Systematic and Discretionary Hedge Funds: Classification and Performance Comparison

Hui-Ching CHUANG
College of Management, Yuan Ze University
Chung-Ming KUAN
Department of Finance & CRETA, National Taiwan University

Workshop 3: Asset Pricing and Risk Management
Institute for Mathematical Sciences, NUS
Aug. 28, 2019

Outline

1. Introduction
2. Fund Classification
   - Classifier Training and Prediction
3. Fund Performance and Comparison
   - Testing Significance of Fund Performance
   - Testing Stochastic Dominance
4. Concluding Remarks

Machine vs. Man

MAN VS. MACHINE

Man Vs. Machine: Editor's Introduction

By: Bob Pisani

8:55 AM ET Tue, 14 Sept 2010

There's a whole new universe of new places to trade and new players doing the trading. How has it changed, and why? What does it mean for you?

» Read More

RagingBull RAGINGBULL INSIDER OUR TRAINERS FREE OPTIONS HANDBO

Man vs. Machine and The Future of Trading

Jeff Bishop · May 4, 2017 at 8:42 PM

I'll keep this short and sweet. Trading stocks that have news in them is not easy. After all,

Forbes Billionaires Innovation Leadership Money Consumer Industry Lifestyle BrandVoice

2,603 views | Jun 26, 2018, 10:15am

Man Vs. Machine: Credit Suisse Finds Renewed Value In Human Traders

We Are Interested in

- Classify hedge funds into systematic or discretionary
  - Model-based "systematic" funds vs. human-based "discretionary" funds (Harvey et al., 2017)
    - Similar classifications: "quantitative" vs. "qualitative"("discretionary" ,"fundamental"), and "machine" vs. "man"(Chincarini,2014; Abis, 2018; Evans et al.,2018)
- Evaluate the performance of classified funds
  - Is fund performance due to authentic skills or sampling luck?
  - Do systematic funds (as a group) outperform discretionary funds?

This Paper

- A new approach to classifying funds
  - Textual analysis is applied to convert text of investment strategies into numeric data and extract "features" from such data.
  - Statistical learning methods are invoked to determine a "classifier" based on the extracted features: regressions, LDA, KNN, SVM, classification tree, random forest, etc. (e.g., Hastie, et al., 2009).
  - Our approach captures strategy similarities and avoids subjective judgement or choice of keywords (cf. Harvey et al., 2017; Abis 2018).
- Evaluating fund performance
  - Bootstrap analysis of Kosowski et al. (2006, JF): Testing significance of the factor-adjusted returns at different quantiles.
  - Stochastic dominance test of Linton et al. (2010, JoE): Testing the ordering between two distributions of the factor-adjusted returns.

Related Researches

- Chincarini (European Financial Management, 2014)
  - uses (i) the classification of Hedge Fund Research (HFR) to bifurcate funds or (ii) word count: *algorithm, automate, econometric, mathematical, model, quantitative, statistic*
  - quantitative hedge funds have higher alphas than qualitative ones.
- Harvey et al. (J. of Portfolio Management, 2017)
  - word count approach. *algorithm, approx, computer, model, statistical,and system* are keywords used in their paper.
  - performances are similar.
- Abis (working paper, 2018)
  - collected 2,607 mutual funds' "Principal Investment Strategies" in prospectuses from SEC.
  - classified manually a sub-sample of 200 prospectuses into two types.
  - apply machine learning methods to 200 (training sample) to classify the reaming funds (prediction sample).
  - compare stock picking/timing and holding performance and justify her empirical findings by a theoretical model.

HFR Classification

- As Harvey et al. (2017), we only consider two main strategies (Equity Hedge and Macro) and their six sub-strategies in HFR.

Equity Hedge	Macro
Equity Market Neutral	Active Trading
Quantitative Directional	Commodity: Metals
Fundamental Growth	Commodity: Agriculture
Fundamental Value	Commodity: Energy
Sector: Energy/Basic Materials	Commodity: Multi
Sector: Healthcare	Currency: Discretionary
Sector: Technology	Currency: Systematic
Short Bias	Discretionary Thematic
Multi-Strategy	Systematic Diversified
	Multi-Strategy
Testing	Training

Classification: Training Sample

- HFR has natural candidate for training sample.
  - Systematic Diversified Macro funds: "investment processes that typically are functions of mathematical, algorithmic, and technical models, with little or no influence from individuals over the portfolio positioning."
  - Discretionary Thematic Macro funds: "primarily reliant on the evaluation of market data, relationships and influences, as interpreted by an individual or group of individuals who make decisions on portfolio positions."
- Training sample: Binary variable $y_i = 1$ if the i-th fund is a Systematic Diversified Macro fund and $y_i = 0$ if it is a Discretionary Thematic Macro fund; the feature matrix (explanatory variable matrix) of Macro funds as inputs to train classifiers.
- Our approach is free from subjective judgement of investment strategies/keywords.

Constructing Feature Matrices

- Consider $9,408$ investment strategies, with $7,174$ for Equity Hedge funds (Test set) and $2,234$ for Macro funds (Training set), from the HFR database.
- Tokenizing each document (*tidytext*, *RTextTools*)
  - Stop-words (e.g., is, the, and) are excluded.
  - Each word is stemmed to its root using the Porter stemmer algorithm. (e.g., cats, catty, cat will become cat)
  - A total of $7,923$ common tokens based on "bigrams", two consecutive words.
- Constructing two $N \times M$ feature matrices, where $N = 7174$ or 2234 (documents) and $M = 7,923$ (tokens).

Feature Matrix

The term frequency (tf) of the token j in the document i is:

$$\text{tf}_{ij} = \frac{\text{Number of times that token } j \text{ appear in the document } i}{\text{Total number of all tokens in the document } i},$$

and every tf is weighted by the inverse-document frequency (idf):

$$\text{idf}_j = \log \frac{\text{Total number of documents}}{\text{Number of documents that contain token } j}.$$

The larger the idf, the less frequently the token j is observed in these documents (Manning et al., 1999). The (i, j)-th element of a feature matrix is:

$$f_{ij} = \text{tf}_{ij} \cdot \text{idf}_j, \quad i = 1, \ldots, N, \; j = 1, \ldots, M.$$

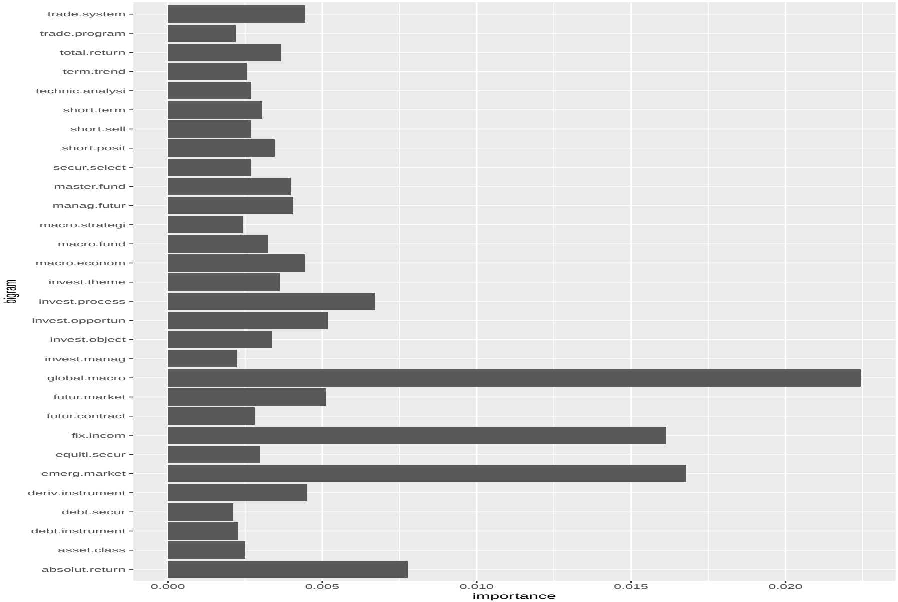
Classifier

- Statistical learning methods (Hastie et al., 2009; James et al., 2013) for supervised learning:
  1. Logistic regression (*stats*)
  2. Linear discriminant analysis (*lda*)
  3. k-nearest neighbour with $k = 1, 3, 5, 7$: (*caret*)
  4. Support vector machine with linear, radial, polynomial, and sigmoid kernels (*e1071*)
  5. Classification tree (*tree*); boosting (*gbm*); bagging and random forest (*ranger*)

Accuracy Ratios: 10-fold Cross Validation

Fold	LOG	LDA	KNN_1	KNN_3	SVM_l	SVM_s	TRE	BAG	RF	BOT
1	0.52	0.85	0.77	0.69	0.87	0.90	0.76	0.87	0.85	0.75
2	0.41	0.89	0.90	0.79	0.91	0.90	0.80	0.89	0.90	0.80
3	0.48	0.83	0.81	0.67	0.83	0.86	0.76	0.84	0.86	0.74
4	0.63	0.89	0.80	0.69	0.88	0.90	0.79	0.89	0.89	0.79
5	0.53	0.86	0.82	0.72	0.87	0.87	0.83	0.87	0.87	0.79
6	0.48	0.89	0.85	0.75	0.88	0.87	0.75	0.85	0.89	0.77
7	0.45	0.86	0.82	0.71	0.86	0.88	0.83	0.93	0.93	0.80
8	0.44	0.87	0.84	0.72	0.89	0.88	0.81	0.91	0.91	0.80
9	0.48	0.89	0.82	0.71	0.89	0.86	0.78	0.84	0.87	0.76
10	0.45	0.88	0.83	0.73	0.88	0.88	0.78	0.88	0.88	0.80
Mean	0.49	0.87	0.83	0.72	0.88	0.88	0.79	0.88	0.89	0.78
Min	0.41	0.83	0.77	0.67	0.83	0.86	0.75	0.84	0.85	0.74
Max	0.63	0.89	0.90	0.79	0.91	0.90	0.83	0.93	0.93	0.80

Top Features of the Random Forest Classifier



Top Features

- These features characterize similarities of the investment strategies of Macro funds.
- Importance measure of a feature: The average decrease in model accuracy of the out-of-bag samples when the values of that feature are randomly permuted.
- Top 5 features selected by random forest classifier are *global macro*, *emerg market*, *fix incom*, *absolut return*, and *invest process*.
- These features are quite different from those of Harvey et al. (2017) and Chincarini (2014) which are researcher-dependent. For example, the keywords used by Harvey et al. (2017), *algorithm*, *approx*, *computer*, *model*, *statistical*, and *system*, do not appear in our list of leading features.

Performance Comparison

- $2,149$ Equity Hedge funds and 603 Macro funds:
  - Reporting style: net of all fees
  - Assets under management (AUM): at least $10 million
  - At least 36 consecutive monthly returns

  A total of 239 monthly returns from Jan. 1996 to Nov. 2015.
- For fund classification, Equity Hedge funds are classified by the random forest classifier, and Macro funds are based on the HFR classification.
- Risk factors taken from:
  - Fama French website: MKT, SMB, HML, MOM, and Rf;
  - Fung and Hsieh (2004): PTFSBD, PTFSFX, PTFSCOM, PTFSIR, and PTFSSTK;
  - Federal Reserved Bank: CS, Δ10Y.

Estimating Factor Models

We estimate the 1-, 3-, 5-, 7- and 11-factor models based on monthly excess returns (return minus one-month TB rate):

$$r_{i,t} = \alpha_i + \boldsymbol{\beta}_i' \mathbf{f}_t + \epsilon_{i,t},$$

where α_i is the factor-adjusted return of fund i.

Model	$\mathbf{f}_t$
1. CAPM (Jensen, 1968)	MKT-Rf
3. Fama-French (1993)	MKT-Rf, SMB, HML
5. Fung and Hsieh	PTFSBD, PTFSFX, PTFSCOM, PTFSIR, PTFSSTK
7. Fung and Hsieh (2004)	MKT, SMB, CS, Δ10Y, PTFSBD, PTFSFX, PTFSCOM
11. Bali et al. (2014)	7 + HML, MOM, PTFSIR, PTFSSTK

Summary Statistics of HFR Main Strategies

	Equity Hedge			Macro		
	Dis.	Sys.	Diff.	Dis.	Sys.	Diff.
N	778	1371	593	234	369	135
Mean	6.98	7.88	0.90	5.91	6.94	1.03
STD	13.51	13.27	−0.24	12.55	14.13	1.58
SR	0.41	0.49	0.07	0.29	0.30	0.01
ACF	0.14	0.12	−0.02	0.07	−0.02	−0.08
F1	4.76	5.69	0.92	3.87	4.71	0.83
F3	4.88	5.78	0.89	3.79	4.59	0.80
F5	5.12	6.42	1.31	3.73	4.39	0.65
F7	3.17	4.61	1.45	3.08	3.77	0.68
F11	4.31	5.99	1.67	3.30	3.76	0.45

Summary Statistics of EH SubStrategies

	Market Neutral			Fundamental Growth		
	Dis.	Sys.	Diff.	Dis.	Sys.	Diff.
N	81	274	193	284	381	97
Mean	3.84	5.59	1.74	7.19	9.25	2.06
STD	7.89	7.24	−0.65	16.18	17.99	1.80
SR	0.20	0.54	0.33	0.35	0.46	0.11
ACF	0.09	0.10	0.02	0.16	0.14	−0.02
F1	1.54	3.24	1.70	5.07	7.19	2.12
F3	1.69	3.13	1.44	5.16	7.24	2.08
F5	1.50	3.76	2.26	5.54	8.15	2.61
F7	0.63	2.78	2.15	3.00	5.37	2.37
F11	1.24	3.55	2.31	4.69	7.51	2.82

Summary Statistics of EH SubStrategies (conti.)

	Fundamental Value			Quantitative Directional		
	Dis.	Sys.	Diff.	Dis.	Sys.	Diff.
N	388	629	241	25	87	62
Mean	7.59	8.13	0.54	5.39	7.32	1.93
STD	12.85	13.26	0.41	11.72	11.69	−0.03
SR	0.52	0.49	−0.03	0.22	0.43	0.21
ACF	0.14	0.12	−0.02	0.10	0.11	0.02
F1	5.31	5.90	0.59	3.29	5.32	2.02
F3	5.45	6.08	0.63	3.40	5.52	2.13
F5	5.76	6.59	0.83	2.01	6.02	4.01
F7	3.94	4.94	1.00	1.30	4.75	3.46
F11	4.88	6.13	1.25	1.23	5.94	4.71

Luck versus skill: Bootstrap analysis (Kosowski et al.,2006, JF)

Step 1 For each fund i, estimate a factor model and store $\hat{\alpha}_i$, $t(\alpha_i)$, $\hat{\boldsymbol{\beta}}_i'$, and residuals $\{\hat{\epsilon}_{i,s}\}$, $s = 1, \ldots, T$.

Step 2 Re-sample $\{\hat{\epsilon}_{i,s}\}$ and $\{\hat{\epsilon}^b_{i,1_b}, \ldots, \hat{\epsilon}^b_{i,T_b}\}$.

Step 3 Generate pseudo pure-luck (zero α) fund returns:

$$r^b_{i,s_b} = \hat{\boldsymbol{\beta}}_i' \mathbf{f}_{s_b} + \hat{\epsilon}^b_{i,s_b}, \quad s_b = 1_b, \ldots, T_b,$$

and estimate a factor model by regressing r^b_{i,s_b} on the intercept term and $\mathbf{f}_{s_b}$ and store the resulting $t(\alpha^b_i)$.

Step 4 Do Steps 1–3 for all funds to generate $t(\alpha^b_i)$, $i = 1, \ldots, N$, and compute quantiles of $t(\alpha^b_i)$.

Step 5 Repeat Steps 1–4 for $B = 10,000$ times to obtain the empirical distributions of the quantiles computed in Step 4.

Example: Median of $t(\alpha_i^b)$

$$
\begin{array}{cccc}
t(\alpha_1^{b=1}) & t(\alpha_1^{b=2}) & \cdots & t(\alpha_1^{b=10000}) \\
t(\alpha_2^{b=1}) & t(\alpha_2^{b=2}) & \cdots & t(\alpha_2^{b=10000}) \\
\vdots & \vdots & \ddots & \vdots \\
t(\alpha_N^{b=1}) & t(\alpha_N^{b=2}) & \cdots & t(\alpha_N^{b=10000}) \\
\\
\Big\downarrow & \Big\downarrow & \cdots & \Big\downarrow \\
\\
\mathrm{med}(t(\alpha^{b=1})) & \mathrm{med}(t(\alpha^{b=2})) & \cdots & \mathrm{med}(t(\alpha^{b=10000}))
\end{array}
$$

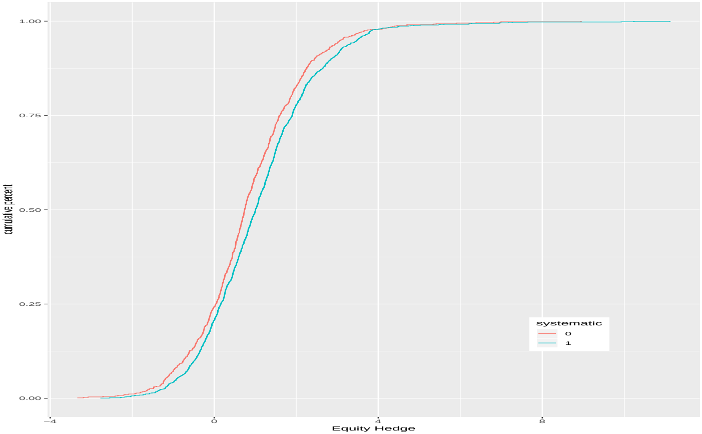
Results of the 7-factor Model: Equity Hedge

	Discretionary			Systematic			
Quantile	Act.	Avg.	p-value	Act.	Avg.	p-value	Diff.
Min.	-3.34	-4.04	0.15	-2.78	-4.39	0.00	0.56
1%	-2.11	-2.60	0.00	-1.69	-2.57	0.00	0.42
5%	-1.19	-1.78	0.00	-0.93	-1.76	0.00	0.26
10%	-0.75	-1.37	0.00	-0.49	-1.36	0.00	0.27
30%	0.20	-0.55	0.00	0.32	-0.54	0.00	0.12
50%	0.74	0.01	0.00	1.01	0.01	0.00	0.27
70%	1.43	0.58	0.00	1.65	0.57	0.00	0.22
90%	2.45	1.43	0.00	2.85	1.40	0.00	0.40
95%	3.08	1.86	0.00	3.48	1.82	0.00	0.41
99%	4.70	2.73	0.00	5.16	2.67	0.00	0.46
Max.	8.95	4.18	0.00	11.11	4.36	0.00	2.17

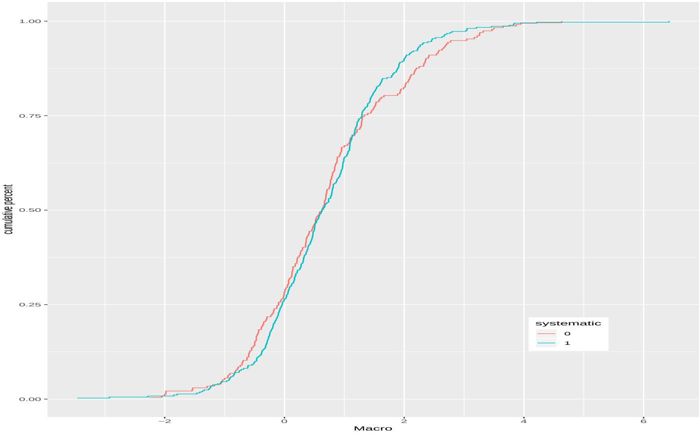
Results of the 7-factor Model: Macro

	Discretionary			Systematic			
Quantile	Act.	Avg.	p-value	Act.	Avg.	p-value	Diff.
Min.	-2.26	-3.37	0.01	-3.46	-3.57	0.41	-1.20
1%	-1.99	-2.50	0.02	-1.81	-2.41	0.00	0.18
5%	-1.01	-1.73	0.00	-0.92	-1.67	0.00	0.09
10%	-0.67	-1.33	0.00	-0.47	-1.29	0.00	0.20
30%	0.05	-0.54	0.00	0.11	-0.52	0.00	0.06
50%	0.63	0.00	0.00	0.63	0.00	0.00	0.00
70%	1.17	0.54	0.00	1.15	0.51	0.00	-0.02
90%	2.36	1.34	0.00	2.00	1.26	0.00	-0.36
95%	2.87	1.73	0.00	2.46	1.63	0.00	-0.41
99%	3.80	2.50	0.00	3.78	2.34	0.00	-0.02
Max.	4.63	3.28	0.04	6.42	3.32	0.01	1.79

CDFs of $t(\alpha)^S$ and $t(\alpha)^D$: Equity Hedge



CDFs of $t(\alpha)^S$ and $t(\alpha)^D$: Macro



Stochastic Dominance Test Results

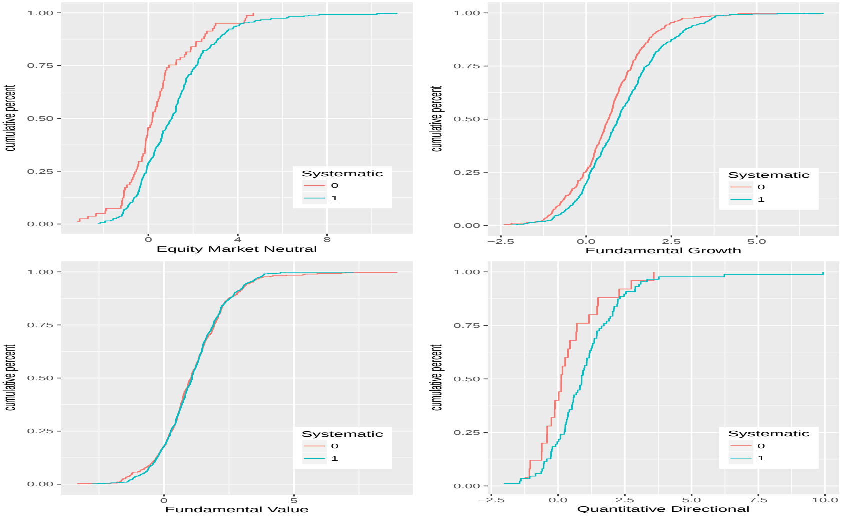
SD test of Linton et al. (2010) on $t(\alpha)$ of the 7-factor model.

$H_0^{(1)}$: $t(\alpha)^D$ stochastically dominates $t(\alpha)^S$.
$H_0^{(2)}$: $t(\alpha)^S$ stochastically dominates $t(\alpha)^D$.

We test first- and second-order SD relations.

	Equity Hedge		Macro	
Hypothesis	(1)	(2)	(1)	(2)
FSD	0.000***	0.991	0.277	0.171
SSD	0.000***	0.782	0.379	0.493

CDFs of $t(\alpha)^S$ and $t(\alpha)^D$: Sub-Strategies of EH



SD Test Results: Sub-Strategies of EH

	Market Neutral		Fundamental Growth	
Hypothesis	(1)	(2)	(1)	(2)
FSD	0.000***	0.982	0.000***	0.983
SSD	0.000***	0.674	0.000***	0.782
	Fundamental Value		**Quantitative Directional**	
Hypothesis	(1)	(2)	(1)	(2)
FSD	0.532	0.573	0.000***	0.896
SSD	0.418	0.539	0.000***	0.545

Summary of SD test

- The SD test significantly rejects $H_0^{(1)}$ (p values ≈ 0) but not $H_0^{(2)}$ (large p values) for Equity Hedge funds. This shows that systematic funds in this category would be preferred because their $t(\alpha)$ stochastically dominate those of their discretionary counterparts.
  - This conclusion also holds for the 3 out of 4 sub-categories of Equity Hedge funds: Market Neutral, Fundamental Growth, and Quantitative Directional.
- The SD test does not reject both hypotheses for Macro funds and for Fundamental Value funds of Equity Hedge funds. Hence, there is no clear SD relation between their systematic and discretionary funds.
- These conclusions remain valid when other factor models are used to estimate alphas.

Concluding Remarks

- This paper introduces a text-driven approach to classifying hedge funds into systematic and discretionary funds. This classification is determined by strategy similarities without subjective judgement.
- We test significance of factor-adjusted returns of the classified systematic and discretionary funds and test the SD relation between the CDFs of their standardized alphas.
- Our empirical results suggest that, while a large portion of the systematic and discretionary funds of Equity Hedge funds exhibits authentic investment skills, the SD test suggests that systematic funds would be preferred to their discretionary counterparts.

Chincarini's classification:

Table 3

Classification 1 of quantitative and qualitative funds

This table reports the HFR hedge fund strategies that were classified as quantitative or qualitative. Main represents the main hedge fund category and sub represents the sub-category within the main category according to HFR. *Source*: HFR strategy descriptions and authors' judgement. For a full description of each fund main category and sub-category, see Chincarini and Nakao (2011).

Categorised			
		Quantitative	Qualitative
Number	Main	Sub	Sub
1.	Equity Hedge		
		EH: Equity Market Neutral	EH: Fundamental Growth
		EH: Quantitative Directional	EH: Fundamental Value
2.	Macro		
		M: Commodity Systematic	M: Commodity Discretionary
		M: Currency Systematic	M: Currency Discretionary
		M: Systematic Diversified	M: Discretionary Thematic

2.3.2. *Classification 2.* In order to check our results against other methods of separating quantitative and qualitative hedge funds, we considered all hedge funds again, but performed a word search on the strategy description of each individual hedge fund in the database. We classified a fund as quantitative if the following words appeared in the fund description: *quantitative*, *mathematical*, *model*, *algorithm*, *econometric*, *statistic*, or *automate*. Also, the fund description could not contain the word *qualitative*. We classified a fund as qualitative if it contained the word *qualitative* in its description or had none of the words mentioned for the quantitative category.

Harvey's Crti:

1. **Chincarini classifies Equity Market Neutral funds as quantitative by default. This is particularly problematic for comparing the equity market exposure (i.e., beta) of quantitative and qualitative funds: His finding that quantitative funds are more market neutral may be a direct result of the chosen categorization method.**
2. **Harvey's: algorithm, approx, computer, model, statistical, system**
3. **Chincarini: algorithm, automate, econometric, mathematical, model, quantitative, statistic**

HFR Classification

Equity Hedge
Equity Hedge: Equity Hedge strategies maintain positions both long and short in primarily equity and equity derivative securities. A wide variety of investment processes can be employed to arrive at an investment decision, including both quantitative and fundamental techniques; strategies can be broadly diversified or narrowly focused on specific sectors and can range broadly in terms of levels of net exposure, leverage employed, holding period, concentrations of market capitalizations and valuation ranges of typical portfolios. Equity Hedge managers would typically maintain at least 50% exposure to, and may in some cases be entirely invested in, equities - both long and short. EH is further subdivided into 7 sub-strategies.
Equity Market Neutral
EH: Equity Market Neutral strategies employ sophisticated quantitative techniques of analyzing price data to ascertain information about future price movement and relationships between securities, select securities for purchase and sale. These can include both Factor-based and Statistical Arbitrage/Trading strategies. Factor-based investment strategies include strategies in which the investment thesis is predicated on the systematic analysis of common relationships between securities. In many but not all cases, portfolios are constructed to be neutral to one or multiple variables, such as broader equity markets in dollar or beta terms, and leverage is frequently employed to enhance the return profile of the positions identified. Statistical Arbitrage/Trading strategies consist of strategies in which the investment thesis is predicated on exploiting pricing anomalies which may occur as a function of expected mean reversion inherent in security prices; high frequency techniques may be employed and trading strategies may also be employed on the basis on technical analysis or opportunistically to exploit new information the investment manager believes has not been fully, completely or accurately discounted into current security prices. Equity Market Neutral Strategies typically maintain characteristic net equity market exposure no greater than 10% long or short.

Fundamental Growth
EH: Fundamental Growth strategies employ analytical techniques in which the investment thesis is predicated on assessment of the valuation characteristics on the underlying companies which are expected to have prospects for earnings growth and capital appreciation exceeding those of the broader equity market. Investment theses are focused on characteristics of the firm's financial statements in both an absolute sense and relative to other similar securities and more broadly, market indicators. Strategies employ investment processes designed to identify attractive opportunities in securities of companies which are experiencing or expected to experience abnormally high levels of growth compared with relevant benchmarks growth in earnings, profitability, sales or market share.
Fundamental Value
EH: Fundamental Value strategies which employ investment processes designed to identify attractive opportunities in securities of companies which trade a valuation metrics by which the manager determines them to be inexpensive and undervalued when compared with relevant benchmarks. Investment theses are focused on characteristics of the firm's financial statements in both an absolute sense and relative to other similar securities and more broadly, market indicators. Relative to Fundamental Growth strategies, in which earnings growth and capital appreciation is expected as a function of expanding market share & revenue increases, Fundamental Value strategies typically focus on equities which currently generate high cash flow, but trade at discounted valuation multiples, possibly as a result of limited anticipated growth prospects or generally out of favor conditions, which may be specific to sector or specific holding.
Quantitative Directional
EH: Quantitative Directional strategies employ sophisticated quantitative techniques of analyzing price data to ascertain

information about future price movement and relationships between securities, select securities for purchase and sale. These can include both Factor-based and Statistical Arbitrage/Trading strategies. Factor-based investment strategies include strategies in which the investment thesis is predicated on the systematic analysis of common relationships between securities. Statistical Arbitrage/Trading strategies consist of strategies in which the investment thesis is predicated on exploiting pricing anomalies which may occur as a function of expected mean reversion inherent in security prices; high frequency techniques may be employed and trading strategies may also be employed on the basis on technical analysis or opportunistically to exploit new information the investment manager believes has not been fully, completely or accurately discounted into current security prices. **Quantitative Directional Strategies typically maintain varying levels of net long or short equity market exposure over various market cycles.**