Journal of Banking & Finance 50 (2015) 608–615

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Generalized runs tests to detect randomness in hedge funds returns

Rania Hentati-Kaffel *, Philippe de Peretti [1]

Centre d'Economie de la Sorbonne (CES), 106-112 Boulevard de l'Hôpital, 75013 Paris, France

ARTICLE INFO

Article history:
Received 10 September 2013
Accepted 25 July 2014
Available online 8 August 2014

JEL classification:
G14
G23
C14

Keywords:
Hedge funds
Randomness
Persistence
Generalized runs tests
ARCH
Breaks

ABSTRACT

The major contribution of this paper is to make use of generalized runs tests (Cho and White, 2011) to analyze the randomness, i.e. the lack of persistence, in both absolute and relative returns of hedge funds. We find that about 42% of the HFR universe exhibit iid absolute returns over the period spanning 2000 to 2012. These funds are mainly found in proportions within the Macro and Equity Hedge strategies. A similar result holds for relative returns. We also find that funds having non-iid returns often exhibit ARCH effects and structural breaks, with largest breaks located within financial crises. Also, only a small percentage displays persistence in their relative performance, 8.2% to 16.7% of the universe, mainly found in proportions within the Relative Value and Event-Driven strategies. The robustness of results is challenged by implementing the tests on a crisis-free period. We find similar results for absolute returns. For relative ones, differences appear across strategies and benchmarks, but still both ARCH and breaks are present. Our work contributes to the hedge fund literature in terms of methodology, portfolio allocation, and performance measurement.

1. Introduction

The recent financial crisis has stressed the crucial role of hedge funds returns analysis in the selection process of managers. Indeed, due to the well known opacity of hedge funds, limiting access to information on implemented positions and on strategies, past performances analysis to detect positive persistence in absolute or relative returns is often one of the main quantitative tools helping investors to pick the right manager. Apart from investors, analyzing performance persistence is also of interest for economists focusing on market efficiency as in Fama and MacBeth (1973).

This paper aims mainly to analyze performance persistence in the light of a new econometric test. One of our major contributions is to make use of generalized runs tests for randomness, introduced by Cho and White (2011). To our knowledge, it is the first time that such tests are used on hedge fund data. We implement generalized runs tests on both absolute and relative returns. In the former case, generalized runs tests return straightforward information about the iid nature of returns. In the latter case, generalized runs tests are to be interpreted as goodness-of-fit tests, and failure to reject the null hypothesis suggests that the assessed fund is unable to outperform its related market.

A large literature has dealt with performance persistence. But, as pointed out by Boyson (2008), this one beginning with the early studies of Sharpe (1966) and Jensen (1968), has returned contradictory results, and drawing clear-cut conclusions is uneasy. Thus, whether or not hedge funds are able to produce persistence remains an open question (see Eling, 2009 for a survey of this vast literature). As an illustration, early studies have supported short-term but not long-term persistence (see e.g. Agarwal and Naik, 2000a,b; Baquero et al., 2005; Brown et al., 2001 or Gyger et al., 2003), whereas late ones have also found some empirical evidences of long-term persistence (see e.g. Fung et al., 2008, Jagannathan et al., 2010 or Kosowski et al., 2007). To deal with these contradictory results, the literature has been extended in at least two directions. The first one tries to relate hedge fund performance persistence to fund characteristics as in Amenc et al. (2003) or in Getmansky (2012). The second one has set the focus on more advanced econometric tools, as the CPR (Agarwal and Naik, 2000a), the Chi-square test (Park and Staum, 1998; Carpenter and Lynch, 1999), the RIC (Herzberg and Mozes, 2003), the Kolmogorov–Smirnov goodness-of-fit test (Agarwal and Naik, 2000a) or the Hurst exponent as in De Souza and Gokcan (2004). At last, recently, Kosowski et al. (2007) have introduced a Bayesian approach to improve the accuracy of alpha estimates in parametric models.

* Corresponding author. Tel.: +33 144078277.
E-mail addresses: rania.kaffel@univ-paris1.fr (R. Hentati-Kaffel), philippe.de-peretti@univ-paris1.fr (P. de Peretti).
[1] Tel.: +33 144078763.

http://dx.doi.org/10.1016/j.jbankfin.2014.07.011
0378-4266/

Our paper clearly relates to this second branch, and our first major contribution is to analyze performance persistence using newly introduced generalized runs tests of Cho and White (2011). Generalized runs tests are powerful tests that allow to check for the null of randomness, i.e. lack of persistence, against a broad and undefined alternative including first and second-order dependence or structural breaks. The second major contribution is to implement generalized runs tests on a selected sample of hedge funds extracted from HFR database (4759 funds) over the period spanning January 2000 to December 2012. Both absolute and relative returns are analyzed to check the iid nature of the former and the ability of a fund to outperform a market in the latter case. Several benchmarks are then used corresponding to either the hedge fund industry, or to an equity market. Results are reported using a break down by primary strategies, i.e. Equity Hedge, Event-Driven, Macro and Relative Value. When the null is rejected, we consider an incomplete mapping of the alternative made of three main rejection factors: AutoRegressive Conditional Heteroskedasticity (ARCH), structural breaks and clustering, the latter being our measure of persistence. This is our third contribution. At last, results robustness is challenged by running tests on a crisis-free sub-period identified by structural breaks.

Using generalized runs tests, we find that for around half of our universe, we fail to reject the null in absolute and relative returns cases. The concerned funds are mainly found in the Equity Hedge and Macro strategies. Back to the three main rejection factors, we find that volatility clustering is common to all considered strategies. Structural breaks do appear but less frequently. In both cases, the results depend on the used benchmark and the fund strategy. Also, few funds are able to outperform one benchmark, i.e. to produce clusters: between 8.15% and 16.7% of our universe. These funds are mainly found in proportions within Relative-Value and Event Driven strategies. When a period excluding crises is considered, volatility clustering and structural breaks are still present, but to a lesser extent suggesting that they are inherent to this kind of industry, and partly not directly related to crisis. Our work has therefore direct implications in terms of portfolio allocation, and performance measurement. At a methodological level, generalized runs appear to be quite promising in these kinds of studies.

This paper is structured as follows. In Section 2, we present the generalized runs tests. In Section 3, we implement the tests on HFR database. Section 4 goes deeper into the alternative, and Section 5 considers a crisis-free period. At last Section 6 concludes and discusses our results.

2. Generalized runs tests

To analyze the randomness of absolute and relative returns we use Generalized Runs (GR) tests. Runs tests have been first introduced by Mood (1940), Wald and Wolfowitz (1940) among others. Granger (1963), Fama (1965) or Dufour (1981) have suggested their use as tests for randomness. They have been recently extended by Cho and White (2011), introducing generalized runs tests, but not yet been applied. Generalized runs tests are a powerful mean to test the iid assumption against an unspecified broad alternative, including first and second order dependence or structural breaks. This is a major difference with classical runs tests in which the alternative is defined, i.e. clustering or mixing.

Define $\left\{r_{it}^{j}\right\}_{t=1}^{T}$ as a track record of absolute or relative returns of a hedge fund i having a strategy j computed as residuals of the linear model $r_{it}^{j}=h(X_t,\theta)$, where $X_t=\left(b_t,r_{it}^{oj}\right), b_t$ is a benchmark of interest at time t, r_{it}^{oj} is the observed return, and θ is a parameter. The assumption we want to test is $\mathcal{H}_0:\left\{r_{it}^{j}\right\}_{t=1}^{T}$ is an iid sequence, against the broad alternative that $\{r_{it}^{j}\}_{t=1}^{T}$ is not an iid sequence. Let $F(.)$ be the cumulative distribution function of $\left\{r_{it}^{j}\right\}_{t=1}^{T}$, and first assume that both θ and $F(.)$ are perfectly known. Then using the notation in Cho and White (2011), the runs are defined in two steps. First, for a given probability p build the set $T_n(p)=\left\{t\in\{1,2,\ldots,n\}|F\left(r_{it}^{j}\right)<p\right\}, n=1,2,\ldots$ that contains all indices such that the percentiles are less than the probability p. Let $M_n(p)$ be the number of elements in the set. Now, sort by ascending order $T_n(p)$, and let $t_{n,r}(p)$ be the element being at the rth position of the sorted set, $r=1,\ldots,M_n(p)$. For a given p, the *p-runs* $R_{n,r}(p)$ are defined as follows:

$$R_{n,r}(p)=\begin{cases}t_{n,r}(p), & r=1;\\ t_{n,r}(p)-t_{n,r-1}(p), & r=2,\ldots,M_n(p).\end{cases}\tag{1}$$

To test for the null hypothesis and for a given $s\in\mathbb{S}$ compute the goodness-of-fit statistic $G_n(p,s)$:

$$G_n(p,s)=\frac{1}{\sqrt{n}}\sum_{r=1}^{M_n(p)}\left(s^{R_{n,r}(p)}-\frac{sp}{1-s(1-p)}\right)\tag{2}$$

Various tests statistics are then derived by integrating $G_n(p,s)$ over $s$for a given p; integrating over p for a given s; integrating over both p and s, or taking the supremum of the function for some p or s.

In this paper, among all these statistics, we focus on $\mathcal{T}_{1,n}^{p}(\mathbb{S}_1)=\int_{\mathbb{S}_1}|G_n(p,s)|ds, \mathbb{S}_1=[-0.99,0.99]$ for $p\in\{0.1,0.3,0.5,0.7,0.9\}$, setting $n=T$. Testing for the null then amounts to computing $\mathcal{T}_{1,n}^{p}(\mathbb{S}_1)$ for various p and then comparing the computed values to critical values at a given threshold. Let $\mathbf{p}_{cv}$ be the (5×1) vector of critical values at a given threshold for the various p, and let $\boldsymbol{\tau}$ be the corresponding vector of $\mathcal{T}_{1,n}^{p}(\mathbb{S}_1)$ statistics. Then, we fail to reject the null if $\boldsymbol{\tau}\ll\mathbf{p}_{cv}$ (critical values are reported by Cho and White (2011)).

Now, if either $F(.)$ or θ, or both are unknown, they are replaced by their estimators $\widehat{F}(r_{it}^{j})=\frac{1}{T}\sum_{t=1}^{T}1_{(r_{it}^{j}\leqslant\hat{r}_{it}^{j})}$ and $\hat{\theta}$. In this case, Empirical Generalized Runs (EGR) tests are used in a similar fashion, replacing $T_n(p)=\{t\in\{1,2,\ldots,n\}|F(r_{it}^{j})<p\}$ by $\widehat{T}_n(p)=\{t\in\{1,2,\ldots,n\}|\widehat{F}(r_{it}^{j})<p\}, t_{n,r}(p)$ by $\hat{t}_{n,r}(p), M_n(p)$ by $\widehat{M}_n(p)$. Thus, (1) and (2) are re-defined as:

$$\widehat{R}_{n,r}(p)=\begin{cases}\hat{t}_{n,r}(p), & r=1;\\ \hat{t}_{n,r}(p)-\hat{t}_{n,r-1}(p), & r=2,\ldots,\widehat{M}_n(p).\end{cases}\tag{3}$$

$$\widehat{G}_n(p,s)=\frac{1}{\sqrt{n}}\sum_{r=1}^{\widehat{M}_n(p)}\left(s^{\widehat{R}_{n,r}(p)}-\frac{sp}{1-s(1-p)}\right)\tag{4}$$

if $p\in(n^{-1},1)$ and $\widehat{G}_n(p,s)=0$ otherwise.

Our test statistics then becomes $\widehat{\mathcal{T}}_{1,n}^{p}(\mathbb{S}_1)=\int_{\mathbb{S}_1}\left|\widehat{G}_n(p,s)\right|ds$. We next turn to empirical applications.

3. Data and results

In this section, we first present the HFR database. We then implement EGR tests to analyze the iid nature of absolute and relative returns.

3.1. HFR database

Growth in the hedge fund industry has resumed since the financial crisis of 2008. The increasing hedge funds data providers

Table 1
Repartition by secondary strategies for Equity Hedge, Event-Driven, Macro and the Relative Value primary strategies.

Main strategy			
Sub-strategy	Percent	Sub-strategy	Percent
Equity hedge strategy		*Event-driven strategy*	
Equity market neutral	11.34	Activist	4.88
Fundamental growth	29.48	Credit arbitrage	6.87
Fundamental value	38.09	Distressed-restructuring	27.93
Multi-strategy	4.52	Merger arbitrage	12.19
Quantitative directional	4.96	Multi-strategy	14.19
Sector energy-basic materials	5.85	Private issue-regulating	1.33
Sector technology-health care	4.38	Special situation	32.59
Short bias	1.06		
Macro		*Relative value*	
Active trading	4.66	Fixed income-asset backed	17.47
Commodity-agriculture	2.20	Fixed income-arbitrage convertible	9.31
Commodity-energy	1.39	Fixed income-corporate	19.77
Commodity metal	3.30	Fixed income-sovereign	7.47
Commodity-multi	9.49	Multi-strategy	27.59
Currency-discretionary	4.15	Volatility	9.19
Currency-systematic	6.52	Yield alternatives-energy infrastructure	4.25
Discretionary thematic	18.05	Yield alternatives-real estate	4.94
Multi-strategy	16.19		
Systematic diversified	34.04		

(Hedge Fund Research, Tass/Lipper, Bloomberg, Hennessee, Managed Accounts Reports) and the all time high of the asset under management show the renewed interest of investors.

In this paper, return series come from HFR database. The choice of this database relies on the following reasons: (i) High coverage rate of the existing hedge funds universe, (ii) HFR indices by strategy attenuated the survivorship bias as liquidated funds are taken into account in indices returns calculation, (iii) The impact of the backfill bias is neglected in the HFR indices. In fact, HFR construction methodology ensures that constituents are selected as unique representative of redundant fund share classes and retains only funds with either \$50 M or 12 months of track record. Beside this, it is clear that biases remain. For example, Fung and Hsieh (2000) estimated the backfill bias in the Tass database to be 1.4% annually. Brown et al. (1999) report a bias of 3%. By comparing the Tass and the HFR database, Liang (2000) examines this survivorship bias in hedge fund returns, and finds that the survivorship bias exceeds 2% per year in the Tass database, while the HFR database survivorship bias equals 0.6%. Caglayan and Edwards (2001) have highlighted the impact of survivorship bias by including in their research funds that have not survived and also excluded the first year of the track record to avoid the instant bias. Using data from 1990 to 1998, they confirmed the presence of persistence for both winners and losers (57%: 30% of losers and 27% of winners). They showed that funds which displayed more persistence are Global Macro (58%) and Neutral Market (63%). At last, in a recent contribution, Joenväärä et al. (2014) make three suggestions in order to deal with biases. In particular they suggest rebuilding an aggregate database from various providers.

In this paper, net-of-fee observed returns data between January 2000 and December 2012 are used.[2] The starting fund universe is constituted of 4759 funds classified within 4 primary strategies: 47% in Equity Hedge, 25% in Global Macro, 18% in Relative Value and 10% in Event-Driven. Table 1 provides the list of sub-strategies within each strategy, which associated proportions.

[2] In this paper, following the Liang (2000) study, we have also implemented tests taking into account the survivorship bias. Since results were not significantly altered compared to those obtained on the raw data, only the latter are reported.

3.2. Results of empirical generalized runs tests

We next implement EGR tests on funds having a track record of at least 100 observations. Also, following Joenväärä et al. (2014), funds exhibiting extreme realizations in their returns or relative returns are not considered. We run the EGR tests on absolute and on relative returns. For the former, series of median-adjusted returns are computed. For the latter, relative returns are computed using four different benchmarks, defined as:

$$r_{it}^{j} = \begin{cases} r_{it}^{oj} - rHFR_t^{j}, \\ r_{it}^{oj} - rHFR_t, \\ r_{it}^{oj} - rSP500_t, \\ r_{it}^{oj} - med_t^{j}. \end{cases} \tag{5}$$

where:

$rHFR_t^j$ is the return in period t computed using the class HFR index corresponding to the main strategy j, where here the four main strategies are $j = 1$: Equity Hedge, $j = 2$: Macro, $j = 3$: Event-Driven, $j = 4$: Relative Value, $rHFR_t$ is the return in period t computed using the overall HFR index for all strategies, $rSP500_t$ is the return in period t computed using the S&P500 index, med_t^j is the median of the returns in period t for funds having a common strategy j, r_{it}^{oj} is the observed return for hedge i having a strategy j at time t.

In Definitions 1 to 4 we force $\theta = -1$. The tests are therefore to be interpreted as goodness-of-fit or adequation tests. Failing to reject the null thus leads to conclude that the discrepancy between the returns of a fund and a benchmark is at random, fund i behaving not differently from its benchmark. For Definition 4 we search if among all funds having a common strategy j, a given fund is randomly distributed within the distribution of the returns or not.

3.2.1. Outcomes based on absolute returns

Table 2 reports the results of EGR tests. Main entries are the proportions of funds within each main strategy for which we fail to reject the null at five percent, and the total proportions regarding our hedge fund universe. We fail to reject the null for 42.26% of the funds included in our universe. In proportions, funds exhibiting iid returns are mainly to be found within the Macro (64.64%) and

Table 2
Proportions of funds (in %) for which we fail to reject the null of iid returns at 5%, within the four main strategies.

	Equity Hedge	Event-Driven	Macro	Relative Value	Total
Percent[a]	44.18	19.61	64.64	21.14	.
Percent[b]	23.47	2.38	13.48	2.93	42.26

[a] Proportions of funds within each strategy.
[b] Proportions of funds with regard to the HFR universe.

Table 3
Proportions of funds (in %) for which we fail to reject the null of iid relative returns at 5%, within the four main strategies.

	Benchmarks			
	$rHFR_t^j$	$rHFR_t$	$rSP500_t$	med_t^j
Equity Hedge[a]	49.25	48.51	41.64	48.66
Event-Driven[a]	33.99	39.87	54.90	28.76
Macro[a]	60.08	56.65	34.98	63.12
Relative Value[a]	21.71	42.86	53.71	20.00
Percent[b]	46.01	48.37	43.53	45.28

[a] Proportions of funds within each strategy.
[b] Proportions of funds with regard to the HFR universe.

Equity Hedge (44.18%) strategies. For Event-Driven and Relative Value, the null is rejected for about 80% of the funds having these strategies.

3.2.2. Outcomes based on relative returns

Results are presented in Table 3. Main entries are the proportions of funds for which we fail to reject the null at 5% for each main strategy and by benchmarks, and the proportions of funds having iid relative returns for each benchmark. Focusing on the latter, regardless of the benchmark, about slightly less than half of the sample has iid relative returns. Concerning the former, results are benchmark and strategy-dependent, which is particularly clear for the Event-Driven, Macro and Relative Value strategies, but less for Equity-Hedge. For instance, focusing on the class index $\left(\text{rHFR}_t^j\right)$, 60.08% of the funds having a Macro strategy do not perform differently from it,whereas for the Equity Hedge, the corresponding figure is 49.25%. For Event-Driven and Relative Value, the proportion are quite different, with lower figures, respectively 33.99% and 21.71%. Proportions are similar when one focuses on the global HFR index, except for the Relative Value strategy, with a much higher percentage. Now, relative to an equity market, i.e. the S&P500 index, respectively 41.64%, 54.90%, 34.98% and 53.71% of the funds having an Equity Hedge, Event-Driven, Macro or Relative Value strategy have iid relative returns. At last, using the median of the returns similar results of that of the class index. Thus, the two major conclusions are that (i) About 50% of the funds does not behave differently from the benchmark, (ii) In proportions, funds behaving as the benchmark are likely to be found within the Macro and Equity Hedge strategies.

4. A deeper look into the alternative

EGR tests return a key information about the iid property of series. Nevertheless, since the alternative is not defined, when rejecting one can not know the reasons why. Hence the need for a deeper look into the alternative.

For the funds for which the null of randomness is rejected at five percent, we consider an incomplete mapping made of three rejection factors: (i) First order dependence and especially clustering, (ii) Second-order dependence, i.e. AutoRegressive Conditional Heteroskedasticity (ARCH), and (iii) Structural breaks.

To detect ARCH effects, we use a classical ARCH-LM test (Engle, 1982). For structural breaks, we use the Andrews and Ploberger (1994) SupF test, where the p-values are computed using the fixed regressors bootstrap of Hansen (2000) to deal with heteroskedasticity. At last, to analyze clustering, we use one-sided runs-based tests (see Gibbons and Chakraborti, 1992). Concerning the latter, define $\left\{d_{it}^j\right\}_{t=1}^T$ as:

$$d_{it}^j = \begin{cases} 1 & \text{if } r_{it}^j \geqslant 0, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

and define a run of one kind of element, say of $1's$ as a successions of $1's$ immediately preceded or followed by at least one 0, or nothing. Let T_1 be the number of $1's$ and T_0 be the $0's$ with $T_1 + T_0 = T$, and let r_{1j} be the number of runs of $1's$ of length j and r_{0j} be the number of runs of $0's$ of length j. Let $r_1 = \sum_j r_{1j}$ be the total number of runs of $1's$, and $r_0 = \sum_j r_{0j}$ the total number of runs of $0's$. At last let $r = r_1 + r_0$ be the total number of runs of both kinds.

1 0 1 0 0 1 1 1 0 1 0 0 0 1

we have $r_{11} = 4, r_{01} = 2, r_{12} = 0, r_{02} = 1, r_{13} = 1, r_{03} = 1, r_1 = 5$, $r_0 = 4$ and $r = 9$.

Focusing on the total of number of runs, the first and second moments are defined as:

$$E[r] = E[r_1] + E[r_0] = \frac{2T_1T_0}{T} + 1 \tag{7}$$

$$V[r] = V[r_1] + V[r_0] + 2cov[r_1, r_0] = \frac{2T_1T_0(2T_1T_0 - T)}{T^2(T-1)} \tag{8}$$

Using the above defined moments, to test for the null of randomness against clustering, a Z-stat is used. Using a continuity correction, this latter is given by:

$$Z_r = \frac{r + 0.5 - 2T_1T_0T^{-1} - 1}{\sqrt{\frac{2T_1T_2(2T_1T_0 - T)}{T^2(T-1)}}} \tag{9}$$

which is asymptotically distributed as a normal standard deviate.

For relative returns, focusing on clustering is of particular importance since it allows us analyzing the number of funds within each strategy able to significantly outperform the market during large periods of time.

We first set the focus on structural breaks and ARCH effects. For the former test, we consider breaks in the intercept and in autoregressive coefficients (if any), where the lag is chosen according to the AIC criterion (Akaike, 1974). For the ARCH test, our test is based on 4 lags in the auxiliary regression, whereas the main regression includes an intercept, and also possibly autoregressive parameters, chosen here again using the AIC criterion. Since structural breaks may cause ARCH effects (see e.g. Russell, 2013) and also to analyze the impact of structural breaks on the variance of the series, we separately report the proportions of funds within each strategy exhibiting ARCH effects, structural breaks, ARCH effects alone, i.e. without structural breaks, structural breaks alone, and both effects appearing together.

4.1. Outcomes based on absolute returns

Table 4 presents results of the ARCH and structural breaks analysis for absolute returns. The first key result unveiled is the large amount of funds exhibiting volatility clustering. This is particularly true for funds in Equity-Hedge, Event-Driven and Relative Value categories. For the latter, around 50% of the funds also show structural breaks. For the other strategies, structural breaks occur in approximately 20% of the cases. Results raise one question, are ARCH effects and structural breaks related? The lower part of the

Table 4
Proportions of funds exhibiting non-iid returns, for which an ARCH effect or a structural break is found at 5%. Results given by main strategies.

	Equity Hedge	Event-Driven	Macro	Relative Value
ARCH	68.18	61.78	43.01	67.39
Structural breaks	20.05	17.88	21.50	47.10
ARCH alone	50.80	48.78	26.88	34.78
Structural breaks alone	2.67	4.88	5.38	14.49
ARCH and structural breaks	17.38	13.01	16.13	32.61

Proportions of funds within each strategy.

Table 5
Proportions of funds exhibiting clustering in their returns at 5%, for (i) Funds with no ARCH effects, and no structural breaks, (ii) Funds with ARCH effects but no structural breaks, (iii) Funds with and without ARCH effects but with no structural breaks.

	Equity Hedge	Event-Driven	Macro	Relative Value
Funds with no ARCH effects nor structural breaks				
Percent	40.37	58.54	39.58	76.01
Funds with ARCH effects but no structural breaks				
Percent	33.16	63.33	28.00	77.08
All funds but with no structural breaks				
Percent	35.79	61.39	35.62	76.71

Proportions of funds within each strategy.

Table 6
Proportions of funds exhibiting non-iid relative returns, for which an ARCH effect or a structural break is found at 5%. Results given by main strategies.

	Benchmarks			
	$rHFR_t^j$	$rHFR_t$	$rSP500_t$	med_t^j
Equity hedge				
ARCH	60.88	0.29	75.19	21.80
Structural breaks	22.94	22.03	30.43	59.30
ARCH alone	43.24	0.29	52.17	41.57
Structural breaks alone	5.29	22.03	7.42	4.07
ARCH and breaks	17.65	0.00	23.02	17.73
Event-driven				
ARCH	72.28	0.00	85.51	21.10
Structural breaks	15.84	21.74	26.09	64.22
ARCH alone	56.44	0.00	62.32	47.71
Structural breaks alone	0.00	21.74	2.90	4.59
ARCH and breaks	15.84	0.00	23.19	16.51
Macro				
ARCH	38.10	0.00	69.59	21.65
Structural breaks	15.24	14.91	11.11	36.08
ARCH alone	28.57	0.00	60.82	23.71
Structural breaks alone	5.71	14.91	2.34	14.52
ARCH and breaks	9.52	0.00	8.77	12.37
Relative value				
ARCH	62.77	1.00	87.65	47.86
Structural breaks	32.85	36.00	18.52	62.14
ARCH alone	39.42	1.00	71.60	26.43
Structural breaks alone	9.49	36.00	2.47	12.14
ARCH and breaks	23.36	0.00	16.05	35.71

Proportions of funds within each strategy.

Table carries partial answer. Except for Relative Value, structural breaks seldom appear alone, conversely to ARCH effects.

Results of the clustering analysis are reported in Table 5 and the upper part of Table 8. Table 5 displays the proportions of funds that do cluster within each strategy, with a break down by ARCH/No-ARCH effects. We thus report the proportions of funds clustering when no ARCH effects are found, when ARCH effects are present, and in both cases. In case of ARCH effects, tests for clustering are implemented on normalized series.[3] The upper part of Table 8 reports the proportions funds exhibiting some clustering in their returns with regard to the whole HFR universe. Clearly, about 20% of the funds in the database cluster and these funds are mainly found within the Event-Driven and Relative Value strategies. For the two other strategies, about one third of the funds do cluster. These results can be partly explained by the illiquidity nature of the underlying assets, and by the way the returns are reported in practice (Getmansky et al., 2004).

4.2. Outcomes based on relative returns

We perform the same analysis on relative returns. Similarly, we begin by studying ARCH effects and structural breaks, and then clustering. Table 6 reports the results, where as previously the two effects are jointly and separately analyzed. Outcomes are clearly benchmark and strategy-dependent, especially for ARCH effects. For example, when the HFR class index is used, 60.88% of the Equity Hedge funds exhibit volatility clustering in their returns, but when using the overall HFR index as benchmark, this figure decreases to 0.29%. Very similar results are found for all strategies for the two benchmarks: The ARCH effect vanishes when one considers the overall HFR index. Note that this result does not hold for structural breaks, since the proportions of funds remains very similar across the two benchmarks.

Considering the S&P500 index results in a sharp increase in volatility clustering, up to 87.65% for the Relative Value. Still the structural breaks are very present. When analyzing the position of a fund within the distribution of relative returns, all having the same strategy, the striking fact is the relatively high number of structural breaks, e.g. 62.14% for the Relative Value, 64.22% for the Event-Driven. This suggests that the relative performances of a fund with regard to the other ones having the same strategy is by far not constant over time. This has deep implications concerning portfolio allocation and performance measurement.

Finally the results of the clustering analysis, results shown in Tables 7 and 8, are heterogeneous: the overall proportions of funds able to significantly outperform the benchmark vary from 8.15% (relative to the S&P 500 index) to 16.7% (relative to the HFR Global index). Looking inside each strategy relative to the corresponding strategy index (HFR class index), it turns out that Relative Value and Event-Driven funds have the highest probability to produce cluster. When returns are calculated relative to the S&P500 index, only Relative Value funds do cluster.

5. Analysis over a crisis-free period

In the previous section, we have shown that over the entire sample, for about half of our universe, we failed to reject the null of the iid assumption for absolute and relative returns. Moreover, ARCH effects and structural breaks were common for those funds, unlike clustering. As our sample spans two major crises, i.e. the end of the so-called dot-com crisis, roughly covering 1999–2001 and the global financial crisis (2007–2008), one may ask to what extent

[3] For this, we fit a simple Generalized Auto-Regressive Conditional Heteroskedastic (GARCH) model to the series, and divide the series by the estimated conditional standard error.

Table 7
Proportions of funds exhibiting clustering in their relative returns at 5%, for (i) funds with no ARCH effects, and no structural breaks, (ii) Funds with ARCH effects but no structural breaks, (iii) For funds with and without ARCH effects but with no structural breaks.

	Benchmarks			
	$rHFR_t^j$	$rHFR_t$	$rSP500_t$	med_t^j
Funds with no ARCH effects nor structural breaks				
Equity Hedge	25.22	30.22	14.71	30.95
Event-Driven	53.57	54.17	0.00	50.00
Macro	23.73	19.59	18.75	28.30
Relative Value	63.16	61.90	62.50	55.56
Funds with ARCH effects but no structural breaks				
Equity Hedge	31.29	100	18.14	32.87
Event-Driven	61.40	0.00	6.98	53.85
Macro	23.33	0.00	33.65	4.35
Relative Value	75.93	100	5.17	72.97
All funds but with no structural breaks				
Equity Hedge	28.63	31.23	18.38	34.75
Event-Driven	58.82	29.17	5.88	56.16
Macro	23.60	21.65	30.92	27.54
Relative Value	70.65	37.50	64.55	72.13

Proportions of funds within each strategy.

Table 8
Proportions of funds exhibiting clustering in their relative returns at 5%. Overall proportions.

	Equity Hedge	Event-Driven	Macro	Relative Value	Total
Returns					
	8.48	4.91	2.06	4.44	19.89
Relative returns					
$rHFR_t^j$	5.94	3.95	1.66	5.15	16.7
$rHFR_t$	6.66	1.66	1.65	1.90	11.87
$rSP500_t$	3.96	0.23	3.72	0.24	8.15
med_t^j	6.50	3.64	1.50	3.48	15.12

Proportions of funds with regard to the HFR universe.

our results depend on these two major events. In other words, are the main rejection factors, as ARCH or structural breaks related to crises or are they inherent to the hedge fund industry itself. Indeed, in a recent contribution, Criton and Scaillet (2011) suggest that structural breaks are common in hedge funds absolute returns, indicating dynamic management. Following Bollen and Whaley (2009), they use multiple break tests, and showed that breaks are located both in and outside crisis periods.

In this section, we address the question of the robustness of our results. We therefore apply the same methodology relying on the EGR tests, and then the same incomplete mapping under the alternative. We first define a testing period between the two crises. For that purpose, for each fund exhibiting structural breaks in the previous analysis, we pick up the main breaking date.[4] This date is the one for which the corresponding F-test is maximal. We then compute the nonparametric (kernel) densities of these dates for absolute and relative returns.

The corresponding results are presented by Fig. 1 for absolute returns and Fig. 2 for relative ones with a break down by strategies. For the former we have (at least) a bimodal distribution, the two modes being October 2001 and November 2008, corresponding to the end of the two crises (at least from the hedge funds point of view). For relative returns, we have also bimodal distributions. When the benchmark is either the class HFR index, or the global

[4] Concerning structural breaks, we have also applied the Bai and Perron (1998) procedure. Based on the $\sup F(t+l|t)$ test, we find that many series exhibit multiple breaks.

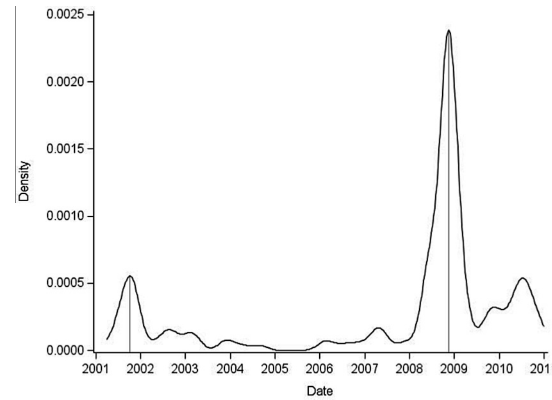

Fig. 1. Kernel density of main breaking dates for absolute returns.

HFR index, the modes are similar: September 2001 and November 2008. In the case of S&P500 index, the modes are October 2002 and February 2009. This reflects the different nature of crisis between hedge funds and traditional markets. Finally when the relative returns are computed by using the medians of all funds following the same strategy, the two modes are August 2001 and December 2008. Thus, in the sequel we will implement our tests on periods starting two months after the first mode of the distribution and ending two months before the second mode. Our analysis therefore spans approximately seven years. Note that for the analysis to be comparable to the previous one, we use exactly the same funds.

Results are given in Tables 9 and 11 for absolute returns, and in Tables 10 and 12 for relative ones. Tables 9 and 10 contain EGR tests results which indicate that for relative returns, changing the sample does not alter our conclusions, and still for about half of the sample we fail to reject the null. Proportions within each main strategy are also equivalent. Table 10 compared to Table 3 unveils slightly different results for relative returns, especially for those computed either using the HFR-strategy or the S&P500 index. In the former case a higher figure is found, and about 65% of relative returns are significantly not different from the benchmark. In the latter case, the corresponding figure is lower, with only 33.62% against 43.53% on the whole sample. Large changes are found for Event-Driven and Relative-Value strategies, with lower percentages. Turning next to rejection factors for absolute returns, Table 11, ARCH and structural breaks are still but less present. At last, as previously found, hedge funds able to produce clusters are still mainly found, in proportions, within the Event-Driven and Relative Values strategies.

For relative returns, Table 12, results are highly dependent on both the strategy and the used benchmark. We note that ARCH effects and structural breaks remain key rejection factors especially when the S&P500 is considered as a benchmark. Note that in this latter case, more funds are used in the tests (66.38% of our universe) compared to our previous analysis (56.47%). It should be noticed that structural breaks found out of crisis periods are highly suggestive of dynamic management. Our result are thus in line with Criton and Scaillet (2011).

At last, hedge funds able to outperform the equity market, represented by the S&P 500 index, are still mainly found in proportions in the Relative Value strategy, and to a lesser extent in the Event-Driven one.

6. Conclusion and discussion

In this paper, we have adopted advanced runs statistics to analyze the statistical properties of both absolute and relative returns of hedge funds in terms of randomness. We have used a two-step

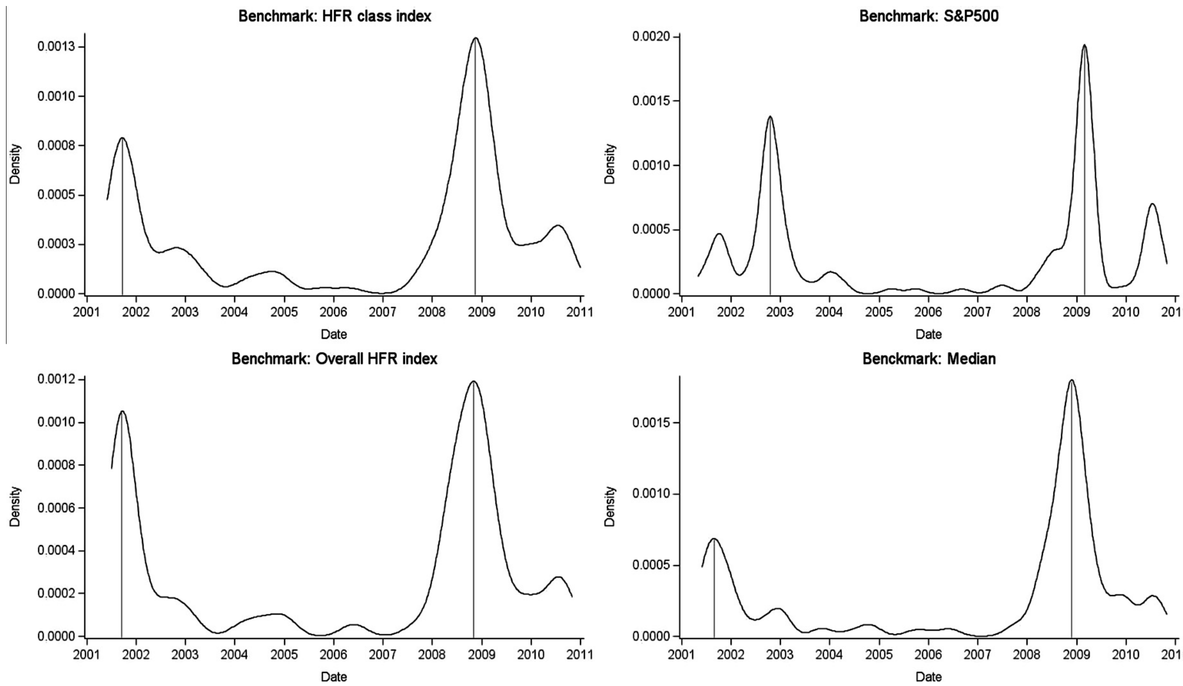

Fig. 2. Kernel densities of main breaking dates for relative returns. Results given by benchmark.

Table 9
Proportions of funds (in %) for which we fail to reject the null of iid returns at 5%, within the four main strategies. Sub-sample.

	Equity Hedge	Event-Driven	Macro	Relative Value	Total
Percent[a]	50.75	19.61	73.00	29.71	.
Percent[b]	26.96	2.38	15.23	4.12	48.69

[a] Proportions of funds within each strategy.
[b] Proportions of funds with regard to the HFR universe.

Table 10
Proportions of funds (in %) for which we fail to reject the null of iid relative returns at 5%, within the four main strategies. Sub-sample.

	Benchmarks			
	$rHFR_t^j$	$rHFR_t$	$rSP500_t$	med_t^j
Equity Hedge[a]	68.66	53.13	35.67	57.61
Event-Driven[a]	54.25	45.10	28.10	49.67
Macro[a]	78.33	56.27	30.04	61.98
Relative Value[a]	41.71	46.86	36.00	34.29
Percent[b]	65.19	51.94	33.62	54.32

[a] Proportions of funds within each strategy.
[b] Proportions of funds with regard to the HFR universe.

Table 11
Proportions of funds exhibiting non-iid returns, for which (i) ARCH effects, (ii) Structural breaks, (iii) Clustering are found. Clustering is analyzed only for funds exhibiting no structural breaks. Results given by main strategies. Sub-sample.

	Equity Hedge	Event-Driven	Macro	Relative Value
ARCH	26.36	23.58	15.49	33.33
Structural Breaks	14.85	13.33	4.23	26.02
Clustering	23.84	45.12	30.88	58.24

Proportions of funds within each strategy.

Table 12
Proportions of funds exhibiting non-iid relative returns, for which (i) ARCH effects, (ii) Structural breaks, (iii) Clustering are found. Clustering is analyzed only for funds exhibiting no structural breaks. Results given by main strategies. Sub-sample.

	Benchmarks			
	$rHFR_t^j$	$rHFR_t$	$rSP500_t$	med_t^j
Equity Hedge				
ARCH	24.29	24.20	63.81	30.99
Structural breaks	15.24	13.38	28.98	17.96
Clustering	23.03	18.01	14.83	28.76
Event-Driven				
ARCH	21.43	17.86	66.36	25.97
Structural breaks	9.71	11.43	25.45	22.16
Clustering	49.15	15.15	8.33	38.60
Macro				
ARCH	22.81	28.70	72.83	24.00
Structural breaks	5.26	6.96	24.67	11.00
Clustering	25.93	19.63	13.85	22.47
Relative Value				
ARCH	43.14	27.96	66.96	46.96
Structural breaks	14.71	13.66	16.43	25.22
Clustering	42.53	38.03	3.33	62.79

Proportions of funds within each strategy.

methodology as follows: (i) Implement EGR tests to classify funds according to the assumption that returns are iid at 5% level. This will segment the hedge fund universe studied into two groups: One for which we fail to reject the null, and the complement one, for which the iid assumption does not hold, (ii) For the second group, we have considered three possible rejection factors, i.e. ARCH effects, structural breaks and clustering.

On the overall sample we find that less than 50% of the sample exhibit iid returns, absolute or relative. We also find that most funds for which the null is rejected exhibit volatility clustering and/or structural breaks. Persistence in relative returns, measured by the ability to cluster, occurs between 8.15% and 16.7% of the cases, and funds producing clusters are mainly found in proportions within the Relative Value and Event-Driven strategies. At last, our results are both benchmark and strategy-dependent. On a crisis-free period, identified by structural breaks tests, nearly the same proportion of fund exhibits randomness in absolute returns.

Focusing on rejection factors, still ARCH effects and structural breaks are found, but to a lesser extent. For relative returns, results are not that similar, and vary according to strategies and benchmarks. Nevertheless, volatility clustering and structural breaks are still present.

These empirical findings are important because they emphasize key facts of these specific strategies. Concerning clustering, i.e. persistence, managers try to arbitrage upon inefficiencies and opportunities of the market and thus managing is dependent to the cycle's trends. This clustering can also be due to the illiquidity nature of these strategies (i.e. Relative Value or Event-Driven). In the contrary, fewer funds within Equity Hedge and Macro strategies do cluster. Our results cast some doubts on the ability of the hedge fund industry to over-perform the market, and have strong implications in terms of portfolio allocation. Also, they allow us to rearrange the universe of hedge funds into trend-following funds and mean-reversion one. According to the ARCH effects and structural breaks factors, 60.88% of Equity Hedge strategy exhibit volatility clustering in their returns (when compared to the HFR Equity Hedge index). This strengthens the assumption that states that this style managing is based on long volatility strategies, in contrast to global macro strategies (43%). For structural breaks, 47% of Relative Value exhibits breaks. This could coincide with market crashes and correspond to substantial changes in the risk exposures of returns.

The contributions of our work are several. First, we have tested and analyzed a new framework to deal with randomness and persistence of hedge funds returns. Second, our methodology based on large sample of hedge funds, gives practical and theoretical answers to understand better the performance of certain hedge funds. Third, by investigating and testing clustering, ARCH and structural breaks, we pointed and investigate whether the risk exposures change with market conditions.

There are several avenues for future researches in this area as exploring the differences in the risk exposures of the 4 main strategies analyzed, or enlarge factors to complete the mapping answers. Also, from an econometric point of view, we have presented results based on raw data, i.e. not biased-corrected. Testing the robustness of our results could therefore be of particular interest using for instance the methodology suggested by Joenväärä et al. (2014) or by Hentati and Prigent (2011).

Acknowledgment

This project has received funding from the European Union's Seventh Framework Programme (FP7-SSH/2007–2013) for research, technological development and demonstration under grant agreement no 320270 - SYRTO.

References

Agarwal, V., Naik, N.Y., 2000a. Generalized style analysis of hedge funds. Journal of Asset Management 1, 93–109.
Agarwal, V., Naik, N.Y., 2000b. Multi-period performance persistence analysis of hedge funds. Journal of Financial and Quantitative Analysis 35, 327–342.
Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19, 716–723.
Amenc, N., El Bied, S., Martellini, L., 2003. Predictability in hedge fund returns. Financial Analysts Journal 59, 32–46.
Andrews, D.W K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62, 1383–1414.
Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66, 47–78.
Baquero, G., ter Horst, J., Verbeek, M., 2005. Survival, look-ahead bias and the persistence in hedge fund performance. Journal of Financial and Quantitative Analysis 40, 493–517.
Bollen, N.P.B., Whaley, R.E., 2009. Hedge fund risk dynamics: implications for performance appraisal. Journal of Finance 64, 985–1035.
Boyson, N.M., 2008. Hedge fund performance persistence: a new approach. Financial Analyst Journal 64, 27–44.
Brown, S., Goetzmann, W., Ibbotson, R., 1999. Offshore hedge funds: survival and performance: 1989–95. The Journal of Business 72, 91–117.
Brown, S., Goetzmann, W., Park, J., 2001. Careers and survival: competition and risk in the hedge fund and CTA industry. Journal of Finance 56, 1869–1886.
Caglayan, M., Edwards, F., 2001. Hedge fund performance and manager skill. Journal of Futures Markets 21, 1003–1102.
Carpenter, J.N., Lynch, A.W., 1999. Survivorship bias and attrition effects in measures of performance persistence. Journal of Financial Economics 54, 337–374.
Cho, J.S., White, H., 2011. Generalized runs tests for the IID hypothesis. Journal of Econometrics 162, 326–344.
Criton, G., Scaillet, O., 2011. Time-varying Analysis in Risk and Hedge Fund Performance: How Forecast Ability Increases Estimated Alpha. Working Paper, Swiss Finance Institute at HEC, University of Geneva.
De Souza, C., Gokcan, S., 2004. Hedge fund investing: a quantitative approach to hedge fund manager selection and de-selection. The Journal of Wealth Management 6, 52–73.
Dufour, J.-M., 1981. Rank tests for serial dependence. Journal of Time Series Analysis 2, 117–128.
Eling, M., 2009. Does hedge fund performance persist? overview and new empirical evidence. European Financial Management 15, 362–401.
Engle, R.F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 987–1007.
Fama, E., 1965. The behavior of stock market prices. Journal of Business 38, 34–105.
Fama, E.F., MacBeth, J.-D., 1973. Risk, return and equilibrium: empirical tests. Journal of Political Economy 81, 607–636.
Fung, W.K.H., Hsieh, D.A., 2000. Performance characteristics of hedge funds and CTA funds: natural versus spurious biases. Journal of Quantitative and Financial Analysis 35, 291–307.
Fung, W., Hsieh, D., Narayan, Y., Ramadorai, T., 2008. Hedge funds: performance, risk, and capital formation. Journal of Finance 63, 1777–1803.
Getmansky, M., 2012. The Life Cycle of Hedge Funds: Fund Flows, Size, Competition, and Performance. Working Paper, University of Massachusetts.
Getmansky, M., Lo, A.W., Makarov, I., 2004. An econometric model of serial correlation and illiquidity in hedge fund returns. Journal of Financial Economics 74, 529–609.
Gibbons, J.D., Chakraborti, S., 1992. Nonparametric Statistical Inference. Marcel Dekker, New York.
Granger, C., 1963. A quick test for serial correlation suitable for use with nonstationary time series. Journal of the American Statistical Association 58, 728–736.
Gyger, S., Gibson, R., Bares, P.-A., 2003. Performance in the hedge funds industry: an analysis of short- and long-term persistence. Journal of Alternative Investments 6, 25–41.
Hansen, B., 2000. Testing for structural change in conditional models. Journal of Econometrics 97, 93–115.
Hentati, R., Prigent, J.-L., 2011. VaR and omega measures for hedge funds portfolios: a copula approach. Bankers, Markets and Investors 110, 51–64.
Herzberg, M.M., Mozes, H.A., 2003. The persistence of hedge fund risk: evidence and implications for investors. The Journal of Alternative Investments 6, 22–42.
Jagannathan, R., Malakhov, A., Novikov, D., 2010. Do hot hands exist among hedge fund managers? An Empirical Evaluation. The Journal of Finance 65, 217–255.
Jensen, M., 1968. The performance of mutual funds in the period 1945–1964. Journal of Finance 23, 389–416.
Joenväärä, J., Kosowski, R., Tolonen, P., 2014. Hedge Fund Performance: What do we Know? Available at SSRN: <http://ssrn.com/abstract=1989410>.
Kosowski, R., Naik, N., Teo, M., 2007. Do hedge funds deliver alpha? a bootstrap and Bayesian approach. Journal of Financial Economics 84, 229–264.
Liang, B., 2000. Hedge funds: the living and the dead. Journal of Financial and Quantitative Analysis 35, 309–326.
Mood, A.M., 1940. The distribution theory of runs. The Annals of Mathematical Statistics 11, 367–392.
Park, J.M., Staum, J.C., 1998. Performance Persistence in the Alternative Investment Industry. Working Paper. Long Island University and Columbia University.
Russell, B., 2013. ARCH and Structural Breaks in United States Inflation. SIRE Discussion Paper 115.
Sharpe, W.F., 1966. Mutual fund performance. Journal of Business 39, 119–138.
Wald, A., Wolfowitz, J., 1940. On a test whether two samples are from the same population. Annals of Mathematical Statistics 2, 147–162.