Months Overlapping Returns						
Low	Q1	0.69	0.79***	34	07-2006	06-2011
	Q2	0.46	0.56***	34	07-2006	06-2011
	Q3	0.27	-0.19	34	07-2006	06-2011
	Q4	0.43	0.06	34	07-2006	06-2011
High	Q5	0.28	0.15	34	07-2006	06-2011
	Q1-Q5	0.41	0.64***			
Panel C: 48 Months Overlapping Returns						
Low	Q1	0.73	0.86***	42	07-2007	06-2011
	Q2	0.45	0.48***	41	07-2007	06-2011
	Q3	0.27	-0.11	42	07-2007	06-2011
	Q4	0.45	0.13	41	07-2007	06-2011
High	Q5	0.26	0.25*	42	07-2007	06-2011
	Q1-Q5	0.47	0.61***			

Table 10. The relation between momentum and *SDI* quintile performance

Quintiles are formed monthly by sorting CTAs by *SDI* measured over the prior 24 months. Months are partitioned into two subsamples *Pos* and *Neg* depending on the sign of the one-month return of the momentum factor *TSMOM*, which is positive 68.5% of the time. CTA performance is then assessed and averaged over the months in the *Pos* and *Neg* subsamples. Results are reported for average returns and four performance measures: Alpha, Appraisal, Sharpe ratio and MPPM. Tests for significant performance and significant differences between quintiles Q1 and Q5 are conducted within the two subsamples. The bottom row reports differences in Q1-Q5 across the two subsamples. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

		Average Return		Alpha		Appraisal		Sharpe		MPPM	
		Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg
SDI quintile		*TSMOM*	*TSMOM*	*TSMOM*	*TSMOM*	*TSMOM*	*TSMOM*	*TSMOM*	*TSMOM*	*TSMOM*	*TSMOM*
Low	Q1	1.99***	-1.82***	1.87***	-1.48***	0.55***	-0.40***	0.35***	-0.56***	18.12***	-30.45***
	Q2	1.60***	-1.52***	1.39***	-1.24***	0.50***	-0.45***	0.36***	-0.59***	14.33***	-23.18***
	Q3	1.18***	-0.81***	0.87***	-0.63**	0.39***	-0.26**	0.31***	-0.54***	9.48***	-13.07***
	Q4	0.54***	-0.16	0.35***	-0.17	0.25***	-0.11	0.19**	-0.26**	3.00***	-3.76***
High	Q5	0.15**	-0.10	-0.06	-0.19	-0.07	-0.18*	-0.06	-0.28***	-0.98	-1.90*
	Q1-Q5	1.84***	-1.72***	1.93***	-1.29***	0.61***	-0.22	0.42***	-0.29**	19.11***	-28.55***
	Pos-Neg	3.56***		3.22***		0.84***		0.70***		47.65***	

Table 11. Performance of 5-fund CTA portfolios by *SDI* quintile

Listed are performance levels of trading strategies that select 5 CTAs at random annually from quintiles formed by *SDI* measured over the prior 24 months. Results are averaged across 1,000 simulations. For each quintile the annualized average return, Fung and Hsieh alpha, Fung and Hsieh appraisal ratio, Sharpe ratio and MPPM are reported. Tests for significant performance and significant differences between quintiles Q1 and Q5 are conducted. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

SDI quintile		Average Return	Alpha	Appraisal	Sharpe	MPPM
Low	Q1	6.30***	4.80***	0.31***	0.23***	-0.60
	Q2	5.16***	3.11*	0.23*	0.20*	-0.60
	Q3	4.47***	1.99	0.17	0.19	-0.42
	Q4	3.39**	1.44	0.15	0.10	-0.36
High	Q5	0.66	-1.72	-0.25	-0.26	-2.55
	Q1-Q5	5.65***	6.52***	0.55***	0.49***	1.95***

Table 12. CTAs and 3-Factor Momentum

Listed for quintiles of CTAs sorted annually on their *SDI* measure are average regression coefficients, the percentage of CTAs with statistically significant (at the 10% level) regression coefficients, as well as the percentage of the variation explained by each of the three factors, from a regression of CTA returns on the Baltas and Kosowski (2013) time series momentum factors $TSMOM_{12M,1M}$, $TSMOM_{8W,1W}$ and $TSMOM_{15D,1D}$, estimated over the prior 24 months. Also reported is the average adjusted R^2. Tests for significance of the average regression coefficients and significant differences between quintiles Q1 and Q5 are conducted. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

SDI quintile		$TSMOM_{12M,1M}$	% Sig.	% Var. Exp.	$TSMOM_{8W,1W}$	% Sig.	% Var. Exp.	$TSMOM_{15D,1D}$	% Sig.	% Var. Exp.	R^2
Low	Q1	0.27**	0.48	0.31	0.68***	0.54	0.46	0.37***	0.28	0.22	0.46
	Q2	0.21***	0.33	0.29	0.51***	0.37	0.41	0.33***	0.30	0.30	0.34
	Q3	0.11**	0.24	0.30	0.33***	0.27	0.38	0.25***	0.24	0.32	0.22
	Q4	0.01	0.15	0.31	0.15***	0.16	0.37	0.10***	0.17	0.32	0.09
High	Q5	-0.03***	0.13	0.30	0.02	0.10	0.34	0.03	0.16	0.36	0.02
	Q1-Q5	0.30**	0.35***	0.01	0.67***	0.44***	0.13***	0.34***	0.12***	-0.14**	0.44***