

Diversification In The Managed Futures Universe

Derek G Nokes

Monday, May 25, 2015


```
## [1] TRUE
```

Introduction & Motivation

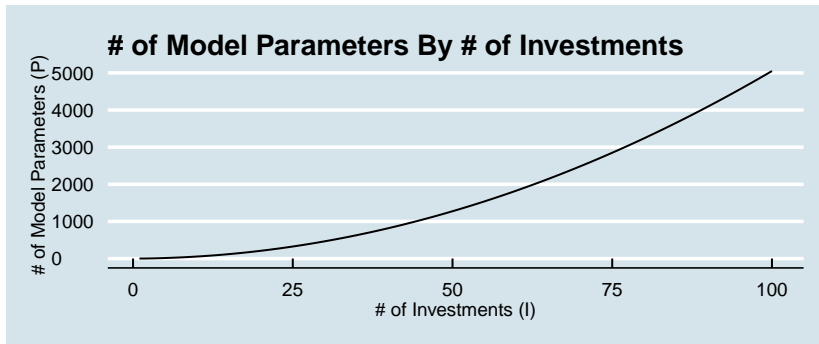
Objective

Maximize investors' future wealth by determining how to allocate capital among a set of available managed futures investments in such a way as to maximize *compound* growth subject to a set of constraints

- ▶ Reducing variability of returns has as much impact on total return as increasing magnitude of returns
- ▶ Portfolio return variability is a function of the co-variability of investment component returns
- ▶ Select sets of investments with *future* positive average returns and low co-variability

Problem: Too Many Moving Parts to Understand Intuitively!

As size of a portfolio increases, number of inter-relationships between components explodes



- Becomes increasingly difficult to understand drivers of portfolio return as number of components rises
- Too many moving parts (particularly during a crisis)

Solution

- Group investments that tend to move together
- Focus on trying to find groups that are independent
 - ▶ One common way to reduce dimension of portfolio allocation problem
- Can be accomplished through use of statistical *factor models*

Data

- Extracted data for all managed futures programs from the Altergis website (<http://www.managedfutures.com/>)
- Scraped managed futures program profiles are found here:
 - ▶ http://www.managedfutures.com/program_profiles.aspx

Raw Data Extraction, Transformation, and Loading (ETL)

For each managed futures program we extract:

[1] Manager Info

- ▶ CTA Name / Address

[2] Program Info

- ▶ Program Name
- ▶ Investment Methodology
- ▶ Instruments/Sectors/Geographical Focus
- ▶ Holding Periods (Short/Medium/Long)
- ▶ Investment Terms and Info

[3] Performance Track Record

- ▶ Monthly Returns

Data Exploration

Systematic – 63.6% of the programs are 100% systematic, while 82.6% claim that the proportion of their operation that is systematic is 90% or above

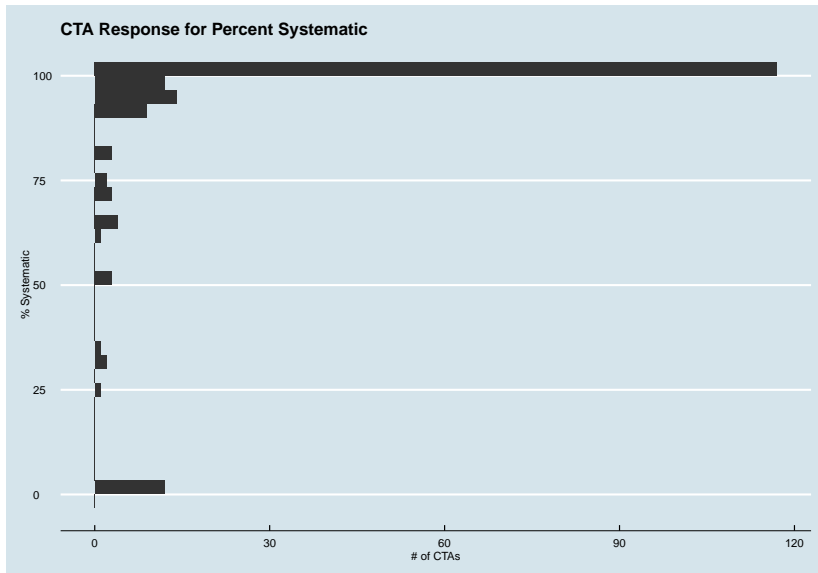
Region of Operations – Vast majority of programs are operated out of either the US or UK (81.55%)

- ▶ 71.84% US-Based
- ▶ 9.71% UK-Based
- ▶ 1.94% do not provide information about geographical location

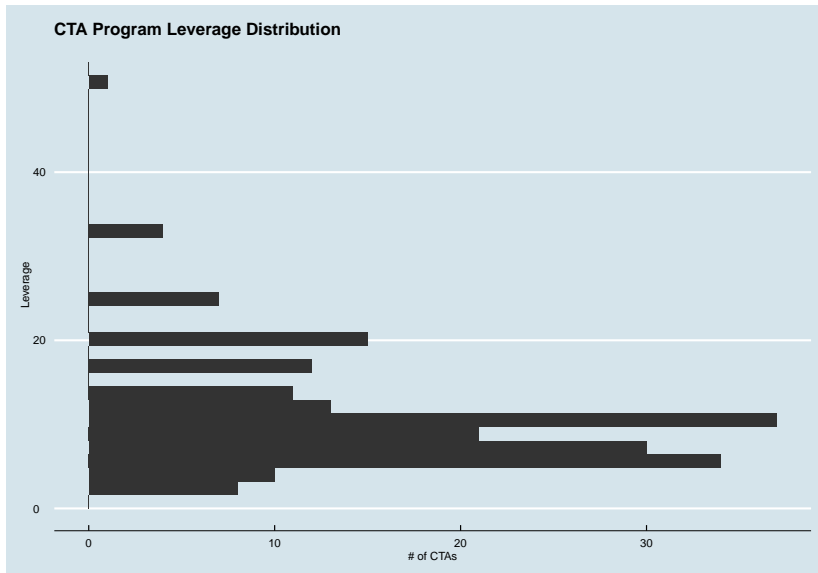
Margin-to-Equity & Leverage – Typical program employs about 9x leverage (i.e., margin-to-equity of 11.11%)

- ▶ Varies a lot across programs (1.7x to 50x)
- ▶ Concentrated around 7x and 12x

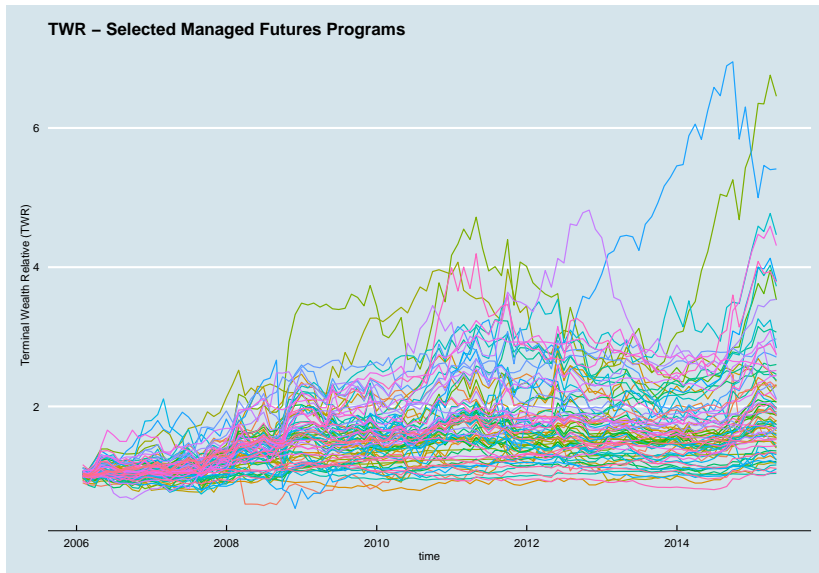
Systematic



Leverage



Monthly Returns



Data Cleaning

- Data cleaning of majority of collected data pertaining to manager and program information was beyond scope of this project
 - ▶ Result: None of this data was used in the modeling sector of the paper
- manager and program information collected is somewhat unstructured and visual inspection of the managed futures website reveals many reporting inconsistencies across managers
- Quick exploratory analysis confirms that data is reported somewhat inconsistently by CTAs
- In particular, there appears to be very little validation of manager and program information submitted by CTAs
 - ▶ Result: this part of the collected data set requires a lot of cleaning and standardization before it can be used effectively in our modeling

Data Preprocessing: Standardizing Returns

- Standardization rescales a variable while preserving its order
- Denote monthly return of i^{th} investment for m^{th} month as $r_{i,m}$ and define standardized return as:

$$\hat{r}_{i,m} = \frac{(r_{i,m} - \bar{r}_{i,M})}{\sigma(r_{i,M})}$$

- ▶ $\hat{r}_{i,m}$ = standardized return of i^{th} investment for m^{th} month using data over time interval 1 to M
- ▶ $r_{i,m}$ = observed return of i^{th} investment for m^{th} month
- ▶ $\bar{r}_{i,M} = \frac{1}{M} \sum_{m=1}^M (\hat{r}_m) =$ mean of return stream of i^{th} investment over time interval 1 to M
- ▶ $\sigma(r_{i,M}) =$ standard deviation of returns for i^{th} investment over time interval 1 to M

Correlations

Represent standardized returns as an $I \times M$ matrix \hat{R} with an empirical correlation matrix C defined as:

$$C = \frac{1}{M} \hat{R} \hat{R}^T$$

- ▶ T denotes the matrix transform
- Correlation matrix (C) of returns (\hat{R}) and covariance matrix ($\Sigma_{\hat{R}}$) of standardized returns (\hat{R}) are *identical*

Principal Component Analysis (PCA)

Objective: Find linear transformation Ω that maps a set of observed variables \hat{R} into a set of uncorrelated variables F . Define $I \times M$ statistical factor matrix as:

$$F = \Omega \hat{R}$$

- Each row f_k ($k = 1, \dots, N$) corresponds to a factor F of \hat{R}
- Transformation matrix Ω has elements $\omega_{k,i}$.
- First row of ω_1 (which contains first set of factor coefficients or 'loadings') chosen such that first factor (f_1) is aligned with direction of maximal variance in I -dimensional space defined by \hat{R} .
- Each subsequent factor (f_k) accounts for as much of remaining variance of \hat{R} as possible (subject to constraint that ω_k are mutually orthogonal)
- ω_k constrained by requiring that $\omega_k \omega_k^T = 1$ for all k .

Principal Component Analysis (PCA) - Continued

Correlation matrix C is an $I \times I$ diagonalizable symmetric matrix that can be written in the form:

$$C = \frac{1}{M} E D E^T$$

- ▶ D = diagonal matrix of eigenvalues d
 - ▶ E = orthogonal matrix of corresponding eigenvectors
- Eigenvectors of correlation matrix C correspond to directions of maximal variance such that $\Omega = E^T$
- Statistical factors / principal components F are found using the diagonalization above

Proportion of Variance

Covariance matrix Σ_F for statistical factor matrix F written as:

$$\Sigma_F = \frac{1}{M} F F^T = \frac{1}{M} \Omega \hat{R} \hat{R}^T \Omega^T = D$$

Total variance of standardized returns \hat{R} for I investments is:

$$\sum_{i=1}^I \sigma^2(\hat{r}_i) = \text{tr}(\Sigma_{\hat{R}}) = \sum_{i=1}^I d_i = \sum_{i=1}^N \sigma^2(f_i) = \text{tr}(D) = I$$

- ▶ $\Sigma_{\hat{R}}$ = covariance matrix for \hat{R}
- ▶ $\sigma^2(\hat{r}_i) = 1$ = variance of vector \hat{r}_i of standardized returns for investment i

Proportion of total variance in \hat{R} explained by k^{th} factor is:

$$\frac{\sigma^2(f_k)}{\sum_{i=1}^I \sigma^2(\hat{r}_i)} = \frac{d_k}{\sum_{i=1}^I d_k} = \frac{d_k}{I}$$

Inverse Participation Ratio (IPR)

IPR_k of k^{th} factor ω_k is defined as:

$$IPR_k = \sum_{i=1}^I (\omega_{k,i})^4$$

- IPR quantifies reciprocal of the number of elements that make a significant contribution to each eigenvector
- IPR is bounded by two cases:
 1. An eigenvector with identical contributions $\omega_{k,i} = \frac{1}{\sqrt{I}}$ from all I investments has $IPR_k = \frac{1}{I}$
 2. An eigenvector with a single factor $\omega_{k,i} = 1$ and remaining factors equal to zero has $IPR = 1$

Participation Ratio (PR)

- Inverse of IPR provides more intuitive measure of significance of a given factor

$$PR = \frac{1}{\sum_{i=1}^I (\omega_{k,i})^4}$$

- Large PR indicates that many investments contribute to the factor; small PR indicates that few investments contribute to the factor
- PRs facilitate identification of statistical factors that represent macroeconomic scenarios, namely those with many participants
- Also help us identify factors that represent microeconomic scenarios, namely factors with few participants

Portfolio Return & Variability

- Portfolio compounded return for interval from months 1 to M :

$$r_{P,M} = \left(\prod_{1=m}^M \left(1 + \left(\sum_{i=1}^I (r_{i,m} w_{i,m}) \right) \right) \right) - 1 = \left(\prod_{1=m}^M (1 + r_{P,m}) \right) - 1$$

- Assume:

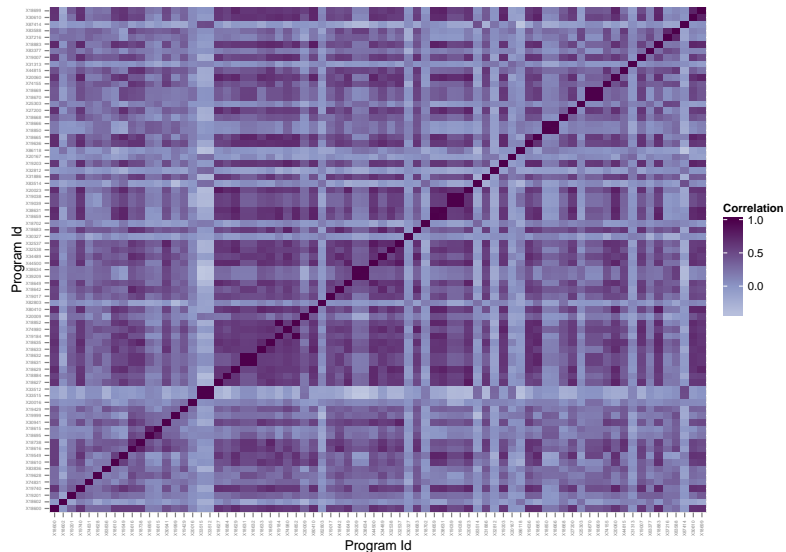
- ▶ Component returns are normally distributed
- ▶ Portfolio returns are multivariate normally distributed

- Standard deviation of portfolio returns (using matrix notation as_):

$$\sigma_{P,M} = \sqrt{\text{Var} \left(W_m^T R_m \right)} = \sqrt{W_m^T \Sigma W_m}$$

- ▶ W_m = vector of portfolio component weights for month m , T denotes transpose operator, R_m = vector of month m component returns, and Σ = return covariance matrix

Correlation Matrix



Top 10 Factors

-First 10 factors explain a very significant proportion of total variance

Table 1:Top 5 Factors

Factors	% of Variance	Cumulative % of Variance
Factor 1	36.8	36.8
Factor 2	8.8	45.7
Factor 3	6.5	52.1
Factor 4	4.0	56.1
Factor 5	3.6	59.8
Factor 6	3.0	62.8
Factor 7	2.5	65.3
Factor 8	2.2	67.6
Factor 9	2.2	69.8
Factor 10	2.0	71.8

1st Factor

- Sort factor loadings and look at top and bottom
 - ▶ Long- and medium- term trend-following programs make strongest positive contributions
 - ▶ Volatility selling, short-term and relative value programs having small or negative loadings
- Participation ratio indicates that 52 components make significant contributions to first factor
- In strong contrast to other factors where number of components making significant contributions is between 2 and 34

10 Largest Factor Loadings for Factor 1

Table 2: Largest Factor 1 Loadings

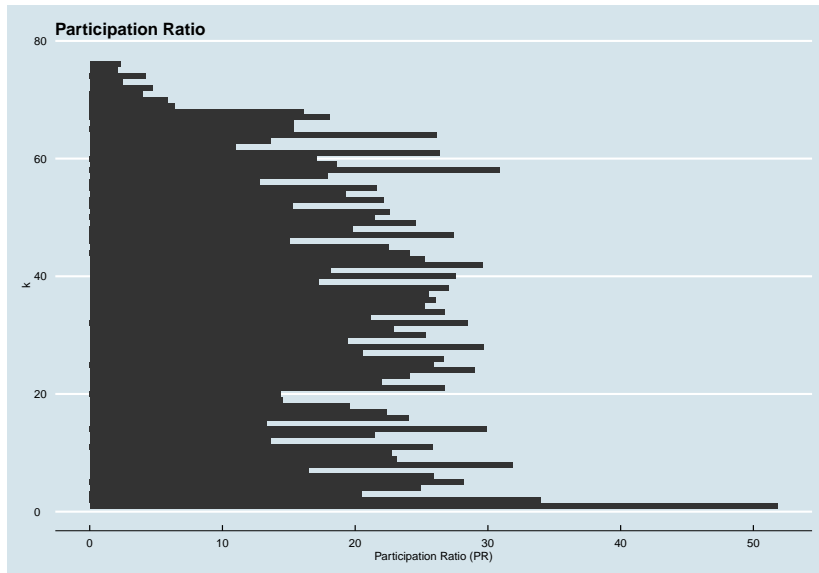
	Program Name	Factor 1
75	Global Directional Portfolio	0.1514
41	Genesis	0.1516
66	Diversified Futures	0.1522
28	Alpha Trend	0.1522
22	World Monetary and Agriculture (WMA)	0.1525
1	Abraham Diversified	0.1530
49	Global Diversified	0.1550
38	Systematic	0.1563
56	AlphaQuest Original (AQO)	0.1586
26	Classic	0.1603

10 Smallest Factor Loadings for Factor 1

Table 3:Smallest Factor 1 Loadings

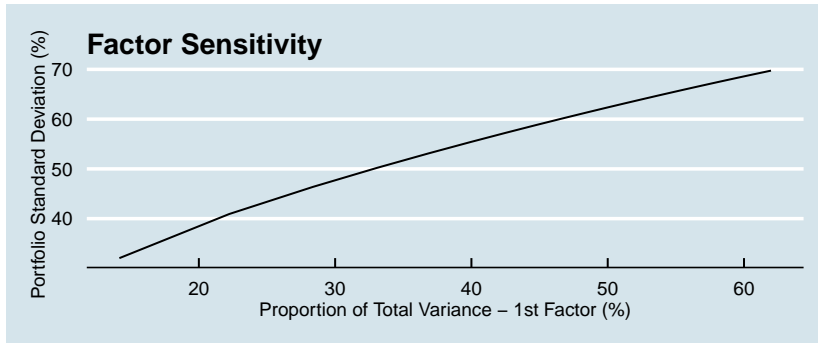
	Program Name	Factor 1
18	Relative Value Volatility 2X	-0.0434
19	Relative Value Volatility 1X	-0.0410
74	Strategic Fund	-0.0228
52	Contrarian 3X Stock Index	-0.0209
50	S&P 500 Option Overwriting	-0.0039
54	Global	0.0023
42	Kinkopf S&P	0.0034
68	Systematic Alpha Futures	0.0044
2	Ag Trading	0.0052
32	Goldman Management Stock Index Futures	0.0079

Participation Ratio (PR)

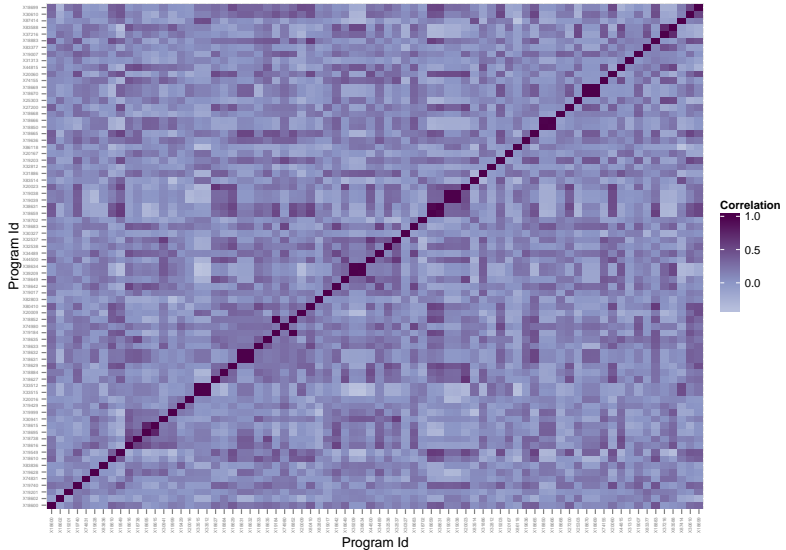


Determining Impact of Factors on Portfolio Variability

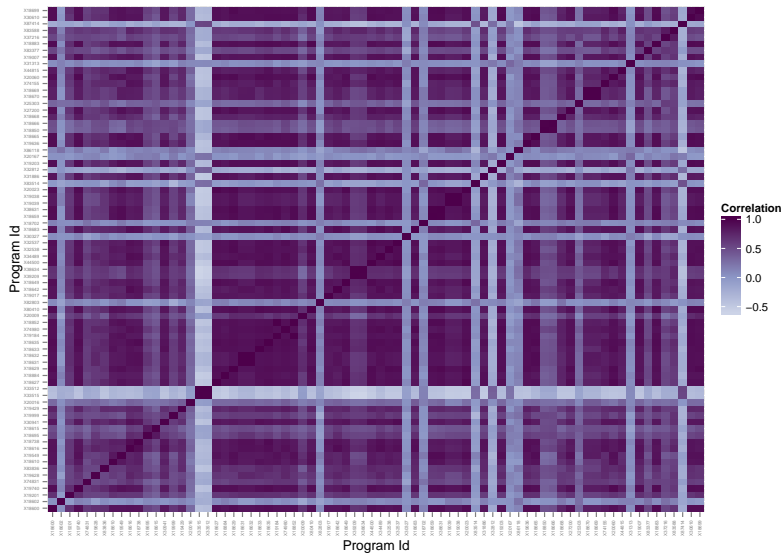
- Perturb 'importance of first factor' up and down
- Use equation for portfolio standard deviation to determine impact on portfolio



Scenario 1: Importance of First Factor Falls Hard - Way More Diversification



Scenario 1: Importance of First Factor Rises Hard - Way Less Diversification



Conclusions

- Factor analysis revealed a very significant proportion of the total variance of modeled managed futures universe can be captured by a single statistical factor
- First factor corresponds to a very intuitive scenario.
- Sensitivity analysis developed - given the intuitive interpretation of first factor - can be used to better understand variation in managed futures universe
- Can be used to identify programs that provide better diversification in different market regimes
- Will use my sophisticated versions of this to help my employer understand risk and return drivers of our \$3.5 billion fund

References

1. C. Bacon [2008], Practical Portfolio Performance Measurement and Attribution, 2nd Ed, John Wiley & Sons, Inc.
2. D. J. Fenn, N. F. Johnson, N. S. Jones, M. McDonald, M. A. Porter, S. Williams [2011], Temporal evolution of financial-market correlations, Physical Review E 84, 026109
3. F. J. Fabozzi, S. M. Focardi, P. N. Kolm [2010], Quantitative Equity Investing: Techniques and Strategies (Frank J. Fabozzi Series), John Wiley & Sons, Inc.
4. N. Fenton, M. Neil [2013], Risk Assessment and Decision Analysis With Bayesian Networks, CRC Press
5. D. Koller, N. Friedman [2009], Probabilistic graphical models: principles and techniques, MIT press.
6. A. Golub and Z. Guo [2012], Correlation Stress Tests Under the Random Matrix Theory: An Empirical Implementation to the Chinese Market
7. A Meucci [2009], Risk and Asset Allocation, 1st Ed, Springer Berlin Heidelberg

References - Continued

8. R. Rebonato [2010], Plight of the Fortune Tellers: Why We Need to Manage Financial Risk Differently, Princeton University Press
9. R. Rebonato [2010], Coherent Stress Testing: A Bayesian Approach to the Analysis of Financial Stress , John Wiley & Sons, Inc.
10. R. Rebonato and A. Denev [2014], Portfolio Management Under Stress: A Bayesian-net Approach to Coherent Asset Allocation, Cambridge University Press
11. D. Skillicorn [2007], Understanding Complex Datasets: Data Mining with Matrix Decompositions, Chapman and Hall/CRC
12. R. Vince [2007], The Handbook of Portfolio Mathematics: Formulas for Optimal Allocation and Leverage, John Wiley & Sons, Inc.

Appendix A: GitHub Repository

All of the R code used to produce this paper can be found in the following github repository:

https://github.com/dgn2/managed_futures

The github repository also includes the .Rmd file used to generate the .pdf working paper file and includes code to:

- ▶ extract CTA manager, program, and monthly return data from the Altegris managed futures website
- ▶ create a MySQL database with tables to store extracted CTA manager, program, and monthly return data
- ▶ load CTA manager, program, and monthly return data to the MySQL database
- ▶ conduct limited exploratory analysis of the data
- ▶ conduct limited cleaning of the data used in subsequent statistical modeling
- ▶ estimate statistical factors based on the monthly returns of a select set of CTA programs
- ▶ compute sensitivities

Contact Information

Thank you! Please do not use this Altegris data for commercial applications!

Derek G Nokes

`dgnokes@gmail.com`

`https://github.com/dgn2/managed_futures`