Comparisons of
Commodity and Managed Futures Benchmark Indices

Thomas Schneeweis*
Richard Spurgin**

CISDM Working Paper Series
Original: July 8, 1995
Revised: August 2, 1996

*Professor of Finance, University of Massachusetts at Amherst
**Assistant Professor of Finance, Clark University, Worcester, Mass.

Please Address Correspondence to:
Thomas Schneeweis
CISDM/School of Management
University of Massachusetts
Amherst, Massachusetts 01003

Phone: (413) - 545-5641
Fax: (413) - 545-3858
Email: Schneeweis@UMASS.SOM.EDU

This research is supported by a grant from the Chicago Board of Trade and by the University of Massachusetts Center for International Security and Derivative Markets (CISDM). The results of this study reflect only the views of the author(s) and not necessarily the institutions with which the author(s) are affiliated.

Comparisons of
Commodity and Managed Futures Benchmark Indices

Abstract

Benchmark indices are commonly used in studies of investment performance to provide a performance index which reflects the particular style of an investment manager. While benchmark indices are common in the areas of stock and bond investment, in recent years futures and options based commodity trading products have grown rapidly as alternative means of investing in cash market commodity and financial securities. Various commodity and futures market based indices have existed for years (e.g., CRB, Dow Jones futures index). However, indices based on the performance of commodity futures traders (e.g., MAR, Barclay, TASS indices) have also recently been used in discussions of benchmark performance or asset class determination. The purpose of this paper is to review the risk/return performance and the relative tracking error of the principal commodity and futures markets based benchmark indices as well as that of the indices used to track various commodity trading advisors and managed futures based trading products.

Results indicate that existing commodity and managed futures benchmark indices have unique risk/return patterns and thus are of use to investors who wish to monitor these return patterns. However, results also indicated that as for traditional equity benchmark indices such the Dow Jones and S&P 500 indices for which risk and return characteristics differ, various commodity and managed futures indices likewise offer differing risk/return attributes. Thus, as for stock and bond investment, one must understand the unique structural and computational differences in commonly used benchmarks and must expect commodity and managed futures indices to continue to be created which attempt to track unique asset class characteristics.

Comparisons of Commodity and Managed Futures Benchmark Indices

I. Introduction

Security market benchmark indices are commonly used to provide a performance index which reflects the particular style of an investment manager. While benchmark indices are common in the areas of stock and bond investment, in recent years futures and options based commodity trading products have grown rapidly as alternative means of investing in cash market commodity and financial securities. While commodity and futures market based indices (e.g., Dow Jones Commodity index) have existed for years, other commodity based indices (Goldman Sachs, JP Morgan) and managed futures indices (e.g., MAR, Barclay, TASS, and EACM indices) based on the performance of commodity futures trading advisors (CTAs) have also recently been used in discussions of benchmark performance or asset class determination. Other performance indices have been created which attempt to provide an active futures market based return index which attempts to mimic the performance of active managed futures traders who hold both long and short positions. For instance, the Mount Lucas Management (MLM) index uses a trend-following technical trading rule to determine whether a particular futures contract should be held as a long or short position.

The purpose of this paper is to review the risk/return performance and the relative tracking error of the benchmarks based on commodity prices as well as the indices that track various managed futures based trading products. In the following section, we review the composition and structure of popular commodity and managed futures based indices. In section III the data and methodology employed in this study are briefly reviewed. Results are presented in section IV. Results indicate that commodity and managed futures benchmark indices have sources of risk and return that are distinct from traditional assets such as stocks and bonds and offer investors an important area of diversification. However, commodity and managed futures indices differ in significant ways. There appears to be little connection between the returns of major commodity indices and the returns of CTA indices. Unlike stock and bond markets, where broad indices

provide suitable benchmarks for the performance of an actively managed portfolio, results presented suggest that CTA-based indices provide a better benchmark for an actively managed futures portfolio than either passive or active commodity based performance indices. Results also show unique tracking error between similar CTA indices, thus, despite the high correlation between CTA based performance indices, a month-to-month comparison of the index returns reveals considerable discrepancy in certain months. Moreover, results indicate that the actual efficient frontier created from the inclusion of alternative CTA and commodity indices with an existing stand alone investment may differ depending on the CTA or commodity index used. Investors must therefore realize that to the degree that their own investment portfolio differs from the established benchmarks or for which the index is noninvestible, the relationships describe above may not be met. Conclusions and suggestions for future studies are discussed in Section V.

II. Commodity and Managed Futures Indices in Asset Management

The increased use of commodity trading vehicles and managed futures products in investment management has led practitioners to create indices that offer performance benchmarks for purchasers of physical commodities as well as investors in managed futures products such as commodity funds, pools, and segregated commodity trading advisor accounts. As is true for stock and bond performance indices, commodity and managed futures-based indices have a variety of uses. They are a source of information on cash commodity and futures commodity market trends, are used as performance benchmarks for evaluation of commodity trading advisors, and provide a historical track record useful in developing asset allocation strategies.

Classification of Benchmark Indices

Benchmark indices may be either noninvestable, and thus useful primarily as a measure of relative performance, or investible and of use both as a relative return index as well as a surrogate for actual investment performance. An index is termed inevitable if it is easily replicated through investment in the

underlying securities. Until recently, most futures-based indices were not investable. However, most of the new commodity based futures price indices are designed to be investable, and many commodity trading firms have begun to offer investment products based on the performance of managed futures indices.[1]

The principal commodity and managed futures indices fall into two categories: The first category is commodity indices based on the returns of futures contracts and/or cash markets. Included in this group are the Dow Jones, CRB, Goldman Sachs, JPMorgan, Bankers Trust, and MLM.[2]. The second group contains indices based on the actual performance of commodity trading advisors. This group consists of MAR, Barclay, TASS, and Evaluation Associates (EACM). The indices in the first group provide returns comparable to passive long positions in listed futures contracts or, in the case of the MLM index, of long and short positions determined by a moving average trading rule. Of these indices, all but the CRB and Dow Jones indices are investable. The CRB is not generally considered investable because the formula for apportioning investment across the different contract months too cumbersome. The Dow Jones index requires daily rebalancing of each position, making it expensive to replicate.

In addition to commodity indices based on the performance of actual cash/futures market contracts, a number of managed futures indices have been derived to track the performance of a select group of actively traded managed futures funds, pools, or commodity trading advisors (CTAs). At present, none of these indices is in investable form. There are two reasons for this. First, the firms that create these indices (e.g. MAR) do not release the exact composition of the indices. Second, since management and incentive fees for managed futures accounts are generally negotiable and CTAs may manage accounts differently for different clients, knowing the exact composition of the index is not sufficient to replicate the reported returns.

Commodity Indices:

Commodity indices attempt to replicate the return available to holding long positions in agricultural, metal, energy, or livestock investment. Since the cost of carry model insures that the return on a fully

margined position in a futures contract should mimic the return on an underlying spot deliverable, futures contract returns are often used as a surrogate for cash market performance. Futures contract based commodity indices have three separate sources of return: price, roll, and collateral return. Price return derives from changes in commodity futures prices. Roll return arises from rolling long futures positions forward through time and may capture a risk premium over time if futures prices are a downwardly biased forecast of future spot prices. Collateral return assumes the full value of the underlying the futures contracts are invested at a risk-free interest rate. This is equivalent to assuming an investor posts 100% margin with Treasury bills. The Goldman Sachs, JPMorgan, and MLM indices include all three return components. Other commodity indices such as the CRB, Dow Jones Futures Index, and Bankers Trust Commodity Index (BTCI) and are based on price return alone.[4] Commodity indices, however, differ widely in purpose, composition, and weighting scheme. The major differences are outlined in Table 1 and discussed below. A fuller description of the indices is in Appendix I.

Commodity indices use either arithmetic or geometric averaging to calculate the index return from the component returns. The JPMCI, GSCI, and BTCI use an arithmetic average of the commodity returns. Investors attempting to replicate the index must rebalance their portfolios monthly to maintain constant dollar weights. The CRB uses geometric averaging of the component returns (which are themselves calculated with arithmetic averaging).

Insert Table 1 about here

Composition and Weightings

Each of the major indices contains a different group of underlying assets. The JPMCI and the BTCI are composed entirely of futures contracts based on industrial commodities (energy and metals). The Dow Jones, CRB, and Goldman Sachs indices include industrials, grains, and soft commodities. The MLM

index contains those groups as well as interest rate and currency futures. The composition of each index is periodically reviewed, typically annually, and contracts may be added or deleted based on previously announced criteria. For example, the composition of the MLM index depends on volume and open interest in the underlying contracts, and the GSCI depends on the importance of the underlying commodity to world production.

Commodity indices also differ in the relative emphasis placed on different commodities and the procedure used to determine the weightings in the index. The weights of the JPMCI are determined by an optimization procedure that maximizes an objective function that includes positive correlation with economic growth and unexpected inflation, negative correlation with stock and bond returns and a high Sharpe ratio. Equal weighting is used in the CRB and Dow Jones futures index. A world production weighting is used in the GSCI. Weights in the BTCI are approximately fifty percent energy and fifty percent non-energy industrial commodities.[5] Composition and weighting schemes are driven partly by liquidity considerations. An equal or production-weighted index that includes relatively illiquid contracts may be difficult to replicate. For example a $1 billion investment in the GSCI would represent an investment in live hog futures equal to 27% of the average open interest in live hog futures during 1993, making it difficult to roll a position forward without influencing the price.

Managed Futures Indices:

Managed futures indices also differ widely in purpose, composition, and weighting scheme. The major differences are outlined in Table 2 and discussed below. A fuller description of the indices is in Appendix I. Differences in composition arise because some indices require CTAs to have longer historical track records or larger asset size before inclusion. Indices also differ by type of managed futures accounts followed (e.g. CTAs, funds, pools, hedge funds). Managed Account Reports (MAR) tracks the performance of individual CTAs as well as CTA Funds and Pools while TASS and Barclays offer performance measurement for CTAs only.[6] MAR produces both equal-weighted and dollar-weighted

indices while TASS offers dollar weighted indices and Barclay and EACM provide equal-weighted indices. The proper benchmark for an investor thus depends in large part on whether the portfolio contains a large number of smaller CTAs, in which case an equally weighted index is appropriate, or invests primarily in the large, well-known CTAs, which calls for a dollar-weighted benchmark. Dollar-weighted indices are easier to replicate, however, as monthly rebalancing is not required. This is an important consideration, as rebalancing requires shifting assets among CTAs based on their end-of-month performance. Performance figures are often not publicly available for several days after month-end, and many CTAs require advance notice of redemptions. Thus replicating an arithmetic index will be very difficult.

Both Barclay and MAR provide subindices based on market sector (e.g., currency, energy, financial, stock, agricultural) and trading sector (e.g., discretionary and systematic). MAR also provides broad indices for commodity funds and pools and sector subindices (e.g., guaranteed, offshore, multi-advisor, single advisor, public, private).

Insert Table 2 about here

TASS, MAR, and Barclay are broad-based, and attempt to include all CTAs or Pools which meet listing requirements. TASS and MAR have less stringent listing requirements than Barclay. TASS and MAR require a one year historical track record, while Barclay requires four. The returns of an individual CTA or pool are calculated as a weighted average composite of the different accounts managed by each CTA, including public commodity funds, private pools, and individual managed accounts. Moreover, basing an index on composite returns and including a large number of CTAs with few assets under management may result in an index that is very difficult to replicate and may not closely track a typical institutional managed futures portfolio. EACM tracks the monthly performance of a preselected set of CTAs and hedge funds. The EACM equal weighted index is rebalanced only once each year. The initial

investment is assumed to be invested with each trader for the full year. EACM recognizes that a lack of liquidity and fund investment rules prevent monthly rebalancing.

The EACM index is more susceptible to survivor bias than the general indices because much of the early return series for the index is hypothetical. Since inclusion in the index is a subjective decision made by EACM management, a historical reconstruction of the index is difficult. The EACM index is comparable to the Dow Jones Industrial average in that the CTAs and hedge funds included in the index tend to be the dominant players in the industry. Furthermore, EACM maintains frequent contact with the fund managers, and would likely remove a fund from the index well before the fund dissolved. A final difference between EACM and the other indices is the method of calculating returns. MAR, TASS, and Barclay rely on the CTAs to report monthly returns, either directly to the index managers or through disclosure documents, so data quality may be an issue. The returns in the EACM index are generated by its own tracking of investment managers who are following specifically designated programs. Thus potential problems resulting from self-reporting and from using composite returns are lessened.

III. Data and Methodology

Monthly returns were derived for a series of stock, bond, commodity, and managed futures indices for the time period from January, 1987 through December 1995.[7] Data was obtained for each of the indices and relevant subindices described in Appendix I (JPMCI, GSCI, BTCI, CRB, DJCI, MAR, TASS, Barclay, EACM, MLM), as well as the Standard and Poors 500 Stock Index, the Salomon Brothers U.S. Bond Index, one-month Treasury bill yields, the US Dollar Index, and the U.S. Producer Price Index. Stock, bond, commodity, currency and inflation indices are obtained from Datastream. CTA indices are obtained from respective index providers.

Returns for all data series are expressed as monthly holding period returns. The test period is also divided into two intervals, 1/87-12/90 and 1/91-12/95, to provide a means of comparing the stationary of managed futures and commodity index performance and to permit complete analysis of several managed

futures indices that started after 1990. Statistical tests include presentation of descriptive risk and return characteristics and return correlations between each of the primary and subindices.

Tracking error measures the degree that a managed futures or commodity index matches the performance of some other index. Tracking error an important factor in selecting a benchmark, as an appropriate benchmark should closely track the target portfolio. There are several methods of determining the degree of tracking error. These include (1) the mean error and standard deviation of the monthly return differences, (2) autocorrelation coefficients of the return differences, which determines if a lead/lag relationship exists between the indices, and (3) the mean absolute return tracking deviation and the standard deviation of the absolute monthly tracking deviation.

The historical returns available for many indices, both commodity price indices and managed futures indices, predate the creation date of the index. Historical returns and correlations prior to the official launch date of the index are less reliable indicators of future performance than returns subsequent to the launch date because these returns overlap the optimization period used to select the inputs and the weighting scheme. Many of the commodity indices have been recently introduced, so it is impossible to exclude the testing period from the analysis without excluding the majority of the indices, The date of actual index inception is noted in Appendix I. In addition, several of the indices have been reconstructed at various times to better reflect a particular performance history. As a result the performance of an index reported in this paper may not strictly reflect the performance of that index in other research.

IV. Results

A. Analysis of Commodity and Managed Futures Indices as Stand-Alone Investments

In Table 3 the average monthly arithmetic returns, geometric returns, and standard deviations of monthly arithmetic returns for the sample of futures products, stock, bond, and commodity indices over the 1/87-12/95 period are presented. Results in Table 3 show that the mean arithmetic monthly returns of the broad-based commodity indices (JPMCI, 1.05%; GSCI, 1.14%), all of the CTA indices (MAR CTA -

Equal Weighted (MARCTAeq), 1.24%, MAR CTA - Dollar Weighted (MARCTA$), 1.45%, Barclay, 1.11%, TASS, 1.10%), and the trading index (MLM, 0.81%), were greater[8] than the average return of the commodity indices that do not include collateral return (CRB. 0.17%, DJCI, 0.23%, Bankers Trust 0.19%.[9] Returns on the MAR CTA indices outperformed traditional financial benchmarks such as the S&P 500 (1.22%), the Salomon Brothers U.S. Government Bond index (0.71%), and the T-bill index (0.44%). However, as in previous tests (Elton et al., (1992) and Irwin et al., (1992)), the average return of the MAR dollar weighted Public Fund/Pool Index (1.06%) and MAR Equal Weighted Fund/Pool Index (0.83%) were below stock and bond returns, reflecting the additional performance costs of commodity fund and pool investment relative to direct CTA investment.[10]

Insert Table 3 and Figure 1 About Here

Results in Table 1 also show that the monthly standard deviation of each of the CTA-based indices {MARCTAeq, 4.75%, MARCTA$, 4.21%, Barclay, 4.77%, and TASS, 4.23%), were greater than most of the futures-based and cash market-based commodity indices (MLM, 1.65%, JPMCI, 4.73%, GSCI, 4.42%, CRB, 2.37%, DJCI, 2.50%, and BTCI, 5.07%). CTAs indices were also more volatile than the returns of the S&P 500 (4.27%), the US Bonds (1.36%), and T-bills (0.14%). Figure 1 plots the performance of the indices in return/standard deviation space. Most of the CTA measures (circles) are clustered in an area that represents an attractive risk/return relationship as compared to commodity indices (triangles) and other asset classes (squares).

The performance of managed futures indices and the MLM index during the second subperiod (1991/1995) is worse than the first subperiod (1987-1991); MARCTA$ (2.02%/1.00%); Barclay (1.99%/0.40%); TASS (1.86%/.49%); MLM (1.16%/0.52%). The return of the GSCI and JPMCI in the second subperiod below that of the first; JPMCI (2.04%/.255%); GSCI (2.32%/.19%). The lower returns

of the GSCI and JPMCI is due in part to lower short-term interest rates, which reduced the imputed return on cash balances. However, this accounts for only a small part (0.18%) of the difference. Returns to the other commodity indices were similar between the two subperiods; DJCI (0.20%/0.25%), CRB (.17%/.16%), suggesting the lower second period returns in the GSCI and JPMCI were due to the relatively larger weightings in energy and industrial commodities, which experienced lower prices during that period..

CTAs can take long and short positions, and may be expected to perform well in both up and down markets, and are generally believed to benefit from volatility. The lower returns of managed futures indices in the second subperiod may be linked to the lower volatility of commodity futures and financial assets during this period (S&P 500, 5.55%/2.92%, US Bonds, [1.50%/1.25%, JPMCI, 5.77%/3.55%).

B. Managed Futures and Stock and Bond Correlation

The decision to add an investment product to an existing portfolio depends on the relative means and standard deviations of the investment vehicle and the existing portfolio as well as the correlation between the investment vehicle and the portfolio. The mean and standard deviation of returns for various investment categories are reported in Table 3 and the correlation between each of the comparison groups is given in Table 4 (parts A and B). The correlation between the S&P 500 and the CTA-based managed futures indices and subindices are between 0.15 and -.30. While most correlations are negative, only a few are statistically significant[11]. Among the futures- and cash-based commodity indices, the JPMorgan and Bankers Trust indices are significantly negatively correlated with the S&P 500, while the GSCI, CRB, Dow Jones, and MLM have a negative sign but are not significant. It will be interesting to see whether the negative correlation in the JPMCI continues in the future. Negative correlation with equity returns was one of the objectives of the index creators, so the ex post correlation is not surprising. In the two years since the index was launched the correlation between SP500 and JPMCI has been a statistically insignificant -0.06.

Insert Table 4 about here

The correlation coefficients between the Salomon Brothers U.S. Government bond index and the CTA-based managed futures range from .37** (MAR Guaranteed) to -.23* (MAR Energy).[12] The high correlation in the MAR guaranteed subindex is consistent with the fact that a guaranteed fund typically place the bulk of its assets in zero-coupon bonds to ensure investors the return of principal at some point in the future. Likewise, the negative correlation with the energy subindex is not surprising given the positive correlation between energy costs and interest rates. Among futures indices, the JPMCI and GSCI are negatively correlated with the bond index (-.24** and -.15 respectively). The CRB (-.26**), Bankers Trust (-.28**), and Dow Jones (-.24**) are also negatively correlated bonds. The larger negative coefficients in the second group is probably due to the lack of collateral return in these indices, which is positively correlated with the bond return.

C. CTA Index Correlation

Among CTA based indices, the broad indices have very high correlations. The MARCTA$, MARCTAeq, Barclay, TASS. and EACM indices have correlations above 0.85. There is considerable variation among some of the subindices, however, particularly those which have a small population of CTAs. Correlation between MAR and TASS financial (0.56**) and energy (0.39**), and discretionary (0.44**) subindices are relatively low, while, currency (0.91**), diversified (.91**) and systematic/trend following (0.97**) subindices are high. Similarly, the MAR and EACM discretionary subindices have a lower correlation (0.51**) than the systematic/trend following subindices (0.95**).

When subindices are compared to broad indices, the MAR energy index (0.11) offers the lowest correlation with the MARCTA$ index, while CTA trend-followers (.96**) have the highest correlation with the MARCTA$ index.[13] Barclay subindex results are similar. These results suggest there is little difference between broad indices constructed by MAR, TASS, Barclay, and EACM, and furthermore that these

indices would provide suitable benchmarks for the performance of a diversified or trend-following CTA. However, these broad indices will provide poor benchmarks for CTAs who concentrate in a single market (e.g., energy traders). Lastly, the low correlation offered by single-market CTAs suggests they offer more diversification benefit to a typical CTA portfolio than a trend-following or diversified CTA.

The broad CTA indices (MAR, Barclay, TASS, and EACM) are relatively uncorrelated with commodity indices. The MARCTA$ index has no correlation with the JPMorgan (.05), Goldman Sachs (.02), or Dow Jones (.14) indices. It is somewhat more highly correlated with the CRB (.29**). Even though statistically significant, the latter index explains only a small percentage of the variation in the MAR index. Low correlation between CTA returns and commodity index returns (correlations are below .25 across all cash and futures- based commodity indices) reflects fundamental differences in the return-generating processes: First; CTAs take long, short, and spread positions in both futures and options contracts, while indices are strictly long futures and do not include options. Second, commodity indices generally do not include fixed income, currency, or equity futures, while many CTAs actively trade in these contracts. An exception is the MLM Index, which takes short and long positions (but excludes spreads and options). This index correlates more highly with the MARCTA$ index (0.45**) and other CTA indices, reflecting the predominance of trend-following traders in the MAR CTA index (Lukac [1985]).[14]

D. Tracking Error

Tracking Error: Mean and Standard Deviation

Tracking error is the difference between the return on a portfolio and the return on a market index. Among indices that have similar risk/return characteristics, the index that minimizes tracking error with a given portfolio may be regarded as the more suitable benchmark. Prior studies of tracking error have generally focused on stock or bond indices. Reilly et al. (1992) study Treasury bond performance indices. They define several measures of tracking error, the principal measures being the average monthly tracking

error, the standard deviation of the monthly tracking error, the mean absolute value of the tracking error, and the autocorrelation of the tracking error over time.

Each of these measures sheds light on a particular aspect of the tracking process. Mean tracking error identifies whether an index outperforms the comparison index. Standard deviation and absolute value of tracking error measure the dispersion of tracking error over time. Autocorrelation studies test whether an index leads or lags the comparison index. Each of these tracking measures is calculated for a sample of commodity and CTA indices and subindices. We also report the minimum and maximum monthly deviation.

There is less tracking error among the principal commodity and CTA indices than there is among commonly used indices of U.S. stocks. For instance, the standard deviation of the monthly tracking error between the S&P 500 and the Russell 2000 index of small stocks was 2.97% during the 1987-1995 test period. Thus a 95% confidence interval for the difference between SP500 and Russell 2000 returns in a given months would be about ±6%. The SP500 and the Dow Jones industrials tracked more closely, with a standard deviation of 1.15, for a confidence interval of about ±2.3%.

Results in Table 5, indicate the MARCTA$ and TASS indices track with a monthly standard deviation of just 1.05%. Barclay and TASS indices, in contrast, have a higher standard deviation of monthly tracking error of 2.27%. The higher standard deviation is largely explained by a single month when the Barclay and TASS indices deviated by 12.64%. The MLM index tracks the CTA indices poorly. Its standard deviation with the MARCTA$ index is 3.76%. The high correlation noted in the previous section did not translate into low tracking error. The minimum and maximum monthly tracking error during the period were -5.84% and 12.97%. By comparison, the tracking error between the S&P 500 and Dow Jones industrials ranged from -3.82% to 2.96%. The MLM index cannot be regarded as a suitable benchmark for CTAs, although the high correlation suggests the index shares some common factors with the CTA indices.[15]

Insert Table 5 about here

Among commodity indices, tracking error was smallest between JPMCI and GSCI with a 2.01% standard deviation and 5.68% maximum. Tracking error between the JPMCI and DJCI was larger, with a standard deviation of 5.00% and a maximum of 21.84%. Tracking error between the GSCI and DJCI was similarly large, as was the error when the JPMCI and GSCI were compared to the CRB. The data indicate the JPMCI and GSCI are very similar, the CRB and DJCI are similar, but the two groups differ dramatically.

Insert Table 6 about here

The differences in index performance are even more evident when CTA subindices are compared to commodity subindices. For instance, as indicated in Table 6, the tracking error between the MAR energy subindex and the JPMCI energy subindex ranged from -40.31% to 25.97%, and the error between the MAR energy subindex and the Barclay energy subindex was as low as -16.88% and as high as 9.45%. This degree of tracking error is surprising considering how closely the overall MAR and Barclay indices tracked. MAR and Barclay energy subindices also had large tracking errors with the MLM index. Within the energy subgroup, MLM tracked the CTAs better than the commodity indices. It does not appear that commodity indices are useful benchmarks for CTA performance, even when the assets traded by the CTA exactly match the assets in the underlying index.[16] Results were similar when this analysis was repeated for other energy and agriculture subindices as well as subindices of discretionary traders compiled by MAR, Barclay, and EACM.

Tracking Error: Autocorrelation

Autocorrelation in tracking error may occur if a portfolio's returns lead or lag the benchmark index. Autocorrelation coefficients in the monthly tracking error are examined and no clear evidence of tracking error autocorrelation is reported. The MLM index displayed marginally significant second-order tracking error autocorrelation when compared to CTA indices, but this can be attributed to autocorrelation in the underlying MLM index that carries through to the tracking error. Similarly, the CRB index displays first-order autocorrelation that carries over to the tracking error with the JPMCI and GSCI. However, the autocorrelation is not present in the tracking error between the DJCI and CRB.

Absolute Tracking Error

The mean tracking error may be affected by averaging across various months. Absolute tracking error, therefore, may offer a more consistent picture of the expected monthly error. In Tables 5 and 6, the absolute tracking error is reported for major CTA and commodity indices as well as CTA and commodity subindices. As for the reported mean errors, the CTA indices with the lowest mean error (MARCTA$ and TASS) have the lowest absolute error (.78%). This is not surprising since both MAR and TASS are dollar-weighted CTA indices covering large CTA populations. In contrast, the mean absolute error between the equal weighted Barclay CTA index and the TASS and MARCTA$ indices are above 1.25%. When the EACM 100 index is compared to the MARCTA$, TASS, and Barclay CTA indices, mean absolute errors are above 2.5%, possibly due to the inclusion of hedge funds in the EACM index. As indicated in Tables 5 and 6, monthly absolute errors in excess of 2% to 3% are common for most commodity indices, and commodity subindices, and CTA subindices. The level of noise in these data compounds the investors problem. Even if an investor selects a benchmark with appropriate ex ante risk/return characteristics, it may take several years before enough data can be collected to demonstrate the point with any degree of confidence.

E. Portfolio Performance

The impact of the relative tracking error for CTA and passive commodity benchmarks and the reported correlations between the alternative CTA and commodity benchmarks and traditional investments such as the S&P 500 may results in differential portfolio performance when CTAs or commodity benchmarks are considered as additions to existing assets. While previous studies [Schneeweis et. al., 1996; Edwards and Park 1996] have shown that CTAs should be added to existing stock and bond indices, previous studies have not reviewed the impact of adding various CTA or commodity indices on the full risk and return frontier.[17] However, small differences in tracking and correlation relationships may result in differential judgements as to the level of investment or the investability of hedge funds. In Tables 7a, the impact on relative return and risk (e.g., Sharpe ratio) of the MAR$, MAREW, Barclay, and TASS CTA index classifications combined with the S&P index is shown. The results show that as the the MAR$, MAREW, Barclay, and TASS index returns are added to a 100% investment in the S&P 500 index, the relative return/risk relationhsips differ. Thus, the use of hedge funds as part of an investors overall asset mix may be impacted by the differential use of alternative CTA fund indices. Similary, as shown in Table 7b, similar results are shown for the impact on relative return and risk for popular passive commodity indices when considered as part of a stand alone stock portfolio.

Insert Tables 7a and 7b about here

In Figures 2a and 2b, the various efficient frontiers are given for adding the alternative active CTA indices and passive commodity indices to an existing S&P 500 position. For the CTA indices, while all CTA indices resulted in reductions in risk when considered as additions to the existing S&P 500 position,

the impact on return varies greatly. The most positive impact was reported for the Managed Accounts Reports $ weighted CTA index, with lower risk/ return impacts reported for the MAR equal weighted and the Barclay and TASS CTA indices. One reason for the differential returns among the alternative CTA indices is due to their alternative CTA selection rules. For instance, Barclay CTA index excludes CTAs with less than a four year track record, while TASS as the least restrictive (one month) listing restriction. Thus the reported efficient frontiers may vary greatly when alternative CTA indices are used. The degree of impact will, of course, depend on the time period of analysis. As important, however, is that for a typical level of CTA investment (approximately 10%) the impact on risk and return are total risk and return are similar. Similarly, given the similar approximately zero correlation with the S&P 500 the maximum reduction in risk for all indices is at the approximately 50% CTA investment level.

Insert Figure 2a and 2b About Here

For passive commodity indices, the impact of the use of alternative commodity indices likewise impacts the reported efficient frontier. For the GSCI and JPMCI the reported efficient frontiers are similar. However, the reported efficient frontiers for the inclusion of the CRB index or Bankers Trust index are dramatically different from that of the GSCI and JPMCI. While in all cases the reported risk is reduced with passive commodity indices are included with the S&P 500, the degree of risk reduction and return performance varies depending on which commodity index is used. For the CRB and Bankers Trust index, the principal difference for the lower reported return is that neither index includes a Treasury bill return on the collateral. If the monthly return to Treasury bills (.44%) were added to their return, the relative efficient frontiers would be more similar. Thus, as for CTAs, assumed large investments in the respective indices may results in differences in investor's use of commodity indices as alternative investments.

V. Conclusions

This study reviewed the risk/return performance and relative tracking error of the various commodity indices and indices used to track managed futures performance. Results indicate that commodity and managed futures benchmark indices have sources of risk and return that are distinct from traditional assets such as stocks and bonds. Assuming this pattern continues, these indices offer investors an important area of diversification. Each index studied also has a unique construction that results in differential return correlation with alternative assets and thus each may be useful as a performance benchmark for unique portfolios, or, if the index is tradeable, as an addition to an existing portfolio.

However, commodity and managed futures indices differ in significant ways. There appears to be little connection between the returns of major commodity indices and the returns of CTA indices. Unlike stock and bond markets, where broad indices provide suitable benchmarks for the performance of an actively managed portfolio, the return to a buy and hold commodity investment strategy provides a poor forecast of CTA returns. Results presented therefore suggest that CTA-based indices provide a better benchmark for an actively managed futures portfolio than either passive or active commodity based performance indices. Results also show unique tracking error between similar CTA indices, thus, despite the high correlation between CTA based performance indices, a month-to-month comparison of the index returns reveals considerable discrepancy in certain months. Moreover, results indicate that the actual efficient frontier created from the inclusion of alternative CTA and commodity indices with an existing stand alone investment may differ depending on the CTA or commodity index used. Investors must therefore realize that to the degree that their own investment portfolio differs from the established benchmarks or for which the index is noninvestible, the relationships describe above may not be met.

Bibliography

Chance, D. Managed Futures and Their Role in Investment Portfolios (Institute of Chartered Financial Analysts, 1994).

Edwards, F. and J. Park. “Do Managed Futures Make Good Investments,” The Journal of Futures Markets, Vol. 16, No. 5 (August, 1996), 475-517.

Bailey, Jeffrey. “Evaluating Benchmark Quality,” Financial Analysts Journal (May-June, 1992), 33-39.

Bodie, Zvi, and Victor Rosansky. "Risk and Return in Commodity Futures," Financial Analysts Journal 36 (May-June 1980), 3-14.

Edwards, F. and J. Park. “Do Managed Futures Make Good Investments,” The Journal of Futures Markets, Vol. 16, No. 5 (August, 1996), 475-517.

Elton, E.J., M.J. Gruber, and J.C. Rentzler. "Professionally Managed, Publicly Traded Commodity Funds," Journal of Business (April 1987), 177-199.

Elton, E.J., M.J. Gruber, and J.C. Rentzler. "New Public Offerings, Information, and Investor Rationality: The Case of Publicly Offered Commodity Funds," Journal of Business, Vol. 6, No. 1, (1989) 1-15.

Elton, E.J., M.J. Gruber, and J. Rentzler, “The Performance of Publicly Offered Commodity Funds,” in C. Peters ed. Managed Futures (Probus Publishing, 1992), 387-401.

Epstein, Charles ed., Managed Futures in the Institutional Portfolio, (Wiley, 1992).

Irwin, S. “Further Evidence on the Usefulness of CTA Performance Information in Public Commodity Pool Prospectuses and a Proposal for Reform,” Advances in Futures and Options Research, Vol. 7 (1994).

Irwin, S., T. Krukemyer, and C. R. Zulaf, “ Are Public Commodity Pools a Good Investment, “ in C. Peters ed. Managed Futures (Probus Publishing, 1992),405-433.

McCarthy D, T Schneeweis,., and R. Spurgin. "Investment through CTAs: An Alternative to Public Commodity Funds", Journal of Derivatives (Summer, 1996), 36-47.

New York Stock Exchange. Commodity Futures as an Asset Class Report by Powers Research Associates, (January, 1990).

Peters, C.C. ed. Managed Futures: Performance Evaluation and Analysis of Commodity Funds, Pools and Accounts (Probus Publishing, 1992).

Reilly, F., G.W. Kao, and D. Wright, “Alternative Bond Market Indices”, Financial Analyst Journal (May-June, 1992), 44-58.

Reilly and Rashid A. Akhtar, “The Benchmark Error Problem with Global Capital Markets,” The Journal of Portfolio Management (Fall, 1995), 33-53.

Roll, R. " Performance Evaluation and Benchmark Error I.," Journal of Portfolio Management (Summer, 1980), 5-12.

Roll, R. "Performance Evaluation and Benchmark Error II." Journal of Portfolio Management (Winter, 1981), 17-22.

Rudd, A. and B. Rosenberg, "The Market Model in Investment Management," Journal of Finance (Vol. 35. No. 2 (1980), 597-607.

Schneeweis, T. The Benefits of Managed Futures. European Managed Futures Association, August, 1996.

Schneeweis, T., U. Savanayana, and D. McCarthy, "Alternative Commodity Trading Vehicles: A Performance Analysis," Journal of Futures Markets (August, 1991), 475-490.

Schneeweis, T., U. Savanayana and D. McCarthy. "Multi-Manager Commodity Portfolios: A Risk/Return Analysis," in C. Epstein ed. Managed Futures (Wiley, 1992), 81-102.

Schneeweis, T. and R. Spurgin. "Comparisons of Hedge Fund Indices," Working Paper (CISDM, SOM, University of Massachusetts, 1996).

Schneeweis, T. and R. Spurgin. "Multi-Factor Models in CTA/Hedge Fund Return Estimation", Working Paper (CISDM, SOM, University of Massachusetts, 1996).

Schneeweis, T., R. Spurgin, and D. McCarthy. "Survivor Bias in Commodity Trading Advisor Performance" Journal of Futures Markets (Forthcoming, 1996).

Schneeweis, T., R. Spurgin, and D. McCarthy. "Informational Content in Historical CTA Performance" Journal of Futures Markets (Forthcoming, 1997).

Appendix I

Description of Commodity and Managed Futures Indices

Dow Jones Futures and Spot Commodity Index is comprised of twelve commodities: cattle, coffee, copper, corn, cotton, gold, hogs, lumber, silver, soybeans, sugar, and wheat. The weightings of the indices is neutral; that is, each commodity's price on a given day is divided by its price on the base date and the results are totaled. The total is divided by twelve and multiplied by 100 to yield the index. To estimate the price of a commodity five months in the future, two contract months are used, one expiring in fewer than 150 days and one expiring in more than 150 days. It is assumed that each contract expires on the 15th of its delivery month. For each commodity, weights are assigned to the two contracts, based on the number of days between the 150th and the theoretical expiration date of the contract. Then the price of each contract is multiplied by its weight, the results are added, and the sum is divided by the number of days between the expiration dates of the two contracts. The result is the estimated price of the commodity for delivery in exactly five months, or 150 days.

Commodity Research Bureau Index (CRB) futures price index represents an unweighted geometric average of 21 component commodity prices. Each of these
21 commodity prices has been arithmetically averaged for all actively traded contracts expiring on or before the end of the ninth calendar month from the current date, excluding noncycle months. After these two averaging techniques have been performed, the resulting value is converted to a percentage of the base year value.

Goldman Sachs Commodity Index (GSCI) is an arithmetic measure of the performance of actively traded, dollar-denominated nearby commodity futures contracts. As of January 9, 1995, there were 22 commodities in the index.
The weights assigned to individual commodities are based on a five-year moving average of world production. Weights are determined each July and are made effective the following January. All contracts are rolled on the fifth business day of the month prior to the expiration month of the contract. Subindices are calculated for agricultural, energy, industrial, livestock, and precious metals contracts. Two versions of the indices are available: a total return version, which assumes that capital sufficient to purchase the basket of commodities is invested at the risk-free rate, and a spot version, which only tracks movements in the futures prices. This study uses the total return measure. The GSCI was officially launched in 1992.

JPMorgan Commodity Index (JPMCI) is comprised of 11 highly liquid industrial commodity futures contracts. It excludes "softs," relying exclusively on energy, precious metals, and industrial metals. The two nearby contracts for each commodity are used, and rolls are conducted over a five-day period from the fifth to the ninth business day of the month. Component weights are rebalanced monthly according to a scoring system that seeks to maximize risk-return performance, track unexpected changes in inflation, provide a hedge against stock and bond investments, and correlate with economic growth measures. Sub-indices for energy, precious metals, and industrial metals are also available. The JPMCI is published in both spot and total return formats. This study uses the total return measure. The JPMCI was officially launched on September 21, 1994.

Bankers Trust Commodity Index (BTCI) is based on spot rather than futures prices. It assumes ownership of a basket of five physical commodities: crude oil, gold, aluminum, heating oil, and silver. 55% of the weight is given to energy components, and the remaining 45% to the three metals. The basket is priced daily using spot quotes such as the London gold fix. Front-month futures quotes, which are equivalent to cash market quotes, are used for energy prices. BTCI thus reproduces the price changes of a basket of physical assets without the storage and holding costs. The BTCI was launched on July 18, 1994.

Mount Lucas (MLM) Index differs from other indices in two important ways. First, it allows both long and short positions in the underlying futures contracts. Second, it incorporates financial and currency futures (but not stock index futures) into the index, along with the commodities tracked by other indices. The index is an equally weighted average of the monthly returns from 25 separate futures contracts. Within each market (e.g., corn futures), the index will be long or short depending on whether the contract is above or below its trailing 12-month moving average. MLM is a total-return index. It was launched in May, 1989.

Managed Account Reports (MAR) tracks the performance of individual CTAs as well as CTA Funds and Pools that invest in individual CTAs. MAR produces several performance indices, the dominant being the CTA equal-weighted and dollar-weighted indices. MAR classifies CTAs into a number of different groups, and publishes each group's performance index. These groups are currency, energy, financial, diversified,

discretionary, and trend-following. MAR also reports the following subindices for fund and pool performance: guaranteed, multi-advisor, single-advisor, private pools, and public pools. It was restructured in 1992.

Barclay Trading Group, publisher of the Barclay Managed Futures Report, also creates CTA performance indices. Indices are based on monthly returns of CTAs with established track records. Barclay publishes an equal weighted index of all CTAs as well as the following subindices: agricultural, currency, diversified, energy, financial/metal, discretionary, and systematic.

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

EACM offers historical managed futures/hedge fund performance on an array of indices and subdices designed to capture the return to unique managed futures/hedge fund strategies. These indices include relative value, event-driven, equity hedge funds, global asset allocators and short selling. Subindices include long/short equity, convertible hedge, bond hedge, rotational, deal arbitrage, bankruptcy, and multi-event. For managed futures, the principal subindices include discretionary and systematic. It has been publicly available since January, 1996.

[1] See Managed Accounts Reports March, 1996.

[2] Merrill Lynch and other firms have constructed their own futures indices based on traded futures contracts. In fact, given the alternative weighting and design characteristics, the number of alternative futures index forms are unlimited. Likewise, REFCO and other firms involved in CTA analysis have created their own managed futures indices. This paper limits its analysis to those indices which have gained the greatest public attention.

[4] Adding the T-Bill return to commodity indices makes them a surrogate for a full investment in the underlying spot. However, this is only true if the cost of carry model fully captures the relevant price relationship. From 1979 to 1994, the JPMCI total return was 14.8%, of which only 4.8% was price return, 1.3% was roll return and 8.7% was collateral yield.

[5] Equal weights may disproportionately reflect changes in smaller commodity markets. For example, the CRB assigns equal weight to crude oil and orange juice. Likewise, production weights disproportionately reflect current market cycles. The GSCI has only marginal allocations to the base (7%) and precious (3%) metal sectors.

[6] MAR, TASS, Barclays, and EACM are continually improving their index process. For instance, TASS is in the process of introducing indices which adjust for underlying performance fees and interest costs. Similarly, EACM is in the process of further refining their subindices.

[7] Data prior to the launch date of an index are hypothetical, not actual, returns.

[8] The high volatility of commodity indices and the relatively small sample size make it difficult to reject the null hypothesis that any of the total return indices studied have the same mean return. However, since these results are consistent with other longer-term studies, it is reasonable to assume that CTAs will outperform unmanaged commodity indices.

[9] The CRB, DJCI, and BTCI, are spot indices and do not impute a risk-free return. Adding the one-month T-Bill return of 0.44% to the monthly returns provides a more accurate comparison. If this is done, the CRB return would be 0.61%, the DJCI 0.67%, and the BTCI 0.63%, still below the other indices but quite a bit closer.

[10] Research (Schneeweis, 1996) has shown that MAR CTA and MAR public commodity fund and private pool indices have similar minimum returns over similar time periods, and that MAR CTA indices outperform MAR fund and pool indices by about 4% in terms of maximum returns. These results are consistent with CTAs charging lower performance fees, which are charged only for positive performance.

[11] With a sample size of 108, correlations with absolute value greater than .19 are significant at the 5% level (*) and greater than .24 significant at the 1% level (**).

[12] The MAR Energy subindex has only 72 observations, thus it requires higher correlation to achieve the same statistical significance.

[13] The trend-following subindex has been estimated to reflect the principal trading style of most CTAs, which explains the larger correlation with the overall index.

[14] An exception is the MAR energy subindex, which correlates more closely with GSCI and JPMCI than the MLM.

[15] The MLM index, for example, reported outstanding hypothetical returns from the 1970's to the 1989 official launch date, but has since returned approximately the risk free rate.

[16] Results also indicated that as the investment holding period is increased from one month to four that while the absolute tracking error increased, the monthly standardized average and absolute tracking error between all the tested pairs decreases. This result is consistent, with short term deviations in return movements between indices becoming dominated over longer time periods by the general positive return across all indices. To the degree that the bias between any two indices in any one month is random, the use of longer time intervals may result in a closer correlation between measured returns. Thus results indicate that for investors wishing to evaluate CTA or passive commodity indices over longer investment intervals, the relative impact of choosing different indices within a class of active or passive futures indices lessens as the investment interval increases.

[17] Both Schneeweis et. al, (1996) and Edwards and Park (1996) both use an excess benchmark return methodology used by Elton et. al. (1987) and others. This methodology indicates if inclusion of the alternative asset would have increased the Sharpe ratio of the stand alone asset. In both of their studies, CTAs inclusion in various stand alone asset portfolios would have improved an investors' return/risk tradeoff.