3-2011

Stochastic Dominance Analysis of CTA Funds

H.H. Lean
Universiti Sains Malaysia

Kok Fai PHOON
Singapore Management University, kfphoon@smu.edu.sg

W.K. Wong
Hong Kong Baptist University

Stochastic Dominance Analysis of CTA Funds

Hooi Hooi Lean
School of Social Science
Universiti Sains Malaysia
11800 Minden,
Penang, Malaysia.
learnmy@gmail.com

Kok Fai Phoon
Lee Kong Chian School of Business
Singapore Management University
50 Stamford Road
Singapore 178899
kfphoon@smu.edu.sg
Tel: 65-68280146

Wing-Keung Wong
Department of Economics
Hong Kong Baptist University
Kowloon Tong
Hong Kong
awong@hkbu.edu.hk

Mar 2011

Stochastic Dominance Analysis of CTA Funds

ABSTRACT

In this paper, we employ the stochastic dominance approach to rank the performance of commodity trading advisors (CTA) funds. An advantage of this approach is that it alleviates the problems that can arise if CTA returns are not normally distributed by utilizing the entire returns distribution. We find both first-order and higher-order stochastic dominance relationships amongst the CTA funds and conclude that investors would be better off investing in the first-order dominant funds to maximize their expected utilities and expected wealth. However, for higher-order dominant CTA, risk-averse investors can maximize their expected utilities but not their expected wealth. We conclude that the stochastic dominance approach is more appropriate compared with traditional approaches as a filter in the CTA selection process given that a meaningful economic interpretation of the results is possible as the entire return distribution is utilized when returns are non-normal.

Acknowledgments: The first author would like to acknowledge Universiti Sains Malaysia (RU Grant No. 1001/PSOSIAL/816094),

JEL Classification: G11, G15

Key words: commodity trading advisors funds, stochastic dominance, risk-averse investors, performance measurement.

Stochastic Dominance Analysis of CTA Funds

1. INTRODUCTION

Commodity trading advisors (CTAs) are professional money managers investing in global futures and options markets. CTAs have evolved to investing in more diversified holdings that include currency, financial, and other more liquid derivative contracts. Hence, CTAs are also referred to a “Managed Futures.”

An often cited reason for investing in the alternative universe is that an investor holding traditional stocks and bonds can enhance returns without adding to volatility. The attraction of alternative investing has clearly been demonstrated in the dramatic growth in assets managed by hedge funds over the past decade and in managed futures more recently. Assets under management of CTAs in the CSFB/Tremont CTA index grew from less than $10 billion at year-end 2001 to $28 billion by year-end 2003. As explained by Collins (2005), managed futures have proven to be less correlated to equities than other hedge fund strategies and while providing daily transparency and liquidity. Moreover, managed futures tend to be non-correlated with equities in bull markets and negatively correlated in bear markets (Lee et. al., 2004).

CTAs can utilize many strategies though the majority of money allocated to it still falls into the medium to long-term trend following camp. Collins (2005) estimated that 80% to 90% of CTAs are involved in trend following, viz., a strategy that tries to take advantage of up and down trends in various markets. Trend following had a miserable year in 2005 highlighting the paramount importance of understanding the investment risks of CTAs. A well known trend-following fund ran by John Henry had a rough start in 2005, being down 36%, representing a drawdown of 54%. There has been considerable works done to understand the risk of hedge funds. Some of the earlier papers that examine the performance and risk characteristics of hedge funds (for example Fung and Hsieh (2000) included CTAs in their hedge fund universe). In this paper, we focus on CTAs as they are more homogeneous than the large diverse hedge fund universe. Further, CTAs are viewed by some to provide "better" risk-adjusted returns enhancement. Kat (2004a) introduced the possibility of combining CTAs and hedge funds in a portfolio as the positive skewness of CTAs can help reduce the impact of negative skewness which can be a problem in hedge fund strategies. We should note that while superficially, CTAs are mentioned in the same breath as the global macro hedge fund strategy, they differ in the ability to capture trends under different market conditions. Specific differences in these two classes of strategies include the way they trade, manage risk and their investment time horizon.

There remains considerable debate among academics and professionals on assessing alternative investments for inclusion in their portfolios. One point that receives general agreement is that traditional criteria like using the Sharpe ratio will in many cases lead to erroneous selection. One of the main issues is that hedge fund returns are not normality distributed and are often not even close to it. Kat (2004b) highlighted that even the returns of funds of hedge funds are possibly skewed and leptokurtic. He pointed that investors wishing to use funds of hedge funds in risk reduction or yield enhancement must know how to hedge against negative skewness that can be expected when hedge funds are added to their portfolio. Further, Vuille and Crisan (2004) confirmed that CTA return distributions are non-normal and showed that a "buy and hold" multi-factor linear model fails to explain CTA returns.

As most of the traditional approaches to performance and risk measurement rely on the normality assumption, new approaches have been proposed. For example, Lee et al. (2006) proposed a practical approach to filter hedge funds based on past returns. In the use of this approach, investors are assumed to have sophisticated preferences – i.e., they like downside protection, whilst looking for yield enhancement. In this paper, on the other hand, we will rely on a selection methodology couched on traditional expected utility theory. We employ the stochastic dominance (SD) approach to rank the performance of CTAs. An advantage of this approach is that it alleviates the problems that can arise if CTA returns are not normally distributed because it utilizes the entire returns distribution. Our approach also allows for meaningful economic interpretation of the results based on non-satiation and risk-aversion. Section 2 of this paper motivates the study. The data and methodologies employed are described in Section 3. Empirical results are provided in Section 4 and Section 5 concludes.

2. MOTIVATION OF THE STUDY

Asset managers view CTAs as an attractive alternative investment. Including CTAs in their investment portfolio can provide downside protection to extreme events in financial markets. Lee et al. (2004) found evidence suggesting that adding CTAs investments to an equity portfolio provide both the usual portfolio diversifications, and the CTA returns are also negatively correlated with equity indices returns during periods of marked downturns of equity markets.

Vuille and Crisan (2004) documented that positive skewness and excess kurtosis signify that a MV framework is not well suited to analyze CTAs, as the over simplistic assumptions it relies on prevent such a model from capturing some of CTAs' most attractive features. CTAs are viewed by some to provide "better" risk-adjusted returns enhancement. Kat (2004a) introduced the possibility of combining CTAs and hedge funds in a portfolio as the positive skewness of CTAs can help reduce the impact of negative skewness which can be a problem in hedge fund strategies.

Given the trend following nature of most CTA strategies, Kat (2003) noted that modern portfolio theory is too simplistic to deal with CTAs. He maintained that Sharpe ratios and standard alphas could be misleading in analyzing such investments. This makes the use of traditional performance measures questionable.

We recommend the SD approach that allows investors to appropriately rank fund performance without the need for strong assumptions on investors' utility functions or the returns distribution of assets. SD rules offer superior criteria on which to base investment decisions relative to the traditional MV and CAPM analysis because the assumptions underlying SD are less restrictive than those of the MV and CAPM. In addition, SD incorporates information on the entire distribution, rather than the first two moments and requires no precise assessment as to the specific form of the investors' risk preference or utility functions (Taylor and Yoder, 1999).

The SD approach had been used in the evaluation of performance of mutual funds since the 1970s (Levy and Sarnat, 1970; Porter, 1973). Later, Taylor and Yoder (1999) used the SD approach to compare the performance between load and no-load funds during the 1987 crash. Kjetsaa and Kieff (2003) documented that the SD approach provides a collateral and feasible strategy to reveal relative investment preferences by discriminating among and parsing the universe of mutual fund opportunities. In addition, Gasbarro et al. (2007) utilized both the SD approach and the CAPM criterion to compare the performance of 18 country market indices (iShares) and found that SD appears to be both more robust and discriminating than the CAPM in the ranking of the iShares.

We use the Davidson and Duclos (DD, 2000) test to determine if SD occurred among the 56 CTAs during our sample period. Apart from applying the SD approach to CTAs, we are also able to determine if the differences between any two returns cumulative density functions are statistically significant based on the DD test.

We propose using the SD approach to filter CTAs using past returns given that such returns are possibly non-normal. In any analysis on performance using past data, Kat and Menexe (2003) suggested that the benefit of a track record lies in the insights on the risk of one fund relative to another of the same strategy class. This was provided by SD analysis on the CTA class of funds.

3. DATA AND METHODOLOGY

3.1 Data

We use monthly returns of the 56 CTAs reported by the EurekaHedge database for the sample period from January 1995 to December 2004 in this study. As traditional U.S. based fund managers and investors may use the S&P 500 as the equity benchmark, we include the S&P 500 (IX1) in our study. If they are investing internationally, diversification benefits can be measured relative to a regional benchmark like the MSCI World (IX2) constructed by Morgan Stanley. For completeness, we also include Goldman Sachs Commodity (IX3), Lehman Global Aggregate US Universal (IX4), and Lehman US Universal: High Yield Corp. (IX5) indices. The risk-free rate and the global market index are proxied by the 3-month U.S. T-bill rate and the MSCI World (IX2) respectively.

For comparison, this study first employs the MV criterion and CAPM statistics to investigate the performance of CTA. By the MV criterion (Markowitz, 1952; Bai, et al., 2009, 2011a,b), for the returns of any two CTAs Y and Z with means μ_y and μ_z and standard deviations, σ_y and σ_z respectively, Y is said to dominate Z if $\mu_y \geq \mu_z$ and $\sigma_y \leq \sigma_z$. CAPM statistics include the beta, Sharpe ratio, Treynor's index and Jensen (alpha) index to measure performance developed by Sharpe (1964), Treynor (1965) and Jensen (1969). Sharpe (1964), Treynor (1965), Jensen (1969), and Leung and Wong (2008) provide detailed definitions of the indices and statistics.

Let F and G be the cumulative distribution functions (CDFs) and f and g are the corresponding probability density functions (PDFs) of the returns of two CTAs Y and Z respectively with common support of $[a, b]$ $(a < b)$. Define

$$H_0 = h \text{ and } H_j(x) = \int_a^x H_{j-1}(t)\,dt \text{ for } h = f, g\,,\ H = F, G \text{ and } j = 1, 2, 3\,. \quad (1)$$

CTA *Y* would dominate CTA *Z* by first-order SD (FSD) if and only if $F_1(x) \le G_1(x)$; by second-order SD (SSD) if and only if $F_2(x) \le G_2(x)$; and finally, by third-order SD (TSD) if and only if $F_3(x) \le G_3(x)$ for all *x*, and the strict inequality holds for at least one value of *x*; and *Y* has higher expected return than *Z*. Wong and Li (1999), Anderson (2004), Wong (2007), and Wong and Chan (2008) have discussed the definition in Equation (1) in detail.

The existence of SD implies that the expected utilities of investors are always higher when holding the dominant CTA than holding the dominated CTA. Consequently, the dominated CTA should not be chosen. Under FSD, investors will exhibit non-satiation (more is preferred to less); under SSD, investors will have additional characteristic of risk aversion while under TSD they have added decreasing absolute risk aversion (DARA). We note that hierarchical relationship exists in SD (Levy 1992, 1998). This means FSD implies SSD, which in turn implies TSD. However, the converse is not true. Thus, only the lowest dominance order of SD is reported in practice. Wong and Ma (2008) showed that SD criteria also apply for a range of non-expected utility theories of choice under uncertainty.

Recent advances in SD techniques allow the statistical significance of SD to be determined. To date, the SD tests have been well developed, for example, see McFadden (1989), Kaur et al. (1994), Anderson (1996, 2004), Davidson and Duclos (DD, 2000), Barrett and Donald (BD, 2003) and Linton et al. (LMW, 2005). The DD test has been found to be one of the most powerful, but yet less conservative in size (Wei and Zhang, 2003; Tse and Zhang, 2004; Lean et al., 2008); while the BD test is another powerful test instrument and the LMW is useful as it is extended from Kolmogorov-Smirnov test for FSD and SSD by relaxing the iid assumption. We report the results of DD test and skip reporting those of BD and LMW tests as the former is the only SD statistics that test the SD relationship up to the third-order and the results of both BD and LMW tests are consistent with those of the DD test.

For any two CTAs *Y* and *Z* with CDFs *F* and *G* respectively and for a grid of pre-selected points $x_1, x_2 \ldots x_k$, the order-*j* DD statistic, $T_j\ (x)$ (j = 1, 2, and 3), is:

$$T_j\ (x) = \frac{\hat{F}_j\ (x) - \hat{G}_j\ (x)}{\sqrt{\hat{V}_j\ (x)}} \qquad (2)$$

where $\hat{V}_j(x) = \hat{V}_Y^j(x) + \hat{V}_Z^j(x) - 2\hat{V}_{Y,Z}^j(x),$

$$\hat{H}_j\ (x) = \frac{1}{N(j-1)!}\sum_{i=1}^{N}(x-h_i)_+^{j-1},$$

$$\hat{V}_H^j(x) = \frac{1}{N}\left[\frac{1}{N((j-1)!)^2}\sum_{i=1}^{N}(x-h_i)_+^{2(j-1)} - \hat{H}_j\ (x)^2\right], H = F, G; h = y, z;$$

$$\hat{V}_{Y,Z}^j(x) = \frac{1}{N}\left[\frac{1}{N((j-1)!)^2}\sum_{i=1}^{N}(x-y_i)_+^{j-1}\left(x-z_i\right)_+^{j-1} - \hat{F}_j\ (x)\hat{G}_j\ (x)\right]$$

in which F_j and G_j are defined in (1).

It is empirically impossible to test the null hypothesis for the full support of the distributions. Thus, Bishop et al. (1992) proposed to test the null hypothesis for a pre-designed finite numbers of values *x*. Specifically, the following hypotheses are tested:

$H_0 : F_j\ (x_i) = G_j\ (x_i)$, for all x_i, $i = 1, 2, ..., k$;

$H_A : F_j\ (x_i) \neq G_j\ (x_i)$ for some x_i ;

$H_{A1} : F_j\ \left(x_i\right) \leq G_j\ \left(x_i\right)$ for all $x_i, F_j\ \left(x_i\right) < G_j\ \left(x_i\right)$ for some x_i;

$H_{A2} : F_j\ \left(x_i\right) \geq G_j\ \left(x_i\right)$ for all $x_i, F_j\ \left(x_i\right) > G_j\ \left(x_i\right)$ for some x_i.

We note that in the above hypotheses, H_A is set to be exclusive of both H_{A1} and H_{A2}, which means that if either H_{A1} or H_{A2} is accepted, this does not mean that H_A is accepted. Under the null hypothesis, DD showed that $T_j\ \left(x\right)$ is asymptotically distributed as the Studentized Maximum Modulus (SMM) distribution (Richmond,

1982) to account for joint test size. To implement the DD test, the test statistic at each grid point is computed and the null hypothesis is rejected if the test statistic is significant at any grid point. The SMM distribution with k and infinite degrees of freedom, denoted by $M^k_{\infty,\alpha}$, is used to control for the probability of rejecting the overall null hypothesis. The following decision rules are adopted based on 1-α percentile of $M^k_{\infty,\alpha}$ tabulated by Stoline and Ury (1979):

If $|T_j(x_i)| < M^k_{\infty,\alpha}$ for $i = 1,\ldots,k$, accept H_0;
if $T_j(x_i) < M^k_{\infty,\alpha}$ for all i and $-T_j(x_i) > M^k_{\infty,\alpha}$ for some i, accept H_{A1};
if $-T_j(x_i) < M^k_{\infty,\alpha}$ for all i and $T_j(x_i) > M^k_{\infty,\alpha}$ for some i, accept H_{A2}; and
if $T_j(x_i) > M^k_{\infty,\alpha}$ for some i and $-T_j(x_i) > M^k_{\infty,\alpha}$ for some i, accept H_A.

Accepting either H_0 or H_A implies non-existence of any SD relationship, non-existence of any arbitrage opportunity between these two CTAs and neither of these two CTAs are preferred to one another. However, if H_{A1} or H_{A2} of order one is accepted, a particular CTA stochastically dominates another CTA at first-order. In this situation, any non-satiated investor will be better off if s/he switches from the dominated CTA to the dominant one. On the other hand, if H_{A1} or H_{A2} is accepted for order two or three, a particular CTA stochastically dominates the other at second- or third-order. In this situation, arbitrage opportunity does not exist and switching from one CTA to another will only increase investors' expected utilities, but not wealth (Jarrow, 1986; Falk and Levy, 1989).

The DD test compares the distributions at a finite number of grid points. Various studies examined the choice of grid points. For example, Tse and Zhang (2004) showed that an appropriate choice of k for reasonably large samples ranges from 6 to 15. Too few grids will miss information of the distributions between any two consecutive grids (Barrett and Donald, 2003) and too many grids will violate the independence assumption required by the SMM distribution (Richmond, 1982). To make more detailed comparisons without violating the independence assumption, we follow Fong et al. (2005), Lean et al. (2007), and Gasbarro et al. (2007) to make 10 major partitions with 10 minor partitions within any two consecutive major partitions in each comparison and to make the statistical inference based on the SMM distribution for k =10 and infinite degrees of freedom. Lean et al (2008) explained the choice of this methodology. Critical values are: 3.691, 3.254 and 3.043 for 1%, 5% and 10% significance levels tabulated in Stoline and Ury (1979). This allows the examination of the consistency of both magnitudes and signs of the DD statistics between any two consecutive major partitions without violating the independent assumption.

3.2. Market Efficiency and Arbitrage Opportunity

Without identifying any risk index or any specific model, the SD rules can be used to determine if arbitrage opportunities exist, and if the markets are efficient. In examining market data, the criteria that SD employs are: (a) Can investors switch their portfolio choice, say from Y to Z and increase their (expected) wealth? (b) Can some investors switch their investment choice, say from Y to Z and increase their expected utilities?

In the market efficiency hypothesis, if one is able to earn an abnormal return, the market is considered inefficient. Market efficiency can be tested using SD rules as follows: If investors can switch their asset choice and increase their expected wealth, independent of their specific preferences, if market data shows that investors can benefit, then market inefficiency is implied. Jarrow (1986) and Falk and Levy (1989) claimed that if FSD exists, under certain conditions, arbitrage opportunities also exist, and investors will increase their wealth and expected utilities if they shift from holding the dominated asset to the dominant one. However, Wong et al. (2008) showed that if FSD exists statistically, arbitrage opportunities may not exist, but investors can increase their expected wealth as well as their expected utilities if they shift from holding the dominated asset to the dominant one.

In addition, if the market is not 'complete,' even if FSD exists, investors may not be able to exploit any arbitrage opportunities. Also, if the test detects FSD of a particular CTA over another but the dominance only lasts for a short period; the results cannot be used to reject market efficiency. In general, the FSD should not last for a long period of time because market forces induce adjustments to a condition of no FSD if the market is efficient. For example, if Y dominates Z at FSD, then investors would buy Y and sell Z. This will continue, driving up the price of Y relative to Z until the market price of Y relative to Z is high enough to make the marginal investor indifferent between both CTAs. If the FSD does not last for a long period of time, we infer that the market is still efficient.

If the FSD holds for a long time and all investors increase their expected wealth by switching their asset choice, then, we claim that the market is inefficient. Another possibility for the existence of FSD to be held for a long period is that investors do not realize that such dominance exists. It would be interesting to investigate whether FSD relationships among some CTAs disappear over time. If they do not, then this would be considered a financial anomaly.

4. RESULTS

The means and the standard deviations of the returns for all 56 CTAs studied in this paper are plotted in Figure 1. From Figure 1, we find that in general the means and standard deviations move together and thus the results are consistent with modern portfolio theory that higher mean accompanies with higher risk. We also plot the risks versus returns and the corresponding efficient frontier for the 56 CTAs in Figure 2. We find that most of the CTAs are not on the efficient frontier.

Insert Figure 1 and 2 here

Summary statistics of all five indices and the five CTAs with largest or smallest means or standard deviations are provided in Table 1. The five individual CTAs that are summarized in Table 1 differ in investment locations: CTA12 is North America & Asia; CTA13, CTA17 and CTA32 are North America and CTA56 is Asia.

Insert Table 1 here

From Table 1, the average mean and standard deviation of monthly returns of the 56 CTAs are higher than those of the five market indices. These results infer that CTAs generate higher returns with higher risk than equities. In addition, as shown in Table 1, the means and standard deviations vary widely across CTAs. For example, CTA13 possesses the largest monthly mean return (2.0662) and CTA17 possesses the largest standard deviation (12.7135) while CTA32 exhibits the lowest monthly mean returns (0.4617) as well as the smallest standard deviation (0.7229). We run paired t-test and find that CTA13 do not dominate the other four. Therefore, we comment that a fund with the largest mean returns may not be a good investment choice under MV criterion. We also find that no CTA dominate each other by MV rule among the five CTAs.

Insert Table 2 here

Next, we turn to investigate the CAPM measures. All betas are less than one except CTA13, ranging from -0.7416 to 1.2111 and all Sharpe ratios are negative. CTA13 exhibits the largest Sharpe ratio (-0.1772) while CTA32 has the smallest (-4.7599). Furthermore, CTA01 possesses the highest Treynor (994.92) while CTA13 has the highest Jensen (2.1546) measures. A summary of dominance results among the five CTAs measured by MV and CAPM statistics are presented in Table 2. We find that different CAPM measures draw different favourable CTAs, for example CTA13 dominates CTA17 by Sharpe ratio and Jensen index while CTA17 dominates CTA12 by Sharpe ratio and Treynor index.

We also observe that a CTA dominates another CTA by a CAPM statistic but the dominance relation can be reverse if measured by different CAPM statistic(s). For example, CTA12 dominates CTA56 by Sharpe ratios and Jensen index but it is dominated by CTA56 when Treynor index is used. Only CTA12 and CTA13 dominate CTA32 (with the smallest mean, standard deviation and Sharpe ratio) by all the three CAPM statistics. Nonetheless, our results show that some of the return distributions are non-normal and exhibit both negative skewness and excess kurtosis. Specifically, 26 skewness, 29 kurtosis and 31 Jarque-Bera measures are significant at the 0.05 level, highlighting the non-normality feature for the CTAs returns.

Hence, we deduce that the modern portfolio theory is too simplistic to deal with CTA as noted by Kat (2003) and Kooli et al. (2005) in their analysis of hedge funds. Furthermore, CAPM measures may overestimate and miscalculate CTA's performance. As the results drawn by both MV and CAPM statistics could be misleading, we recommend applying SD criterion as the alternative comparison in this paper.

DD stated that the null hypothesis of equal distribution could be rejected if any value of the test statistic, T_j , is significant (see equation 2). In order to minimize the Type II error and to accommodate the effect of almost SD (Leshno and Levy, 2002), we follow Fong et al. (2005, 2008), Lean et al. (2007, 2010) and Gasbarro et al. (2007) to use a conservative 5% cut-off point for the proportion of test statistics in statistical inference. Using a 5% cut-off point, if we find at least 5% of T_j is significantly negative and no portion of T_j is significantly positive then CTA *Y* dominates CTA *Z*. The reverse holds if the CTA *Z* dominates CTA *Y*.

Insert Table 3 and 4 here

Table 3 shows the results of the DD test for the pairwise comparison of the five market indices and the five 'most outstanding' CTAs. Table 4 summarizes the DD test results for those with other CTAs. From the table, we find that there are some FSD among the indices/CTAs, for example CTA12 dominates 3 other indices/CTAs and is dominated by the other two at first-order. This infers that the non-satiation investors can increase their expected wealth and expected utilities if they shift from holding the 3 dominated indices/CTAs to CTA12 or from CTA12 to the other two. In other words there exists arbitrage opportunity under certain conditions as claimed by Jarrow (1986) and Falk and Levy (1989) for investors who are holding this type of portfolio.

Market indices are dominated by 2 – 5 CTAs at first-order. Risk averters would prefer CTA32 most as it dominates fifty five other indices/CTAs and not dominated by others at second-order. On the other hand, CTA17 and CTA13 are the two CTAs that are less preferred by risk-averse investors as they are second-order dominated by thirty five and twenty nine other indices/CTAs respectively.

Insert Table 5 here

We apply equation (2) with the preferable CTA being the first variable (*F*) and the less preferable CTA being the second variable (*G*). If the results are as expected, there will exist some significantly negative T_j, j = 1, 2, 3 with no significant positive T_j. For example, as investors are non-satiated, we presume CTA with the highest mean (CTA13) will be preferred to the CTA with the smallest mean (CTA32). Taking CTA13 as the first variable and CTA32 as the second variable in equation (2), the DD results in Table 5 show that there are 22 (32) percentage of T_1 to be significantly positive (negative). This result shows that CTA13 and CTA32 do not dominate each other at first order. However, we observe that all values of $T_2(T_3)$ are non-negative with 22 (25) percent of $T_2(T_3)$ being significantly positive. Surprisingly, contradicting common belief, an asset with the smallest mean SSD (TSD) an asset with the largest mean. Hence, we deduce that risk averters and risk-averse investors with DARA who make their portfolio choice on the basis of expected-utility maximization will increase their expected utilities by shifting from CTA13 to CTA32.

Insert Figure 3 here

We recall that the MV and CAPM measures show that CTA32 does not dominate CTA13 whereas CTA13 dominates CTA32 by Sharpe Ratio, Jensen index and Treynor index. As CTA13 possesses an insignificantly larger mean but significantly larger standard deviation than CTA32, one should not be surprised that our SD results reveal that CTA32 dominates CTA13 at second- and third-order. This result is consistent with Markowitz (1991) that investors, especially risk-averse investors, worry more about downside risk than upside profit. In addition, together with Figure 3, the results from Table 5 show that 22% of T_1 is significantly positive in the negative domain whereas 32% of T_1 is significantly negative in the positive domain. All these SD results imply that CTA13 and CTA32 do not outperform each other. CTA32 is preferable in the negative domain whereas CTA13 is preferable in the positive domain and, overall, risk averters prefer to invest in CTA32 than CTA13. This result cannot be obtained using MV or CAPM measures.

The traditional measures by comparing a number of assets can only tell investors which asset has performed better under restrictive assumptions. Sometimes these statistics are ambiguous and fail to provide detailed information on neither the dominance relationship nor the preferences of investors. The SD approach is not only assumption free, but also more informative allowing for useful economic interpretation of the performance and risk inherent in a CTA track record.

We note that most of the SD comparisons for assets in the literature stop at answering the question if one asset SSD or TSD another asset (Seyhun, 1993). By applying the DD test, we can also answer the question: if one asset dominates another asset on the downside while the reverse dominance relationship can be found on the upside. This question is in line with the direction of research of Post and Levy (2005) who investigate the behaviors of investors in bull and bear markets.

Insert Figure 4 and 5 here

We further examine the pairwise SD relationship between CTAs of different location focus. We report the results in Table 5 and Figure 4 and 5. We find that only CTA32 with North America focus SSD and TSD both the North America & Asia (CTA12) and Asia (CTA56) focuses. On the other hand, CTA13 and CTA17 with North America strategies are stochastically dominated by both the North America & Asia (CTA12) and Asia (CTA56) strategies. Our results show that location focus is not an important factor for CTA performance.

5. CONCLUSION

This paper introduces an alternative SD test, which is basically assumption free to investigate the characteristics of the entire distribution of returns and test whether rational investors benefit from selecting CTAs to maximize their expected utilities and/or expected wealth. An advantage of this approach is that it alleviates the problems that can arise if CTA returns are skewed and leptokurtic and non-normally distributed. Our approach also allows for a meaningful economic interpretation of the results. The economic interpretation and findings provide useful guides to investors, especially given the relative liquidity and transparency of CTAs compared with other alternative investments.

Based on a sample of 56 individual CTAs, FSD relationship do exist in some CTAs. We also find the existence of the SSD relationship among other CTAs/indices; indicating that the non-satiation and risk-averse investors would maximize their expected utilities, but not their expected wealth by switching from the SSD dominated CTAs to their corresponding SSD dominant ones.

Some authors propose to use higher order (higher than three) SD in empirical application. For example, Vinod (2004) recommended employing the 4th order SD to choose investment prospects amongst 1281 mutual funds. We, however, would like to note that the first three orders are the most commonly-used orders in empirical work on SD, regardless whether the analyses are simple or complicated. We would also like to note that a hierarchy exists in SD relationships whereby findings of the FSD implies the SSD which in turn implies the TSD and the fourth order SD and so on (Levy 1992, 1998). We thus stopped at third order in this paper. We also note that Post and Versijp (2006) developed a new SD test for multiple comparisons recently. It will be an interesting future research to extend to the multiple SD comparison for CTAs.

Table 1: Summary Statistics of CTAs and Five Market Indices

	Mean	Std Dev	Sharpe	Skewness	Kurtosis
S&P 500 Index (IX1)	1.05658	4.51646	-0.63011	-0.64987**	0.47889
MSCI World Index (IX2)	0.60727	4.17056	-0.7901	-0.75939**	0.74925
Goldman Sachs Commodity Index (IX3)	0.61591	5.77761	-0.56884	0.18455	0.46225
Lehman Global Gag. US Universal (IX4)	0.08450	1.10173	-3.4654	-0.44012	1.21000*
Lehman US Universal: High Yield Corp. (IX5)	0.17936	2.27861	-1.63393	0.07035	3.31718**
Average (CTA)	1.19307	5.84364	-0.64177	0.58034	2.24621
Maximum (CTA)	2.06617	12.7135	-0.17724	4.57608**	30.3344**
Minimum (CTA)	0.46167	0.72287	-4.75990	-0.6357**	-0.3890
CTA12 Concepts Currency Fund Ltd (DMC)	1.05308	4.52548	-0.62963	0.78362**	0.98745
CTA13 Red Oak Commodity Advisors	**2.06617**	10.3606	**-0.17724**	0.35809	0.17709
CTA17 Legacy Futures Fund LP	1.63958	**12.7135**	-0.17799	1.77650**	9.09093**
CTA32 Worldwide Financial Futures Program	**0.46167**	**0.72287**	**-4.7599**	0.06434	1.15461*
CTA56 Grinham Diversified Program	0.92183	3.13036	-0.95216	0.15210	0.98950

Note: CTA13, CTA17, and CTA32 are the 'most outstanding funds' in which CTA13 possesses the largest monthly mean return (2.06617) and the largest Sharpe ratio (-0.17724); CTA17 has largest standard deviation (12.7135); CTA32 exhibits the lowest monthly mean return (0.46167), the smallest standard deviation (0.72287), and the smallest Sharpe ratio (-4.7599). CTA12 and CTA56 are included because they represent from different investment location than the three. Results in bold are the extreme values. * $p < 0.05$, ** $p < 0.01$.

Table 2: Pairwise Comparison between CTA by the MV and CAPM measures

	CTA12	CTA13	CTA17	CTA32	CTA56
CTA12		N	J	S, T, J	S, J
CTA13	S, T, J		S, J	S, T, J	S, J
CTA17	S, T	T		S, T	S, J
CTA32	N	N	J		J
CTA56	T	T	T	S, T	

Note: M, S, T, and J indicate dominance by MV criterion, Sharpe ratio, Treynor index, and Jensen index, respectively. N denotes no dominance by MV, Sharpe ratio, Treynor index, and Jensen index. In the table, the rows indicate whether the CTA in the leftmost column dominates any of the CTAs in the top row while the columns show whether the CTA in the top row is being dominated by any of the CTAs in the leftmost column. For example, the cells in the first row (CTA12) and the forth column (CTA32) means that CTA12 dominates CTA32 by Sharpe ratio, Treynor index, and Jensen index. The five CTAs are defined in Table 1.

Table 3: Pairwise Comparison between CTA by the Davidson-Duclos (DD) tests

	IX1	IX2	IX3	IX4	IX5	CTA12	CTA13	CTA17	CTA32	CTA56	Dominates
IX1		ND	ND	ND	ND	ND	SSD	SSD	ND	ND	2
IX2	ND		ND	ND	ND	ND	SSD	SSD	ND	ND	2
IX3	ND	ND		ND	ND	ND	ND	ND	ND	ND	0
IX4	SSD	SSD	SSD		ND	SSD	SSD	SSD	ND	ND	6
IX5	SSD	SSD	SSD	ND		ND	SSD	SSD	ND	ND	5
CTA12	ND	ND	ND	ND	ND		SSD	SSD	ND	ND	2
CTA13	ND	ND	ND	ND	ND	ND		ND	ND	ND	0
CTA17	ND	ND	ND	ND	ND	ND	ND		ND	ND	0
CTA32	SSD	SSD	SSD	FSD	SSD	SSD	SSD	SSD		SSD	9
CTA56	ND	ND	FSD	ND	ND	ND	SSD	SSD	ND		3
Dominated by	3	3	4	1	1	2	7	7	0	1	

Notes: The results in this Table are read based on row versus column. For example, the cell in the forth row IX4 and the first column IX1 tell us that IX4 stochastically dominates IX1 at second-order while the cell in the second row IX2 and the first column IX1 means that IX2 does not stochastically dominate IX1. Alternatively, reading along the row IX1, it can be seen that IX1 dominates 2 other indices/CTAs while reading down the IX1 column shows that IX1 is dominated by 3 other indices/CTAs. The five indices and the five CTAs are defined in Table 1.

Table 4: Summary of the Davidson-Duclos (DD) Test Statistics

	Dominates			Dominated By		
Index / Fund	FSD	SSD	Total	FSD	SSD	Total
IX1	0	5	5	2	4	6
IX2	0	5	5	2	4	6
IX3	0	2	2	5	8	13
IX4	0	47	47	5	0	5
IX5	0	28	28	3	1	4
CTA12	3	7	10	2	3	5
CTA13	0	0	0	1	29	30
CTA17	0	0	0	2	35	37
CTA32	1	55	56	1	0	1
CTA56	1	23	24	0	2	2

Notes: The values indicate the number of indices/funds for each index/fund dominates or the number of indices/funds that it is dominated by. Note that in the table the reported number of SSD excludes the number of FSD. As hierarchical relationship exists in SD, FSD implies SSD. Thus, the total number is the sum of FSD and SSD (exclusive of FSD). For example, IX1 not FSD any others but SSD 5 other indices/CTAs. Thus, it dominates 5 indices/funds (including both FSD and SSD) totally. IX1 is dominated by 2 other indices/CTAs at first-order and 4 other indices/CTAs at second-order. Thus it is dominated by 6 indices/funds in total.

Table 5: Results of Davidson-Duclos (DD) Test for Risk Averters

Sample	FSD		SSD		TSD	
	$\%T_1>0$	$\%T_1<0$	$\%T_2>0$	$\%T_2<0$	$\%T_3>0$	$\%T_3<0$
CTA13 - CTA12	17	25	**19**	**0**	**21**	**0**
CTA17 - CTA12	10	11	**12**	**0**	**14**	**0**
CTA13 - CTA32	22	32	**22**	**0**	**25**	**0**
CTA13 - CTA56	19	29	**20**	**0**	**23**	**0**
CTA17 - CTA56	12	13	**13**	**0**	**16**	**0**
CTA32 - CTA12	26	18	**0**	**21**	**0**	**28**
CTA32 - CTA17	15	13	**0**	**12**	**0**	**15**
CTA32 - CTA56	24	20	**0**	**15**	**0**	**8**
CTA12 - CTA56	0	0	0	0	0	0

Note: DD test statistics are computed over a grid of 100 on monthly CTA returns. The table reports the percentage of DD statistics which are significantly negative or positive at the 5% significance level, based on the asymptotic critical value of 3.254 of the studentized maximum modulus (SMM) distribution. T_j is the Davidson and Duclos (DD) statistic for risk averters with *j =1, 2,* and *3* defined in equation (2) with F to be the first fund and G to be the second fund stated in the first column. The five CTAs are defined in Table 1.

Figure 1: Means and Standard Deviations of 56 CTA

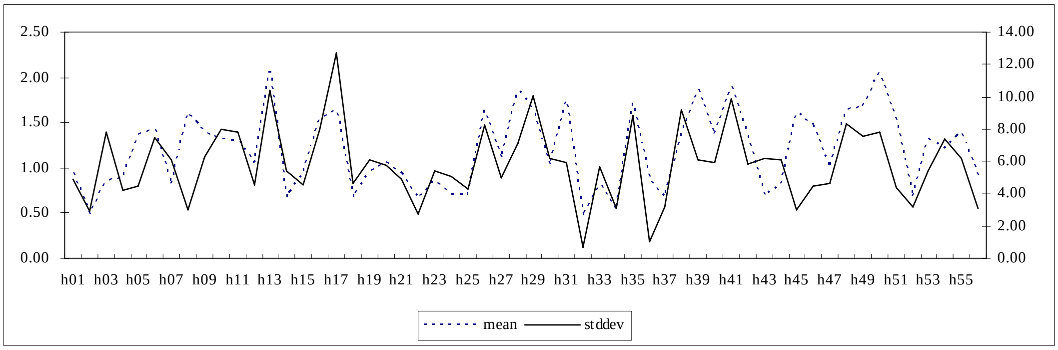
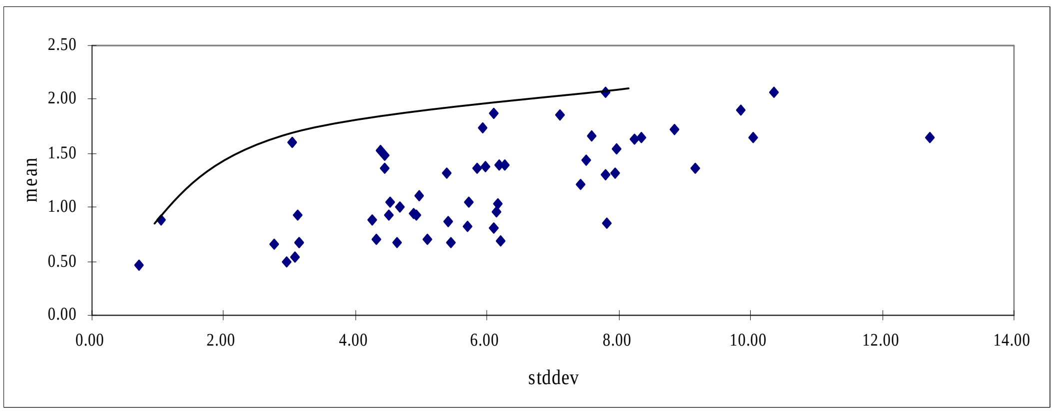


Figure 2: Plot of Risk vs. Returns of 56 CTA



efficient frontier

Figure 3: DD Statistics of CTA13 – CTA32 and their Cumulative Distribution Functions

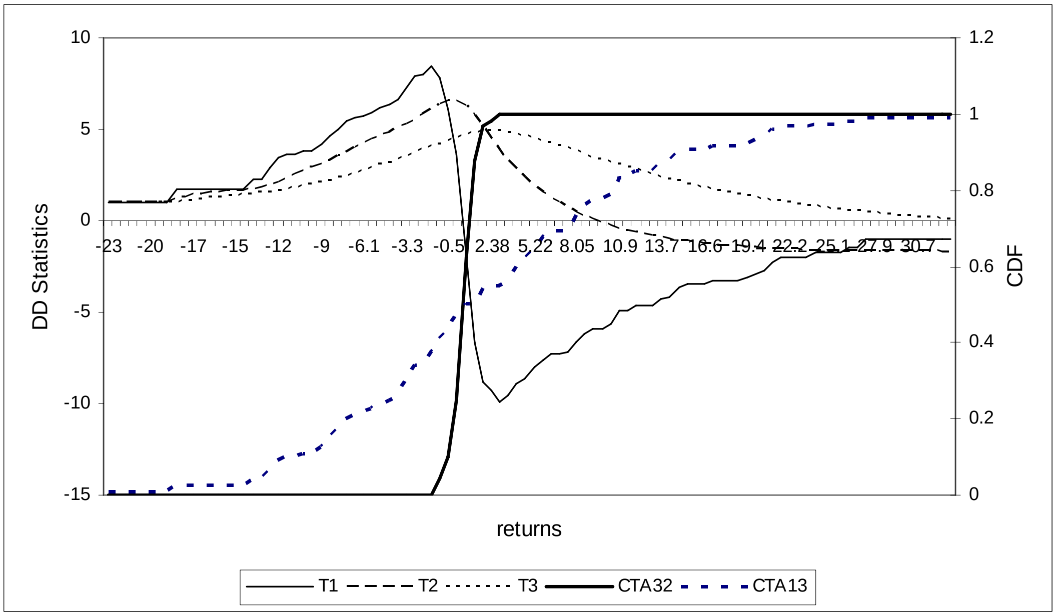


Figure 4: DD Statistics of CTA32 – CTA12 and their Cumulative Distribution Functions

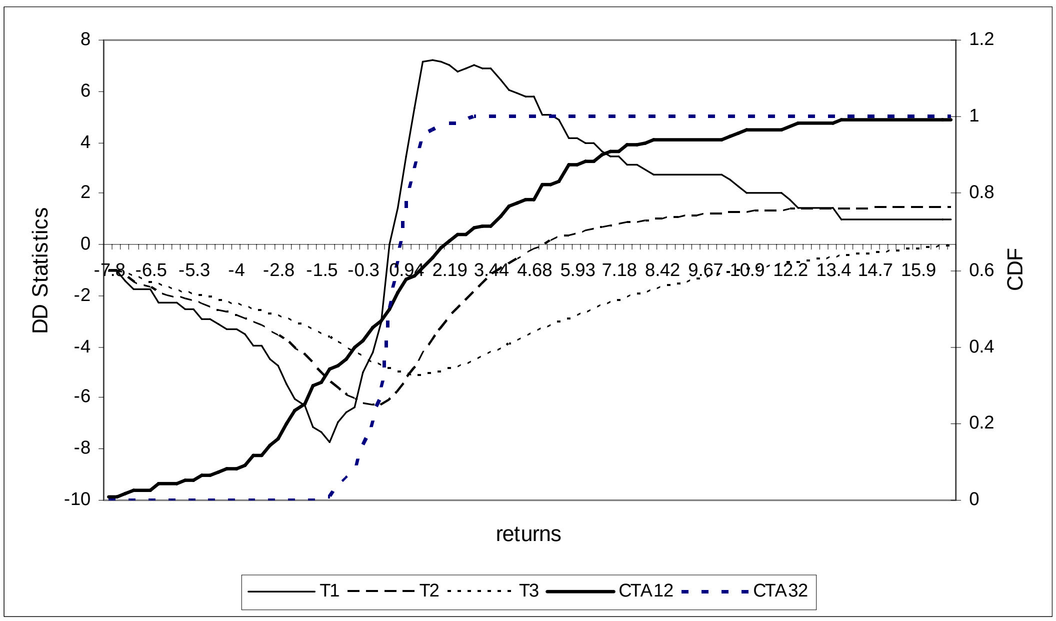
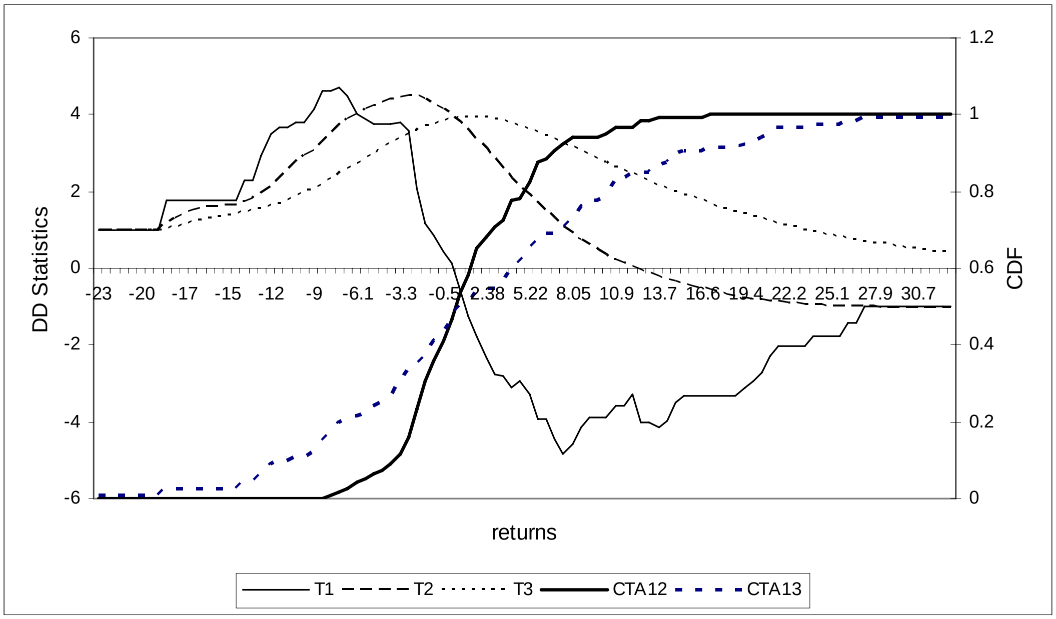


Figure 5: DD Statistics of CTA13 – CTA12 and their Cumulative Distribution Functions



REFERENCES

Anderson, G. (1996), "Nonparametric Tests of Stochastic Dominance in Income Distributions", *Econometrica*, Vol. 64, pp.1183 – 1193.

Anderson, G. (2004), "Toward an Empirical Analysis of Polarization", *Journal of Econometrics*, Vol. 122, pp. 1–26.

Bai, Z.D., Liu, H.X. and Wong, W.K. (2009), "Enhancement of the Applicability of Markowitz's Portfolio Optimization by Utilizing Random Matrix Theory", *Mathematical Finance*, Vol 19, pp. 639-667.

Bai, Z.D., Liu, H.X. and Wong, W.K. (2011a), "Asymptotic Properties of Eigenmatrices of A Large Sample Covariance Matrix", *Annals of Applied Probability*, (forthcoming).

Bai, Z.D., Wang, K.Y. and Wong, W.K. (2011b), "An Improvement of the Sharpe-Ratio Test on Small Samples -- Mean-Variance Ratio Test", *Statistics and Probability Letters*, (forthcoming).

Barrett, G.F. and Donald, S.G. (2003), "Consistent Tests for Stochastic Dominance", *Econometrica*, Vol. 71, pp. 71–104.

Bishop, J.A., Formly, J.P. and Thistle, P.D. (1992), "Convergence of the South and Non-South Income Distributions", *American Economic Review*, Vol.82, pp.262–272.

Collins, D.P. (2005), "How to Select a CTA", *Futures: News, Analysis & Strategies for Futures, Option and Derivatives Traders*, pp.70-74.

Davidson, R., and Duclos, J.Y. (2000), "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality", *Econometrica*, Vol.68, pp.1435–1464.

Falk, H., and Levy, H. (1989), "Market Reaction to Quarterly Earnings' Announcements: A Stochastic Dominance based Test of Market Efficiency", *Management Science*, Vol.35, pp.425–446.

Fong, W.M., Wong, W.K. and Lean, H.H. (2005), "International Momentum Strategies: A Stochastic Dominance Approach", *Journal of Financial Markets*, Vol.8, pp.89–109.

Fung, W. and Hsieh, D.A. (2000), "Performance Characteristics of Hedge Funds and Commodity Funds: Natural vs. Spurious Biases, Journal of Financial and Quantitative Analysis", Vol.35, pp. 291–307.

Gasbarro, D., Wong, W.K. and Zumwalt, J.K. (2007), "Stochastic Dominance Analysis of iShares", *European Journal of Finance*, Vol.13, pp. 89–101.

Jarrow, R. (1986), "The Relationship between Arbitrage and First Order Stochastic Dominance", *Journal of Finance*, Vol. 41, pp.915–921.

Jensen, M.C. (1969), "Risk, The Pricing of Capital Assets and The Evaluation of Investment Portfolios", *Journal of Business*, Vol.42, pp.167–247.

Kat, H.M. (2003), "10 Things that Investors should Know about Hedge Funds", *Journal of Wealth Management*, Vol.5, pp. 72–81.

Kat, H.M. (2004a), "Managed Futures and Hedge Funds: A Match Made in Heaven". *The Journal of Investment Management*, Vol.2, pp. 32-40.

Kat, H.M. (2004b), “In Search of the Optimal Fund of Hedge Funds”. *The Journal of Wealth Management*, Vol.6, pp. 43–51.

Kat, H.M. and Menexe, F. (2003), “Persistence in Hedge Fund Performance: The True Value of a Track Record”, *Journal of Alternative Investments,* Vol. 5, pp. 62–72.

Kaur, A., Rao, B.L. and Singh, H. (1994), “Testing for Second-Order Stochastic Dominance of Two Distributions”, *Econometric Theory,* Vol.10, pp. 849–866.

Kjetsaa, R. and Kieff, M. (2003), “Stochastic Dominance Analysis of Equity Mutual Fund Performance”, *American Business Review*, Vol. 21, pp. 1–8.

Kooli, M., Amvella, S.P. and Gueyie, J.P. (2005), “Hedge Funds in a Portfolio Context: A Mean-Modified Value-at-Risk Framework”, *Derivatives Use, Trading & Regulation*, Vol.10, pp. 373–383.

Lean, H.H., Smyth, R. and Wong, W.K. (2007) “Revisiting Calendar Anomalies in Asian Stock Markets Using a Stochastic Dominance Approach”, *Journal of Multinational Financial Management,* Vol. 17, pp. 125–141.

Lean, H.H., Wong, W.K. and Zhang, X.B. (2008) The Size and Power of Some Stochastic Dominance Tests: A Monte Carlo Study for Correlated and Heteroskedastic Distributions”. *Mathematics and Computer in Simulation*Vol.79, pp.30-48.

Lee, D., Koh, F. and Phoon, K.F. (2004), “CTA Strategies for Returns-Enhancing Diversification”, *in Gregoriou, Karavas, Lhabitant and Rouah, (eds), Commodity Trading Advisors: Risk, Performance Analysis and Selection, Wiley, New York.*

Lee, D., Phoon., K.F. and Wong, C.Y. (2006), “Moments Analysis in Risk and Performance Measurement”, *Journal of Wealth Management*, Vol. 9, pp. 54–65.

Leshno, M. and Levy, H. (2002), “Preferred by “All” and Preferred by “Most” Decision Makers: Almost Stochastic Dominance”, *Management Science*, Vol.48, pp.1074–1085.

Leung, P.L. and Wong, W.K. (2008), “On testing the equality of the multiple Sharpe Ratios, with application on the evaluation of iShares”, *Journal of Risk*, Vol 10, pp. 1-16.

Levy H. (1992), “Stochastic Dominance and Expected Utility: Survey and Analysis”, *Management Science*, Vol.38, pp. 555–593.

Levy, H. (1998), “Stochastic Dominance: Investment Decision Making under Uncertainty”, *Kluwer, Boston.*

Levy, H. and Sarnat, M. (1970), “Alternative Efficiency Criteria: An Empirical Analysis”, *Journal of Finance*, Vol.25, pp. 1153–1158.

Linton, O., Maasoumi, E. and Whang, J.Y. (2005), “Consistent Testing for Stochastic Dominance under General Sampling Schemes”, *Review of Economic Studies*, Vol.72, pp. 735–765.

Markowitz, H.M. (1952), “Portfolio Selection”, *Journal of Finance*, Vol.7, pp. 77–91.

Markowitz, H.M. (1991), “Foundations of Portfolio Theory,” Les Prix Nobel 1990, 292 (Nobel Foundation, Stockholm).

McFadden, D. (1989), “Testing for Stochastic Dominance”, *in Fomby, T., and Seo, T. (eds), Studies in the Economics of Uncertainty, Springer Verlag, New York, 113–132.*

Porter, R.B. (1973), “An Empirical Comparison of Stochastic Dominance and Mean-Variance Portfolio Choice Criteria”, *Journal of Financial and Quantitative Analysis*, Vol.8, pp. 587–608.

Post, T. and Levy, H. (2005), “Does Risk Loving Drive Asset Prices?”, *Review of Financial Studies*, Vol.18, pp. 925-953.

Post, T. and Versijp, P. (2006), “Multivariate Tests for Stochastic Dominance Efficiency of A Given Portfolio”, *Journal of Financial and Quantitative Analysis*, Vol.42, pp. 489-516.

Richmond, J. (1982), “A General Method for Constructing Simultaneous Confidence Intervals”, *Journal of the American Statistical Association*, Vol.77, pp. 455-460.

Seyhun, H.N. (1993), “Can Omitted Risk Factors Explain the January Effect? A Stochastic Dominance Approach”, *Journal of Finance and Quantitative Analysis*, Vol. 28, pp. 195-212.

Sharpe, W.F. (1964), “Capital Asset Prices: Theory of Market Equilibrium under Conditions of Risk”, *Journal of Finance*, Vol.19, pp. 452–442.

Stoline, M.R. and Ury, H.K. (1979), “Tables of the Studentized Maximum Modulus Distribution and An Application to Multiple Comparisons among Means”, *Technometrics*, Vol.21, pp.87–93.

Taylor, W.R. and Yoder, Y.A. (1999), “Load and No-load Mutual Fund Dynamics during the 1987 Market Crash: A Stochastic Dominance Analysis”, *Journal of Economics and Finance*, Vol. 23, pp. 255–265.

Treynor, J.L. (1965), "How to Rate Management of Investment Funds", *Harvard Business Review*, Vol. 43, pp. 63–75.

Tse, Y.K. and Zhang, X. (2004), "A Monte Carlo Investigation of Some Tests for Stochastic Dominance", *Journal of Statistical Computation and Simulation*, Vol. 74, pp. 361–378.

Vinod, H.D. (2004), "Ranking Mutual Funds using Unconventional Utility Theory and Stochastic Dominance", *Journal of Empirical Finance*, Vol.11, pp. 353–377.

Vuille, S. and Crisan, C. (2004), "A Quantitative Analysis of CTA Funds", *SSRN Working Paper No. 623261*.

Wei, S. and Zhang, C. (2003), "Statistical and Economic Significance of Stock Return Predictability: A Mean-Variance Analysis", *Journal of Multinational Financial Management*, Vol.13, pp. 443–463.

Wong, W.K. (2007), "Stochastic dominance and mean-variance measures of profit and loss for business planning and investment", *European Journal of Operational Research*, Vol 182, pp. 829-843.

Wong, W.K. and Chan, R.H. (2008), "Prospect and Markowitz Stochastic Dominance", *Annals of Finance*, Vol.4, pp. 105-129.

Wong, W.K. and Li, C.K. (1999), "A Note on Convex Stochastic Dominance Theory", *Economic Letters*, Vol.62, pp. 293-300.

Wong, W.K. and Ma, C. (2008), "Preferences over Meyer's Location-Scale Family".

Economic Theory. Vol 37, pp. 119-146.