A Commodity Based Economic Model of Risk Factors Affecting the Returns of 11 Global Macro Hedge Fund Categories

L. Mick Swartz[1]

Farrokh Emami-Langroodi[2]

This paper has three contributions to the literature. First, it analyzes the risk and return characteristics for 11 previously unstudied Global Macro hedge fund sub-strategies. These sub-strategies are unique because most of these funds are not investing in stocks or bonds. These strategies invest in commodities so we have analyzed Commodity Based Macroeconomic Factors and models, CBEM. The previous research in academic finance has not examined these areas of asset pricing as much as stock investing. It is not obvious that the asset pricing research in finance or in traditional hedge fund models will be as significant in this area of investing. Therefore, the second contribution in this paper is the introduction of commodity based economic factors. Unlike Swartz and Emami-Langroodi (2018) this is not a behavioral model. Unlike Fung and Hsieh (2004) and Jurek and Stafford (2015) this is not a general model for all hedge fund categories. These new factors are commodity based and are not included in the traditional hedge fund models or the CAPM or the Fama-French models. This commodity based economic information, when included with asset pricing models (CAPM and Fama-French CAPM), is a more powerful method to explain Global Macro hedge fund returns than previous hedge fund models for this category of hedge funds. Using a pool of 20 macroeconomic variables, this study provides evidence that researchers should expand their use of other macroeconomic factors in their analyses of hedge fund returns. According to AIC and SIC criterion, the Global Macro category results for our Commodity Based Economic Model, CBEM, outperforms the traditional hedge fund models in all 11 Global Macro strategies. The empirical results indicate the commodity based specific models not only provide substantially more insight in terms of the risk in different global macro hedge fund strategies, but can also be structured to avoid econometric issues. The third contribution of the model is the advantage of avoiding econometric and modelling problems associated with the other strategies. The empirical results for Commodity Energy, Commodity Metals and Discretionary Thematic demonstrate dramatic differences advantages of our model versus the traditional hedge fund models. In particular the CRB index, Gold index, two types of credit spreads and other macroeconomic variables should be considered in future hedge fund studies.

Marshall School of Business, University of Southern California1

Goethe University Frankfurt, House of Finance2

Keywords: Global Macro, Hedge Fund, Factor Model, Risk Factor, Macroeconomic Financial Indicator

JEL Classifications: G11, G12, G17, G23, C51, C58

*I would like to thank Kenneth Ahern, Mikhail Chernov, Mike Cooper, Harry DeAngelo, Farrokh Emami, Wayne Ferson, Gerard Hoberg, and Gordon Philips for their insightful comments and suggestions. I would also like to thank the participants of the Shanghai National Futures Exchange Conference and China Derivatives Market Conference for their comments and suggestions.

1. Introduction

This paper draws on ideas from Fung and Hsieh (2004) and Jurek and Stafford (2015) regarding hedge fund returns. However, we differ in a few ways in terms of examining the return processes. First, we look more closely at the 11 Global Macro sub-strategies that have not been studied in the past literature. These strategies are interesting because many are not investing in stocks and bonds. The last fifty years of academic financial research has primarily studied stock and bond investing. These investing strategies are different in terms of factors used and the knowledge base in some ways is smaller. Second, we incorporate additional economic and financial information; some of these commodity based variables are more powerful at explaining global macro hedge fund returns than any variables in the previous studies mentioned. In particular, the CRB is considered the market portfolio of the commodity markets and is a factor in many models. This particular factor was described in Swartz and Emami-Langroodi (2018), however, this study differs from Swartz and Emami-Langroodi (2018) because it does not take a behavioral approach. The models developed in this paper are unique to every strategy. The commodity based economic models outperform the traditional hedge fund models in each of the 11 Global Macro Hedge Fund categories. The past performance of capital markets and past knowledge affects investor decision making, however, this paper examines macroeconomic variables, without a behavioral approach, and their impact on hedge fund returns for each Global Macro category.

A combined asset pricing model approach that incorporates the new commodity economic and financial factors is used and compared to traditional hedge fund models for each Global Macro category. The combined model (CBEM) commodity based economic and asset pricing model approach uses the CAPM and Fama-French with macroeconomic variables and this approach outperforms current hedge fund models in every category of Global Macro hedge funds. In addition, this combined approach leads to a more stable model with large econometric advantages over current hedge fund models.

In this paper we provide a comparison of the seven-factor model introduced in Fung and Hsieh (2004) (FH) paper using the data provided by the authors to recreate those seven-factor models for HFRX Macro indices (see Section 6) to show the advantages of incorporating the new variables and to show how specific category risk models have econometric advantages. In addition, this shows how versatile the CAPM and Fama-French structures are in explaining asset returns. The CEBM Commodity Based Economic Model is a different approach not only because of the individual variables but because the CRB index is the market for commodities. This changes the approach of the CEBM model versus the traditional hedge fund model that is based on stock factors and the S&P 500. The empirical results in this paper for the Commodity Energy, Commodity Metals and Discretionary Thematic categories demonstrate that there are huge differences between the FH approach and our approach. Along with the AIC and SIC criterion, the R^2 differences of 28% versus 1% (FH) and 75% versus 7% (FH) and 32% versus 11% (FH), for Energy, Metals and Discretionary categories respectively, indicate that our specific approach can generate much more insight into the actual risks in these categories. In addition, we avoid the econometric problems associated with those studies.

Competition will cause pension funds and large institutions to obtain average levels of returns in many hedge fund categories; therefore, it is critical for these investors to understand the financial and macroeconomic risks that drive the average returns in each hedge fund category. This paper is an explanation of the risk factors that affect each strategy of Global Macro hedge funds. Individual hedge fund returns are not as informative or useful for large institutions because, on average, these institutions will invest in multiple funds in a category and obtain average category returns and risk levels. What types of risk and what levels of risk do investors inherit, on average, when they invest in a Commodity fund or a Currency hedge fund or a Discretionary Thematic hedge fund or a Systematic Diversified hedge fund? These questions are studied in detail for eleven strategies in the HFRX Global Macro category.

This paper includes three types of models generated for each of 11 Macro hedge fund strategies. The first model is a recreation of a traditional hedge fund model, the Fung and Hsieh (2004) seven-factor model using the variables and data available in their database[1]. The second model is the CAPM with economic and financial variables. This model uses a selection of macroeconomic factors and the market index (i.e. the CAPM version). The third model incorporates and value premium and size premium (i.e. Fama-French version) with macroeconomic factors. The second and third model use a series of leading, lagging, and coincidental macroeconomic and financial indicators as explanatory variables as well as the market index (i.e. the CAPM version) and value premium and size premium (i.e. Fama-French

[1] The Fung & Hsieh factors data can be reached at: http://faculty.fuqua.duke.edu/~dah7/DataLibrary/TF-FAC.xls

version). Depending on unique nature of a hedge fund strategy, a specific model type can be more useful in explaining that strategy's return and risk factors.

We have used a pool of 20 *macroeconomic* and *financial* indicators in this paper such as the Commodity Research Bureau (CRB) index return, West Texas Intermediate (WTI) crude oil return, Gold index return, Copper index return, U.S. total balance of trades (BOT), U.S. Dollar index (DXY), U.S. ten-year Treasury bond total return, U.S. industrial production (INDPRO), U.S. personal consumption expenditure (PCE), Case/Shiller U.S. national housing price index, etc.

The approach taken in this paper is different than the seminal papers by Fung and Hsieh (2001, 2004, and 2011), Ammann, Huber, and Schmid (2011), and Jurek and Stafford (2015). Unlike previous papers, this paper examines all the Global Macro sub-strategies individually and does not focus on an aggregate hedge fund index. This approach is following the guidance for future research provided in Fung and Hsieh (2004) suggesting that additional risk factors would be needed for explaining narrower benchmarks and individual funds. It is not sufficient in this time period to classify all hedge funds as having the same level of risk as an aggregated hedge fund index. It is important to break down each category of hedge funds and explain what drives the returns and risks for each individual sub-strategy. Science should make models more precise not less precise.

The time period analyzed is January 1998 to December 2015 and we have utilized HFRX Macro hedge funds index returns instead of HFRI index which is mainly used in previous publications. As a result, our empirical results differ from previous studies regarding hedge fund returns. Capocci and Hubner (2004) document hedge fund outperformance versus the stock indices. Given the time period studied, the hedge fund average return performance in this paper show mixed results for hedge fund category returns outperforming stock indices.

This paper expands the importance of additional economic and financial variables. Unlike Jurek and Stafford (2015) this paper does not use an aggregate index, rather this paper uses 11 sub-strategies in Macro category of hedge funds with an asset pricing approach. Jurek and Stafford (2015) promote a similar line of research as Fung and Hsieh (2004), this paper promotes a different line of study that expands the use of macroeconomic variables or factors and examines the perspective that large institutional investors need to understand related to underlying risk factors for investing in these fund categories. Large institutions have risk management issues that must be addressed specifically for each type of investment. Auditing requires an advanced understanding of the actual risks involved in any investment. The cost of capital for funds should be related to specific risks, as well as correlations, in the return generating process. All of these tasks require a more thorough understanding and a more specific approach to the actual risks in these investment categories. In addition, this approach is not a behavioral approach as in Swartz and Emami-Langroodi (2018). The problems related to auditing, cost of capital, investment factors, and risk controls require that each category of hedge funds be analyzed relative to their own specific category and not relative to a general aggregate hedge fund index, as in most previous studies on hedge fund returns.

In this paper we provide a comparison of the seven-factor model introduced in Fung and Hsieh (2004) paper using the data provided by the authors to recreate those seven-factor models for HFRX Macro indices (see Section 6) to show the advantages of incorporating the new variables and to show how specific category risk models have econometric advantages.

By including insignificant explanatory variables in their models, previous studies by Fung and Hsieh (2004), and Jurek and Stafford (2015) cause the possible existence of time-series assumptions violation in some of their generated models. While Jurek and Stafford (2015) have addressed the autocorrelation issue in their models, they do not seem to have addressed the Conditional Heteroskedasticity and possible severe Multicollinearity issues.

Similarly, Fung and Hsieh (2004) have not addressed the above-mentioned time series issues, causing ten out of 11 and two out of 11 strategies regenerated seven-factor models suffer from Conditional Heteroscedasticity and Serial Correlation respectively which can eventually affect the stability and soundness of their estimated coefficients' signs and values (see *Supplemental Material*, Section 2). All of these time-series assumptions are tested and, if needed, corrected in our generated macroeconomic models in this paper.

We do not use the Option replication method as in Agarwal and Naik (2004) and Fung and Hsieh (2001) to estimate returns because the data for many individual hedge fund categories is not consistent with Put Option pricing concept during a crisis, thus, we do not implement the nine-factor model in Fung and Hsieh (2002) or in Jurek and Stafford (2015). The inconsistency of some of the hedge fund strategy returns with Option replication technique is demonstrated in Section 4 of this paper, by considering the monthly return of the hedge fund strategies during a crisis. This issue is also pointed out in the empirical results in Section 6. These results are consistent with the Fung and Hsieh (2011), however, we demonstrate this is true across Global Macro strategies, not just Long/Short Equity Hedge.

The empirical results will show that the economic models generated in this paper outperform the Fung and Hsieh seven-factor model, in 100% of cases, if we use Schwartz Information

Criterion (SIC) or if we use the Akaike Information Criterion (AIC), as the performance and model selection criteria.

Unlike the Fung and Hsieh (2004) seven-factor model, this paper finds that economic factors tend to be highly dominant in the resulting models. The Commodity Research Bureau (CRB) index return, and the Bullion Gold index return frequently outperform the all purpose hedge fund model. Also, this shows that based on their characteristics a majority of the Global Macro sub-strategies, cannot be explained using just one general model.

For the purpose of our analysis, we have collected and utilized the monthly rate of return data, from January 1998 to December 2015 time period, for the 11 Macro hedge fund strategies extracted from the Hedge Fund Research (HFR) Inc. HFRX indices. The return data for main strategies such as Macro/CTA covers the time period from January 1998 to December 2015.

The return data for Macro: Commodity-Energy covers the period from January 2007 to December 2015. The remaining sub-strategies data cover the period of January 2005 to December 2015.

To consider the macroeconomic effects, several indices are used in our regression analyses such as Standard & Poor’s 500 index (*S&P 500*) monthly total return, Commodity Research Bureau (*CRB*) index monthly total return, high grade Copper (*CU*) index monthly return, Gold Bullion (*XAU*) index monthly return, West Texas Intermediate (*WTI*) crude oil monthly return, U.S. Dollar Trade Weighted index (*DXY*) monthly return, US Dollar (USD) per Euros (*USD/EUR*) rate, USD per 100 Japanese Yen (*USD/JPY*) rate, U.S. government 10-year Treasury bond (T-bond) total return, U.S. Treasury bill (T-bill) total return, U.S. Industrial Production index (*INDPRO*) rate of change, U.S. Personal Consumption Expenditures (*PCE*) rate of change, U.S. Real PCE rate of change, U.S. Consumer Price Index (*CPI*) rate of change, and Chicago

Board Options Exchange (CBOE) Volatility Index (*VIX*). These data are extracted from “Global Financial Data” database.

The three-month and 12-month LIBOR rates based on U.S. Dollar, used as the target rate in calculating Sharpe, Sortino, and other unconventional performance ratios numerators, as well as U.S. ten-year constant maturity Treasury yield are collected from Federal Reserve Bank of St. Louis (FRED) database.

U.S. Total Balance of International Trade (*BOT*) rate of change data is collected from U.S. Census Bureau, Foreign Trade Division. U.S. monthly Unemployment rate (*URATE*) data is collected from U.S. Bureau of Labor Statistics (BLS).

For our Economic models, we have used two versions of the Credit Spread. The first version of credit spread, is called Long-term Credit Spread (*CRSPRD_L*) and is calculated as the difference of Moody’s Baa corporate bond yield and the U.S. government ten-year Treasury bond, collected from Federal Reserve Bank of St. Louis (FRED) database. The second version of credit spread is called Short-term Credit Spread (*CRSPRD_S*) and is calculated as the difference of Moody’s Baa corporate bond yield and the U.S. government three-month Treasury bill. The first version of the credit spread is also used for recreation of the Fung and Hsieh seven-factor models.

As mentioned in Section 1 of this paper, we create both CAPM and Fama-French controlled versions for the Economic models. Then, the version (i.e. CAPM or Fama-French) with higher performance and efficiency is selected as the representation of the corresponding Economic model. Therefore, for constructing the Fama-French controlled models, we add the Size premium and Value premium to the pool of independent variables.

For calculating the Size Premium (Small Cap minus Large Cap), represented by "*SML*", the Russell 2000 Total Return Index and S&P 500 Total Return Index are collected from "Global Financial Data" database and used for Small Cap and Large Cap stock returns respectively. Similarly, Russell 1000 Value Index and Russell 1000 Growth Index are collected from "Global Financial Data" database and used for calculating the Value Premium, High P/M (i.e. Value stock) minus Low P/M (i.e. Growth stock)), represented by "*HML*".

Table 2 summarizes the independent variables used in our generated models. A Combination of financial (CAPM and Fama-French) with economic factors, and the variables used in Fung and Hsieh (2004)[2] seven-factor model.

Table 1
Summary of Fung-Hsieh, CAPM and FF and Economic models' independent variables

Symbol	Independent Variable Description	Variable Model(s)
T10Y	Month end-to-month end change in the Federal Reserve's 10 yr. T-bond yield	Fung-Hsieh[2]

[2] The Fung & Hsieh factors data can be reached at: http://faculty.fuqua.duke.edu/~dah7/DataLibrary/TF-FAC.xls

Bd. Opt.	Return of a portfolio of lookback straddles on bond futures	Fung-Hsieh[2]
FX Opt.	Return of a portfolio of lookback straddles on currency futures	Fung-Hsieh[2]
Com. Opt.	Return of a portfolio of lookback straddles on commodity futures	Fung-Hsieh[2]
S&P 500	Standard & Poor’s 500 index monthly total return	Fung-Hsieh[2], STAT, ECON, COMBO
SML	Russell 2000 total return – S&P 500 total return	Fung-Hsieh[2], STAT, ECON, COMBO
CRSPR-L	Month end-to-month end change in the difference between Moody’s Baa yield and Federal Reserve’s 10 yr. T-bond yield	Fung-Hsieh[2], STAT, ECON, COMBO
CRB	CRB commodity index total return	STAT, ECON, COMBO
HML	Russell 1000 Value total return – Russell 1000 Growth total return	STAT, ECON, COMBO
SD	Standard Deviation of hedge fund return	STAT, COMBO
SKEW	Skewness of hedge fund return	STAT, COMBO

Table 2 (Continued)

Symbol	Independent Variable Description	Variable Model(s)
CRSPR-S	Month end-to-month end change in the difference between Moody’s Baa yield and Federal Reserve’s 3-mo T-bill yield	ECON, COMBO

T3MTR	U.S. government three-month treasury bill total return	ECON, COMBO
T10YTR	U.S. government ten-year treasury bond total return	ECON, COMBO
CPI	U.S. Consumer Price Index monthly rate of change	ECON, COMBO
WTI	West Texas Intermediate (WTI) oil index monthly rate of return	ECON, COMBO
XAU	Gold Bullion index monthly rate of return	ECON, COMBO
CU	High grade Copper index monthly rate of return	ECON, COMBO
INDPRO	U.S. Industrial Production index monthly rate of change	ECON, COMBO
URATE	U.S. monthly Unemployment Rate	ECON, COMBO
HOUSE	Case/Shiller U.S. national housing price index rate of change	ECON, COMBO
BOT	U.S. Total Balance of International Trade rate of change	ECON, COMBO
DXY	U.S. Dollar Trade Weighted Index rate of return	ECON, COMBO
USD/EUR	U.S. Dollar per Euros	ECON, COMBO
USD/JPY	U.S. Dollar per 100 Japanese Yen	ECON, COMBO
PCE	U.S. Personal Consumption Expenditures rate of change	ECON, COMBO
REAL_PCE	U.S. Real Personal Consumption Expenditures rate of change	ECON, COMBO
VIX	Chicago Board Options Exchange volatility index	ECON, COMBO

2. Macro strategies returns descriptive statistics and correlation analysis

Table 3 lists the descriptive statistics for the Macro hedge fund strategies. There are eleven sub-strategies listed.

The Macro–Top Level strategy is created by taking the arithmetic average of the HFRX ten Macro sub-strategies monthly returns for the period of January 1998 to December 2015 (i.e. 216 observations). The mean return per month for this strategy is approximately 0.60% (the highest in the category), with a 2.12% standard deviation and skew of 0.65, while for the same period, The Macro/CTA strategy has the mean monthly return of 0.43% with a 2.35% standard deviation and skew of 0.41.

For the period of January 2005 to December 2015, Active Trading strategy has a 0.45% mean monthly return with a volatility of 1.34% per month and skew of -0.60 (the most negative in the category), while for the same period, Commodity strategy has a 0.30% mean monthly return with a volatility of 2.02% per month and skew of 0.88 as the most positive skew in the category.

Commodity-Agriculture strategy has a mean return of 0.27% with 2.14% standard deviation, 5.90 kurtosis (2.90 excess kurtosis), and 1.04 skew for the period of January 2005 to December 2015.

Commodity-Metals strategy has 0.49% mean monthly return with 6.56% volatility (the highest in the category), 0.10 skew, and 3.14 kurtosis (0.14 excess kurtosis) as the lowest kurtosis in the category for the same period.

For the period of January 2007 to December 2015, Commodity-Metals strategy has a 0.21% per month return with a volatility of 3.32%, -0.59 skew, and 7.02 kurtosis (the highest in category).

For the period of January 2005 to December 2015, Currency strategy has a 0.04% mean monthly return (lowest in the category) with a volatility of 1.30% per month (lowest in category) and skew of -0.17. For the same period, the Discretionary Thematic strategy has the second highest monthly mean return of 0.50%, volatility of 2.33% per month, and -0.32 skewness.

Multi-Strategy index has a 0.42% per month return with a volatility of 1.63%, 0.28 skew, and a 4.38 kurtosis (1.38 excess kurtosis). Systematic Diversified strategy has a 0.36% mean return, 2.77% volatility, 4.02 kurtosis (1.02 excess kurtosis), and 0.63 skewness.

Table 2
Macro strategies returns descriptive statistics

Macro Index	Mean	Median	Std. Dev.	Kurtosis	Skewness	Min	Max	Count	Period
Macro–Top Level	0.60%	0.34%	2.12%	4.63	0.65	-7.05%	8.23%	216	Jan 1998 - Dec 2015
Macro/CTA	0.43%	0.31%	2.35%	4.63	0.41	-7.38%	8.54%	216	
Active Trading	0.45%	0.50%	1.34%	4.78	-0.60	-5.14%	3.78%	132	Jan 2005 - Dec 2015
Commodity	0.30%	-0.14%	2.02%	5.17	0.88	-4.36%	9.11%	132	
Commodity-Agriculture	0.27%	0.10%	2.14%	5.90	1.04	-4.16%	8.39%	132	
Commodity-Metals	0.49%	0.43%	6.56%	3.14	0.10	-17.24%	18.55%	132	
Commodity-Energy	0.21%	0.09%	3.32%	7.02	-0.59	-15.03%	10.07%	108	Jan 2007 - Dec 2015
Currency	0.04%	0.03%	1.30%	4.50	-0.17	-4.66%	4.15%	132	Jan 2005 - Dec 2015
Discretionary Thematic	0.50%	0.36%	2.33%	5.51	-0.32	-9.10%	6.31%	132	
Multi-Strategy	0.42%	0.36%	1.63%	4.38	0.28	-4.31%	6.23%	132	
Systematic Diversified	0.36%	0.09%	2.77%	4.02	0.63	-5.13%	11.56%	132	

Table 4 summarizes the correlation of each Macro sub-strategy monthly return with the S&P 500 market index return and performance ratios such as Sharpe ratio, Sortino ratio.

.

The market index return correlations for Macro categories vary from -23% for Systematic Diversified to 41% for Multi-Strategy, and is outperformed by other ratios correlations in ten out of 11 of the strategies.

Table 3
Macro strategies returns correlations with market benchmark & performance ratios

Macro Index	Market RoR	Sharpe ratio	Sortino ratio	Period
Macro–Top Level	19%	33%	36%	Jan 1998 - Dec 2015
Macro/CTA	10%	31%	34%	
Active Trading	8%	32%	37%	Jan 2005 - Dec 2015
Commodity	8%	40%	40%	
Commodity-Agriculture	14%	20%	17%	
Commodity-Metals	20%	27%	29%	
Commodity-Energy	22%	22%	25%	Jan 2007 - Dec 2015
Currency	15%	11%	3%	Jan 2005 - Dec 2015
Discretionary Thematic	37%	31%	26%	
Multi-Strategy	41%	17%	15%	
Systematic Diversified	-23%	18%	24%	

3. Hedge funds option-replication strategy analysis

Table 5 summarizes the monthly returns of Macro strategies during the 2008 financial crisis. The seven strategies (i.e. Macro-Top Level, Macro/CTA, Commodity, Commodity-Agriculture, Currency, Multi-Strategy, and Systematic Diversified strategies) with returns highlighted in bold are clearly not in any way related to Put-writing strategies. After examining volatility as a significant variable, only one out of eleven Macro strategies have standard deviation as significant factor, specifically this variable appears in the COMBO model of Systematic Diversified strategy. If the September to December cumulative returns during the financial crisis are examined for all eleven categories, then it is questionable that even one strategy is consistent with writing Put Options and/or Call Options. Writing put options during the financial crisis should have produced extremely large losses (greater than 50%). It is unrealistic to believe the entire hedge fund category could exit such positions with perfect timing in an illiquid options market without the average losses in excess of 50% in these types of positions.

While the data from 1997-1998 might show that these strategies behaved similar to writing put Options, the data from 2008 does not indicate this behavior. The cumulative monthly returns for September to December 2008 varied from approximately -15.10% for Commodity-Energy strategy to 15.85% for Systematic Diversified strategy. During the financial crisis, returns from Put-writing strategies should have had much lower returns. Returns from Put-writing would be in the negative range and probably greater than 100%, because most Put writing would occur for At-The-Money or Out-Of-The-Money options to obtain more leverage. It is unrealistic to believe an entire industry would be able to exit from the options markets without large losses. Likewise, if the strategies are consistent with writing Call Options, the returns would have been positive during this time period. The empirical data does not support either writing Put Options or writing Call Options or the replication of such strategies for almost all of these categories.

It is not obvious that a strategy with a -10% return, when the financial markets decline almost 50%, is equivalent to a Put-writing strategy. An investor would expect returns to be much lower if they were equivalent to writing Put Options. A breakdown of the hedge funds by categories finds that the approach of using a model that has Put-writing is not robust and consistent for at least seven out of eleven individual fund categories.

If Put-writing was a robust proxy variable in the individual models, the volatility would be statistically significant frequently. Option pricing models would demand that standard deviation to be a statistically significant variable if a strategy was similar to Put-writing. In this paper, standard deviation is not usually significant and it only appears as significant in one out of eleven strategies (i.e. Systematic Diversified strategy COMBO model). While the Put-writing approach seems to be successful for an aggregate index of hedge fund returns, it is not as useful when breaking the returns down by individual strategies.

Table 4

Macro strategies monthly rate of return and September to December cumulative return (Cum.) for the year 2008

Macro Index	**Jan-08**	**Feb-08**	**Mar-08**	**Apr-08**	**May-08**	**Jun-08**	**Jul-08**	**Aug-08**	**Sep-08**	**Oct-08**	**Nov-08**	**Dec-08**	***Cum.***
Macro–Top Level	**1.98%**	**6.35%**	**-2.50%**	**0.25%**	**1.55%**	**2.89%**	**-2.51%**	**-2.18%**	**-2.42%**	**-2.08%**	**0.66%**	**1.93%**	-1.91%
Macro/CTA	**3.82%**	**8.54%**	**-3.00%**	**-0.65%**	**1.77%**	**3.25%**	**-5.59%**	**-3.94%**	**-0.85%**	**-1.76%**	**1.48%**	**3.26%**	2.13%
Active Trading	0.12%	2.11%	0.22%	3.78%	1.28%	0.83%	2.22%	-0.25%	-5.14%	3.08%	1.08%	-0.22%	-1.20%
Commodity	**3.00%**	**9.11%**	**-1.29%**	**-0.42%**	**0.14%**	**2.70%**	**-2.01%**	**-0.57%**	**-0.12%**	**3.33%**	**-0.14%**	**0.56%**	3.63%
Comm.-Agriculture	**3.14%**	**6.06%**	**-2.11%**	**-0.87%**	**-1.44%**	**5.83%**	**-2.68%**	**-0.88%**	**-4.16%**	**-3.37%**	**-3.68%**	**1.69%**	-9.52%
Comm.-Metals	7.80%	7.97%	-7.64%	-5.50%	1.95%	2.04%	7.70%	-7.11%	-2.19%	-17.24%	8.36%	4.68%	-6.40%
Comm.-Energy	-5.85%	10.07%	-0.96%	4.28%	5.80%	7.20%	-15.03%	-1.80%	-8.41%	-4.98%	-1.32%	-0.39%	-15.10%
Currency	**-1.65%**	**0.30%**	**1.28%**	**1.28%**	**0.57%**	**0.52%**	**0.84%**	**-2.88%**	**-1.69%**	**0.50%**	**0.68%**	**0.83%**	0.32%
Discretionary Thematic	6.31%	4.79%	-5.75%	1.26%	2.36%	1.27%	-3.36%	-2.71%	-0.42%	-9.10%	-2.28%	3.35%	-8.45%
Multi-Strategy	**-0.09%**	**3.04%**	**-3.62%**	**1.45%**	**-0.21%**	**-0.92%**	**-2.09%**	**0.11%**	**-2.80%**	**1.94%**	**-1.99%**	**2.45%**	-0.39%
Systematic Diversified	**3.20%**	**11.56%**	**-2.13%**	**-2.11%**	**3.26%**	**6.20%**	**-5.13%**	**-1.80%**	**1.61%**	**6.77%**	**4.35%**	**3.12%**	15.85%

4. Model estimation, selection, and diagnostics methodology

Hedge fund strategy return models are estimated by Ordinary Least Square (OLS) technique or one of the ARCH family techniques if needed to overcome Conditional Heteroskedasticity issue. To generate robust models, each model is tested and corrected for the following time-series assumptions:

- Stationarity (tested by Augmented Dickey-Fuller Unit Root test, corrected by first differencing),
- Serial-Correlation (tested by Durbin-Watson test, corrected by first order Autoregressive term),
- Multi-Collinearity (tested by Variance Inflation Factor test, corrected by elimination),
- Conditional Heteroscedasticity (tested by ARCH Heteroscedasticity test, corrected by ARCH, GARCH, or EGARCH estimation).
- Heteroscedasticity (tested by White test, corrected by Newey–West HAC estimation).

In our model selection process we focus on creating the most parsimonious model. A parsimonious model is a model that accomplishes a desired level of explanation or prediction with as few predictor variables as possible.

Please refer to Appendix section of this article for detailed description of estimation procedure and diagnostics methodologies used for generating the presented models.

5. Empirical results

Table 6 lists the Macro strategies regression models. Each strategy has two types of models, namely, Fung and Hsieh seven-factor model and the CAPM or Fama-French version with our macroeconomic factors.

Each model is corrected for stationarity, serial correlation, severe multi-collinearity, conditional heteroscedasticity, and unconditional heteroscedasticity. The AIC and SIC information criteria are examined for final model selection, with the SIC chosen if there is a disagreement among the models as the model parsimony is main priority for model-selection in this analysis.

As mentioned before, the Macro–Top Level index is created by taking the arithmetic average of the ten HFRX Macro sub-strategies monthly returns for the period of January 1998 to December 2015 and is used as a benchmark index in the Macro category to compare the resulting significant variables in other sub-strategies.

The Macro–Top Level return model has two significant factors and outperforms the presented Fung and Hsieh seven-factor model, Statistical, and Economic models considering the adjusted R-squared (approximately 32.33%), AIC, and SIC as performance criteria. The significant variables include only economic variables such as CRB index return and Gold index return. These results show that this strategy is influenced by commodity related indicators. In this case, the CAPM and FF variables are not significant. In this category, the FH model has conditional heteroscedasticity problems. This is true in most categories.

Similar to the top level strategy, the Macro/CTA strategy return is driven by the macroeconomic factors, with the CRB and the 10 year treasury return comprising the best model. The CAPM and FF variables are not significant. The macroeconomic model is the best performing model even though the Fung and Hsieh seven-factor model has four significant variables. Unfortunately, the FH model has conditional heteroscedasticity issues.

The Active Trading strategy return has only the VIX as an explanatory variable. Again, both models have a weak R-squared. The FH has conditional heteroscedasticity issues again. This uniqueness may not lend itself to general analysis. This may be an advantage in terms of correlation with other portfolios they manage, however it may make it more difficult for some auditors and risk managers to assess what is actually contained in the portfolio.

The Commodity strategy shows the CRB, Gold and the short term credit spread are significant factors. As a result this strategy's performance is mainly driven by commodity and financial related factors. This outperforms the FH model again. This is the only category that FH does not have a problem with conditional heteroscedasticity. The fact that the option premium is significant in this situation points to a potential insight into the FH model. It is possible, that the option premium factor in the model was picking up conditional heteroscedasticity thru time.

The Commodity-Agriculture strategy model is the best performing model and has only the CRB as a significant variable. The FH model has problems with serial correlation and conditional heteroscedasticity.

The Commodity-Energy strategy model outperforms the FH model in this category considering the AIC and SIC criterion. The economic model has the CRB and the 3 month Treasury Return as significant factors. The R^2 is 28% while the FH R^2 is less than 1 %. The FH model has conditional heteroscedasticity problems again.

The Commodity-Metals strategy has five significant variables including the CRB, Gold, SML, the unemployment rate and the housing index. This model has an R^2 of 75% compared to the FH model with an R^2 of 7%. This model outperforms the Fung and Hsieh seven-factor model using the AIC and SIC criterion. The model has the significant variables including SML (size premium) as a Fama-French variable, as well as economic variables such as CRB index return, Gold index return, U.S. unemployment rate, and Case/Shiller national housing price index return. The FH model has problems with conditional heteroscedasticity again.

The Currency strategy model outperforms the FH model considering the SIC criterion. The only significant variable in the currency strategy model is the 3 month Treasury Return. The FH model has serial correlation and conditional heteroscedasticity issues again.

The Discretionary Thematic strategy model has two significant variables, the CRB index and the short-term Credit Spread. This model outperforms the FH model in terms of AIC and SIC. The R^2 is 32% versus the FH model with an R^2 of 11%. The FH model has conditional heteroscedasticity issues again.

The Global Macro Multi-Strategy model has two significant variables including the S&P 500 return and the unemployment rate. The FH model has a higher R2 of 26% versus 18% for our models, however the AIC and SIC both choose our model over the FH model. The FH model has conditional heteroscedasticity issues again. The CRB is not significant in this model.

The Systematic Diversified strategy model has the long term credit spread and gold as significant variables. This model outperforms the FH model according to AIC and SIC criterion.

The empirical results for Commodity Energy, Commodity Metals and Discretionary Thematic are huge differences between the FH approach and our approach. Along with the AIC and SIC criterion, the R^2 differences of 28% versus 1% (FH) and 75% versus 7% (FH) and 32% versus 11% (FH), for Energy, Metals and Discretionary categories respectively, indicate that our specific approach can generate much more insight into the actual risks in these categories.

It should be clarified that “Serial Correl.” row in the presented regression tables indicates if a model suffers from *uncorrected* Serial Correlation, and similarly, “Cond. Hetero.” row indicates if a model suffers from *uncorrected* Conditional or Unconditional Heteroskedasticity.

While ten out of 11 and two out of 11 regenerated Fung-Hsieh seven-factor models suffer from Conditional Heteroskedasticity and Serial Correlation respectively, all of our generated CAPM and Fama-French version of these economic models have been tested and corrected for stationarity, serial correlation, severe multi-collinearity, conditional heteroskedasticity, and heteroskedasticity, in order to guarantee the robustness and stability of the final generated models. Therefore, our presented models do not suffer from any of these issues.

The evidence of existence of Serial Correlation in the regenerated Fung-Hsieh seven-factor models are demonstrated in Section 2 of *Supplemental Material* document that accompanies this article for more information.

Table 5
Macro category regression models

MACRO TOP-LEVEL			
FH 7-FACTOR MODEL		ECONOMIC MODEL	
Variable	**Coefficient**	**Variable**	**Coefficient**
Intercept	−0.0121	*Intercept*	0.0016**
	(0.0087)		(0.0008)
S&P 500	**0.1209***	*XAU*	**0.1290***
	(0.0331)		(0.0179)
SML	**0.1319***	*CRB*	**0.0710***
	(0.0412)		(0.0195)
T10Y	**4.4778***		
	(1.4630)		
CRSPRD_L	1.6057		
	(2.4711)		
Bd. Opt.	0.0089		
	(0.0097)		
FX Opt.	**0.0139*****		
	(0.0082)		
Com. Opt.	**0.0182*****		
	(0.0100)		
Adj. R^2	13.96%		15.46%
AIC	−4.983		−5.485
SIC	−4.858		−5.391
Est. Method	OLS		GARCH
Serial Correl.	No		No
Cond. Hetero.	Yes		No
Period	Jan. 1998 – Dec. 2015		

Significance Level: *** 10% ** 5% * 1%

Table 6 (Continued)

		MACRO/CTA	
FH 7-FACTOR MODEL		ECONOMIC MODEL	
Variable	**Coefficient**	**Variable**	**Coefficient**
Intercept	−0.0109	*Intercept*	0.0008
	(0.0100)		(0.0010)
S&P 500	**0.0937****	*CRB*	**0.0453****
	(0.0379)		(0.0219)
SML	**0.1107****	*T10YTR*	**0.1402***
	(0.0472)		(0.0538)
T10Y	**4.0561****		
	(1.6752)		
CRSPRD_L	1.0400		
	(2.8296)		
Bd. Opt.	0.0145		
	(0.0111)		
FX Opt.	0.0119		
	(0.0094)		
Com. Opt.	**0.0232****		
	(0.0115)		
Adj. R^2	8.46%		3.11%
AIC	−4.712		−5.012
SIC	−4.587		−4.902
Est. Method	OLS		EGARCH
Serial Correl.	No		No
Cond. Hetero.	Yes		No
Period	Jan. 1998 – Dec. 2015		

Significance Level: *** 10% ** 5% * 1%

Table 6 (Continued)

MACRO: ACTIVE TRADING				
FH 7-FACTOR MODEL			ECONOMIC MODEL	
Variable	**Coefficient**		**Variable**	**Coefficient**
Intercept	−0.0040		*Intercept*	0.0045*
	(0.0073)			(0.0013)
S&P 500	**0.0583*****		*VIX*	**−0.0118****
	(0.0324)			(0.0062)
SML	−0.0623			
	(0.0500)			
T10Y	**3.1769****			
	(1.5989)			
CRSPRD_L	−0.1317			
	(1.8659)			
Bd. Opt.	−0.0068			
	(0.0092)			
FX Opt.	**0.0116*****			
	(0.0070)			
Com. Opt.	0.0080			
	(0.0080)			
Adj. R^2	5.79%			2.82%
AIC	−5.792			−5.805
SIC	−5.618			−5.761
Est. Method	OLS			OLS-HAC
Serial Correl.	No			No
Cond. Hetero.	Yes			No
Period	Jan. 2005 – Dec. 2015			

Significance Level: *** 10% ** 5% * 1%

Table 6 (Continued)

MACRO: COMMODITY			
FH 7-FACTOR MODEL		ECONOMIC MODEL	
Variable	**Coefficient**	**Variable**	**Coefficient**
Intercept	−0.0134	*Intercept*	0.0026
	(0.0106)		(0.0019)
S&P 500	0.0678	*CRB*	**0.1095***
	(0.0467)		(0.0408)
SML	0.0793	*CRSPRD_S*	**22.2668***
	(0.0721)		(7.9059)
T10Y	**6.2483***	*XAU*	**0.0907***
	(2.3093)		(0.0344)
CRSPRD_L	−0.1236		
	(2.6950)		
Bd. Opt.	−0.0049		
	(0.0132)		
FX Opt.	0.0165		
	(0.0101)		
Com. Opt.	**0.0278****		
	(0.0116)		
Adj. R^2	13.72%		19.62%
AIC	−5.057		−5.160
SIC	−4.882		−5.072
Est. Method	OLS		OLS-HAC
Serial Correl.	No		No
Cond. Hetero.	No		No
Period	Jan. 2005 – Dec. 2015		
Significance Level: *** 10% ** 5% * 1%			

Table 6 (Continued)

MACRO: COMMODITY-AGRICULTURE			
FH 7-FACTOR MODEL		ECONOMIC MODEL	
Variable	**Coefficient**	**Variable**	**Coefficient**
Intercept	0.0020	*Intercept*	0.0030**
	(0.0116)		(0.0015)
S&P 500	**0.1041****	*CRB*	**0.0603****
	(0.0514)		(0.0298)
SML	−0.0082		
	(0.0793)		
T10Y	1.6532		
	(2.5384)		
CRSPRD_L	−1.8744		
	(2.9624)		
Bd. Opt.	−0.0057		
	(0.0146)		
FX Opt.	0.0164		
	(0.0111)		
Com. Opt.	**0.0302****		
	(0.0127)		
Adj. R^2	6.99%		6.12%
AIC	−4.868		−5.012
SIC	−4.693		−4.903
Est. Method	OLS		GARCH
Serial Correl.	Yes		No
Cond. Hetero.	Yes		No
Period	Jan. 2005 – Dec. 2015		
Significance Level: *** 10% ** 5% * 1%			

Table 6 (Continued)

MACRO: COMMODITY-ENERGY			
FH 7-FACTOR MODEL		ECONOMIC MODEL	
Variable	**Coefficient**	**Variable**	**Coefficient**
Intercept	0.0041	*Intercept*	0.0052***
	(0.0189)		(0.0028)
S&P 500	**0.1730****	*CRB*	**0.2627***
	(0.0852)		(0.0493)
SML	−0.1446	*T3MTR*	**36.4571***
	(0.1408)		(14.3727)
T10Y	3.1019		
	(4.6810)		
CRSPRD_L	−4.2868		
	(5.0429)		
Bd. Opt.	0.0134		
	(0.0249)		
FX Opt.	−0.0132		
	(0.0200)		
Com. Opt.	0.0089		
	(0.0236)		
Adj. R^2	0.80%		28.55%
AIC	−3.909		−4.294
SIC	−3.711		−4.219
Est. Method	OLS		OLS
Serial Correl.	No		No
Cond. Hetero.	Yes		No
Period	Jan. 2007 – Dec. 2015		
Significance Level: *** 10% ** 5% * 1%			

Table 6 (Continued)

MACRO: COMMODITY-METALS			
FH 7-FACTOR MODEL		ECONOMIC MODEL	
Variable	**Coefficient**	**Variable**	**Coefficient**
Intercept	−0.0754**	*Intercept*	−0.0004
	(0.0357)		(0.0026)
S&P 500	**0.3956****	*XAU*	**0.8971***
	(0.1573)		(0.0723)
SML	0.2862	*CRB*	**0.2253***
	(0.2429)		(0.0759)
T10Y	**21.6997***	*SML*	**0.3150****
	(7.7744)		(0.1473)
CRSPRD_L	10.2759	*URATE*	**4.1259***
	(9.0731)		(1.528)
Bd. Opt.	0.0597	*HOUSE*	**−1.4112****
	(0.0446)		(0.6082)
FX Opt.	−0.0233		
	(0.0340)		
Com. Opt.	0.0423		
	(0.0389)		
Adj. R^2	7.51%		75.60%
AIC	−2.629		−3.972
SIC	−2.455		−3.841
Est. Method	OLS		OLS-HAC
Serial Correl.	No		No
Cond. Hetero.	Yes		No
Period	Jan. 2005 – Dec. 2015		
Significance Level: *** 10% ** 5% * 1%			

Table 6 (Continued)

MACRO: CURRENCY			
FH 7-FACTOR MODEL		ECONOMIC MODEL	
Variable	**Coefficient**	**Variable**	**Coefficient**
Intercept	−0.0018	*Intercept*	0.0014
	(0.0073)		(0.0010)
S&P 500	**0.0590*****	*T3MTR*	**11.2380****
	(0.0324)		(5.1817)
SML	0.0341		
	(0.0499)		
T10Y	0.5023		
	(1.6002)		
CRSPRD_L	0.2885		
	(1.8675)		
Bd. Opt.	0.0006		
	(0.0092)		
FX Opt.	0.0110		
	(0.0070)		
Com. Opt.	−0.0018		
	(0.0080)		
Adj. R^2	0.33%		2.76%
AIC	−5.791		−5.921
SIC	−5.616		−5.833
Est. Method	OLS		GARCH
Serial Correl.	Yes		No
Cond. Hetero.	Yes		No
Period	Jan. 2005 – Dec. 2015		

Significance Level: *** 10% ** 5% * 1%

Table 6 (Continued)

MACRO: DISCRETIONARY THEMATIC			
FH 7-FACTOR MODEL		ECONOMIC MODEL	
Variable	**Coefficient**	**Variable**	**Coefficient**
Intercept	0.0053	*Intercept*	0.0042*
	(0.0124)		(0.0012)
S&P 500	**0.1823***	*CRB*	**0.1347***
	(0.0547)		(0.0272)
SML	0.0539	*CRSPRD_S*	**−18.6058***
	(0.0844)		(6.2068)
T10Y	1.2325		
	(2.7016)		
CRSPRD_L	−2.1888		
	(3.1529)		
Bd. Opt.	−0.0079		
	(0.0155)		
FX Opt.	−0.0027		
	(0.0118)		
Com. Opt.	0.0138		
	(0.0135)		
Adj. R^2	11.35%		32.69%
AIC	−4.743		−5.316
SIC	−4.569		−5.184
Est. Method	OLS		GARCH
Serial Correl.	No		No
Cond. Hetero.	Yes		No
Period	Jan. 2005 – Dec. 2015		

Significance Level: *** 10% ** 5% * 1%

Table 6 (Continued)

MACRO: MULTI-STRATEGY			
FH 7-FACTOR MODEL		ECONOMIC MODEL	
Variable	**Coefficient**	**Variable**	**Coefficient**
Intercept	−0.0140***	*Intercept*	0.0025**
	(0.0079)		(0.0011)
S&P 500	**0.1961***	*S&P 500*	**0.1790***
	(0.0349)		(0.0257)
SML	0.0158	*URATE*	**1.6868***
	(0.0539)		(0.6304)
T10Y	**3.1975*****		
	(1.7240)		
CRSPRD_L	**3.8241*****		
	(2.0120)		
Bd. Opt.	−0.0145		
	(0.0099)		
FX Opt.	**0.0248***		
	(0.0075)		
Com. Opt.	0.0117		
	(0.0086)		
Adj. R^2	26.32%		18.19%
AIC	−5.642		−5.667
SIC	−5.467		−5.535
Est. Method	OLS		GARCH
Serial Correl.	No		No
Cond. Hetero.	Yes		No
Period	Jan. 2005 – Dec. 2015		
Significance Level: *** 10% ** 5% * 1%			

Table 6 (Continued)

MACRO: SYSTEMATIC DIVERSIFIED					
FH 7-FACTOR MODEL				ECONOMIC MODEL	
Variable	**Coefficient**			**Variable**	**Coefficient**
Intercept	−0.0103			*Intercept*	0.0022
	(0.0143)				(0.0019)
S&P 500	−0.0171			*CRSPRD_L*	**39.5040***
	(0.0630)				(12.6784)
SML	−0.0308			*XAU*	**0.0986****
	(0.0973)				(0.0441)
T10Y	**5.4218*****				
	(3.1131)				
CRSPRD_L	0.7020				
	(3.6331)				
Bd. Opt.	**0.0495***				
	(0.0179)				
FX Opt.	0.0174				
	(0.0136)				
Com. Opt.	**0.0309****				
	(0.0156)				
Adj. R^2	16.99%				11.54%
AIC	−4.460				−4.517
SIC	−4.285				−4.385
Est. Method	OLS				GARCH
Serial Correl.	No				No
Cond. Hetero.	Yes				No
Period	Jan. 2005 – Dec. 2015				
Significance Level: *** 10% ** 5% * 1%					

Note. This table represents four estimated models for each strategy, Fung & Hsieh seven-factor (FH 7-Factor), Statistical, Economic, and COMBO Models. The significance level is 5%. The variables in FH 7-Factor model are included regardless of their significance to recreate the results of Fung & Hsieh (2004). Akaike (AIC) and Schwarz (SIC) information criteria are used for model comparison in each strategy with lower values desired. The dependent variable is the corresponding hedge fund strategy monthly return. Values in parentheses are the Std. Errors. The “Serial Correl.” and “Cond. Hetero.” refer to a model that is suffering from *uncorrected* Serial Correlation and *uncorrected* Conditional Heteroskedasticity respectively. See Table 2 for independent variables.

According to the AIC and the SIC criterion, the Global Macro category results for our economic and financial models outperform the Fung and Hsieh (2004) seven-factor model in all 11 strategies (i.e. 100% outperformance). In addition, 10 of the 11 FH models have problems with conditional heteroscedasticity and a few other FH models have serial correlation issues. None of our models have serial correlation or conditional heteroscedasticity issues or multicollinearity issues.

Table 7 demonstrates the most dominant significant factors percentage out of 11 Macro strategies. It can be observed that three factors tend to be highly dominant in all the resulting models. The highly dominant factors include the Commodity Research Bureau (CRB) index return, and the Bullion Gold index return, and the credit spread (either long or short term). The 10 year and 3 month Treasury returns also appear to be useful. This shows that based on characteristics of majority of the Macro category sub-strategies, the combinations of commodity related indicators are crucial for providing the highest level of explanation and influence for the strategies returns in each category.

Comparing these dominant factors with the significant variables appeared in Macro-Top Level index model we can see that the above-mentioned highly and moderately dominant factors are present in this index and are repeated among the other eleven strategies frequently.

Table 6
Dominant significant model factors among eleven macro strategies

Model Type	Dominant Significant Factors	% Significant out of 11 Macro Strategies
ECON,	*CRB*	63%
ECON,	*Gold*	36%
ECON,	Credit Spread	27%

6. Conclusion

This paper examines 11 individual hedge fund strategies in the HFRX Macro category that have never been presented in the previous hedge fund literature. The obtained models are controlled for CAPM and Fama-French variables, however since these hedge fund categories do not typically invest in stocks and bonds very much the S&P 500 or beta is not usually significant. For some categories the size or value premium may be useful. What is significant is the CRB index. The CRB index can be thought of as the "market portfolio" in the commodity markets. The models used in this paper incorporate the CRB as the market or beta in traditional terms. The Fama-French model is included by including the size and value premiums, however, those do not usually add explanatory power and do not show up in the final models. The models presented in this paper are specific to each category and outperform the Fung and Hsieh (2004) seven-factor models 100% of the time, based on AIC and SIC criterion.

The empirical results for Commodity Energy, Commodity Metals and Discretionary Thematic indicate that there are huge differences between the FH approach and our approach. Along with the AIC and SIC criterion, the adjusted R^2 differences of 28% versus 1% (FH) and 75% versus 7% (FH) and 32% versus 11% (FH), for Energy, Metals and Discretionary categories respectively, indicate that our specific approach can generate much more insight into the actual risks in these categories.

Using a pool of 20 macroeconomic variables, this study provides evidence that researchers should expand their use of other macroeconomic factors in their analyses of Global Macro hedge fund returns.Therefore, this paper has shown that depending on the characteristic of a hedge fund strategy, specific macroeconomic/financial variables for each hedge fund category can result in an improved factor model as opposed to using only one group of the variables for every hedge fund category.

This paper also shows that most hedge fund strategies are not simulations of Put Option writing, although a few categories do seem to exhibit characteristics that could be mimicked using Option replication techniques.

For future research in the field of hedge fund returns, the CRB commodity index and the Gold index return should be analyzed. Other additional factors to be considered are two forms of the credit spread, one of which is included in the FH models, and other important financial and economic factors include the unemployment rate, housing index and the ten year and 3 month Treasury returns. Future researchers should consider expanding their scope of economic and financial variables.

APPENDIX

Model estimation, selection, and diagnostics methodology

Based on the Gauss-Markov theorem we check the following assumptions to see whether Ordinary Least Squares (OLS) estimated coefficients are Best Linear Unbiased Estimators (BLUE) or not. Here "best" means giving the lowest variance of the estimate, as compared to other unbiased, linear estimators. The errors do not need to be normal, nor do they need to be independent and identically distributed (only uncorrelated with mean zero and homoscedastic with finite variance).

The requirement that the estimator be unbiased cannot be dropped, since biased estimators exist with lower variance. If these assumptions are violated, then it may be that OLS estimators are no longer “unbiased” or “efficient”. That is, they may be inaccurate or subject to fluctuations between samples.

Assumption (1): $E(\varepsilon_t) = 0$: Expected value of residual error is zero. If this assumption is not satisfied, the Intercept parameter will be biased, but there will be no extreme effect on other parameters.

Assumption (2): $Var(\varepsilon_t) = \sigma_\varepsilon^2 < \infty$: i.e. the variance is constant which is Homoskedasticity assumption, if the errors do not have a constant variance they are said to be heteroskedastic. This assumption is specifically important for cross-section data. If the errors are heteroskedastic, the coefficient estimates would still be the “correct” (assuming that the other assumptions required to demonstrate OLS optimality are satisfied), but the problem would be that the standard errors could be wrong. Therefore, if we were trying to test the hypotheses about the true parameter values, we could end up drawing the wrong conclusions. In fact, for all of the variables except the constant, the standard errors would typically be too small, so that we would end up rejecting the null hypothesis too many times. We have tried to address the unconditional Heteroskedasticity issue by using the Heteroskedasticity robust standard errors which correct for the problem by enlarging the standard errors relative to what they would have been for the situation where the error variance is positively related to one of the explanatory variables. To implement this technique, HAC (Heteroskedasticity and Autocorrelation Consistent) covariance matrix estimation (i.e. Newey–West estimator) is used to provide an estimate of the covariance matrix of the parameters of a regression-type model when this model is applied in situations where the standard assumptions of regression analysis do not apply. The estimator is used to try to overcome autocorrelation, or correlation, and Heteroskedasticity in the error terms in the models. This is often used to correct the effects of correlation in the error terms in regressions applied to time series data. Since we are dealing with time-series data, we give a higher priority to Conditional Heteroskedasticity issue in our residuals and in case of existence of this issue we use Autoregressive Conditional Heteroskedasticity (ARCH) estimation techniques instead of OLS. These techniques, depending on the order of Heteroskedasticity and existence of sign or size bias in under-study data, can vary and in our analysis they include estimation methods such as ARCH, GARCH (Generalized ARCH), or EGARCH (Exponential GARCH).

Assumption (3): $E(\varepsilon_t.\varepsilon_{t-1}) = 0$: It is assumed that the errors are uncorrelated with one another, otherwise there would be Autocorrelation (Serial Correlation). We want our residuals to be random, and if there is evidence of autocorrelation in the residuals, then it implies that we could predict the sign of the next residual and get the right answer more than half the time on average. This assumption is specifically important for time-series data. If this assumption is violated, there would be Autocorrelation (Serial Correlation) among the residuals. Then the value of estimated coefficient is Unbiased but, it is Inefficient meaning that the Standard Error is unknown, so, performing the t-test calculation and hence checking the significance of the coefficients would not be possible. If the form of the Autocorrelation is known, it would be possible to use a GLS procedure. One approach, which was once fairly popular and is used in addressing the autocorrelation issue in our models, is known as the Cochrane--Orcutt procedure. Such method works by assuming a particular form for the structure of the autocorrelation (usually a first order autoregressive, AR(1), process). Existence of Autocorrelation in our estimated models is tested by Durbin–Watson (DW) statistic and checking the existence of positive or negative serial correlation by considering the critical DW values as a test for first order autocorrelation. If the DW statistic is substantially less than 2, there is evidence of positive serial correlation. As a rough rule of thumb, if DW is less than 1.0, there may be cause for alarm. Small values of DW indicate successive error terms are, on average, close in value to one another, or positively correlated. If DW is greater than 2, successive error terms are, on average, much different in value from one another, i.e., negatively correlated.

Assumption (4): Nonexistence of severe Multi-Collinearity between independent variables: This assumption is violated due to very high correlation among independent variables. Some statistical errors, caused by violation of this assumption that we can refer to include, inconsistent regression statistics and/or inconsistent signs of coefficients. This is where the individual repressors are very closely related, so that it becomes difficult to disentangle the effect of each individual variable upon the dependent variable. This causes the estimated coefficients to be Biased, Inefficient and Inconsistent. We test the existence of severe multi-collinearity by performing coefficients diagnostics test of Variance Inflation Factor (VIF) which quantifies the severity of multi-collinearity in an ordinary least squares regression analysis. It provides an index that measures how much the variance (the square of the estimate's standard deviation) of an estimated regression coefficient is increased because of collinearity. In our analyses, the cut-off value of VIF = 10 is used as a [rule of thumb] critical value for existence of severe multi-collinearity. Solving the severe multi-collinearity issue is addressed by dropping one or some of the highly collinear variables, if possible, or by transforming the highly correlated variables into a ratio and include only the ratio and not the individual variables in the regression.

Assumption (5): Stationary Variables: A time-series variable is Stationary if its mean, variance and Covariance are stationary. We check the Stationary assumption using following methods:

1) Checking the existence of any kind of trend (upward or downward) or any kind of evidence to show non-constant mean or variance in the variable graph.
2) Corrologram: As a sign of Non-Stationary data, the Partial Autocorrelation Function's (PACF) first lag should be significant with a value close to 1 while the rest of the lags are insignificant or much smaller than 1, and Autocorrelation Function (ACF) should show numerous significant lags that are gradually decreasing in value.
3) Augmented Dickey-Fuller Unit Root Test: the Null Hypothesis for this test states that the under-study variable has a unit root, i.e. it is Non-Stationary. By checking the P-value of this test, we can decide whether reject the null or fail to reject the null hypothesis for confidence level $\alpha = 5\%$.

Since we are utilizing the Rate of Return (%) as our dependent variable, and after performing the above-mentioned methods, no non-stationarity is observed in our under-analysis dependent variables. In specific situations that we used some of the risk factors such as Skewness or Standard Deviation as our dependent variables, the non-stationary behavior is observed in which case the estimation is performed on the first difference transformation of the under-study variables.

For the purpose of model selection, among possible candidate reliable models, we have utilized the information criteria, AIC and SIC, as the basis of our judgement.

Let:

- n = number of observations (e.g. data values, frequencies)
- k = number of parameters to be estimated (e.g. the Normal distribution has 2: μ and σ)
- L_{max} = the maximized value of the log-Likelihood for the estimated model (i.e. fit the parameters by Maximum Likelihood Estimation (MLE) and record the natural log of the Likelihood.)

SIC (Schwarz Information Criterion):

$$SIC = ln[n]\,k - 2ln[L_{max}] \tag{15}$$

AIC (Akaike Information Criterion):

$$AIC = \left(\frac{2n}{n-k-1}\right)k - 2\,ln[L_{max}] \tag{16}$$

The aim is to find the model with the lowest value of the selected information criterion. The ($-\,2ln[L_{max}]$) term appearing in each formula is an estimate of the deviance of the model fit. The coefficients for k in the first part of each formula show the degree to which the number of model parameters is being penalized. For $n > \sim 20$ or so the SIC (Schwarz, 1997) is the strictest in penalizing loss of degree of freedom by having more parameters in the fitted model. For $n > \sim 40$ the AIC (Akaike, 1974) is the least strict of the two.

In most cases, we prefer the model that has the fewest parameters to estimate, provided that each one of the candidate models is correctly specified. This is called the most parsimonious model of the set. The AIC does not always suggests the most parsimonious model, because the

AIC function is largely based on the log likelihood function. Davidson and MacKinnon (2004) indicates that whenever two or more models are nested, the AIC may fail to choose the most parsimonious one, if that these models are correctly specified. In another case, if all the models are non-nested, and only one is well specified, the AIC chooses the well-specified model asymptotically, because this model has the largest value of the log likelihood function.

The SIC avoids the problem discussed above by replacing $2n/(n - k - 1)$ in the AIC function with the $ln(n)$ term. As $n \to \infty$, the addition of another lag would increase the SIC value by a larger margin. Hence, asymptotically, SIC would pick the more parsimonious model than AIC might suggest.

References

Agarwal, V., Naik, N., 2004. Risks and portfolio decisions involving hedge funds. Review of Financial Studies 17:63–98.

Akaike, H. 1974. A new look at the statistical model identification. System identification and time-series analysis. IEEE Transaction on Automatic Control AC-19:716–723.

Ammann, M., Huber, O.R., Schmid, M.M., 2011. Benchmarking hedge funds: the choice of the factor model. Available at SSRN 1672543.

Capocci, D., Hübner, G., 2004. Analysis of hedge fund performance. Journal of Empirical Finance 11:55–89.

Cohen, G. and A. Kudryartsev, 2012, Investor Rationality and Financial Decisions. Journal of Behavioral Finance 2012 Vol 13 Issue 1

Davidson, R., MacKinnon, J.G., 2004. Econometric Theory and Methods. New York: Oxford University Press.

Ferson, W.E., Schadt, R.W., 1996. Measuring fund strategy and performance in changing economic conditions. Journal of Finance 51:425–461.

Filbeck, G., P. Hatfield, and P. Horvath. Risk Aversion and Personality Type. Journal of Behavioral Finance 2005 Vol 6 Issue 4

Fung, W., Hsieh, D., 2011. The risk in hedge fund strategies: theory and evidence from long/short equity hedge funds. Journal of Empirical Finance 18:547–569.

Fung, W., Hsieh, D., 2004. Hedge fund benchmarks: a risk-based approach. Financial Analyst Journal 60:65–80.

Fung, W., Hsieh, D., 2002. The risk in fixed-income hedge fund styles. Journal of Fixed Income 12:6–27.

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

Jurek, J.W., Stafford, E., 2015. The cost of capital for alternative investments. Journal of Finance 70:2185–2226.

Keller, C. and M. Siegrist. 2006. Money Attitude Typology and Stock Investment. Journal of Behavioral Finance Vol 7 2006 Issue 2

Schwartz, E.S., 1997. The stochastic behavior of commodity prices: implications for valuation and hedging. Journal of Finance 52:923–973.

Semmler, W., 2011. Asset Prices, Booms and Recessions Financial Economics from a Dynamic Perspective, fourth ed. Springer, New York.