

Does Hedge Fund Performance Persist? Overview and New Empirical Evidence

Martin Eling

Institute of Insurance Economics, University of St. Gallen, Kirchlistrasse 2, St. Gallen, Switzerland
E-mail: martin.eling@unisg.ch

Abstract

The contribution of this paper is to provide an overview and new empirical evidence on hedge fund performance persistence, which has been a controversial issue in the academic literature during the last several years. In the first step, we review recent studies and put them into a joint evaluation of hedge fund performance persistence. In the second step, the methodological framework developed in the overview is used to present new empirical evidence. We find different levels of performance persistence depending on the statistical methodology and the hedge fund strategy employed. In our study, performance persistence cannot be explained by the use of option-like strategies, but it can be partially explained by survivorship and backfilling bias. Differences among hedge fund strategies might be explained by return smoothing. Finally, we develop a rationale for choosing between different methodologies to measure performance persistence and conclude that the multi-period Kolmogorov-Smirnov test is the most useful for evaluating performance persistence of hedge funds.

Keywords: performance measurement, performance persistence, hedge funds

JEL classification: G11, G23, L25

1. Introduction

Most prospectuses of investment products note that past performance is not an indicator of future returns. Nevertheless, most investors allocate capital to different funds on basis of the funds' track records (see Capon *et al.*, 1996), which implies that they expect performance to be stable over time and that they expect some fund managers to provide better performance than others. In this context, analysing performance persistence deals with measuring whether some fund managers sustainably achieve higher (or lower) returns than their competitors.

I am grateful to an anonymous referee, Georgi Bontschev, Wolfgang Drobetz, Bill Fung, Marco Navone, Thomas Parnitzke, Hato Schmeiser, Frank Schuhmacher, Richard J. Taffler, and the participants of the 16th annual meeting of the European Financial Management Association (Vienna) and the IRC Conference on Hedge Fund Replication & Alternative Beta (Geneva) for valuable suggestions and comments on an earlier draft of this paper.

Hedge funds are ideal when studying performance persistence and manager skill. Unlike mutual funds, hedge funds do not track a benchmark but instead seek to exploit mispricing and provide absolute returns (see Brown *et al.*, 1999). Thus, analysing performance persistence for mutual fund managers, who follow traditional benchmark tracking strategies, makes less sense than doing so for hedge fund managers, as mutual fund managers have less opportunity than hedge fund managers to display differential skills. Nevertheless, the issue of performance persistence has been extensively studied for traditional mutual funds, e.g., by Grinblatt and Titman (1992), Brown and Goetzmann (1995), and Carhart (1997). Most of these studies confirm that traditional strategies such as investing in mutual funds on average underperform passive investment strategies and that there is only limited evidence for performance persistence with traditional mutual funds (see Droms, 2006, for an overview). This situation might be different with hedge funds, as the few fund managers who have beaten passive strategies tend to move to alternative investments and start their own hedge fund (see Agarwal and Naik, 2000a). For that reason, hedge fund performance persistence has captured a great deal of attention in academic literature. However, in contrast to articles on performance persistence for traditional mutual funds, studies on hedge fund performance come to widely divergent conclusions, possibly due to the use of different databases, investigation periods, performance measures, and statistical methodologies. For example, Agarwal and Naik (2000a) consider 746 funds from the hedge fund research database between 1982 and 1998 and find performance persistence at quarterly horizons. Brown *et al.* (1999) consider 399 funds from the US Offshore Funds Directory between 1989 and 1995 and find no evidence of persistence in hedge fund performance at a yearly horizon. Because of the large differences between these studies it is difficult to get a broad and clear picture of hedge fund performance persistence. What is missing is an overview of performance measures, statistical methodologies, and results.

The first contribution of this paper is to provide such an overview. We will look at 25 studies on hedge fund performance persistence, summarising the databases, performance measures, statistical methodologies, and results. The second contribution of this paper is to use the methodological framework developed in the overview to present new empirical evidence on performance persistence of hedge funds. For this we use the Center for International Securities and Derivatives Markets (CISDM) database, which is one of the largest hedge fund databases ever analysed for this purpose. It contains data on 4,165 hedge funds and 2,021 funds of hedge funds. The analysis covers the years 1996 through 2005, which is advantageous for two reasons. First, the results will not suffer as much from the survivorship and backfilling biases that plague much of the older hedge fund research.¹ Second, this time period contains bullish as well as bearish markets; many other studies are limited to the analysis of bullish

¹ This does not mean that these biases do not exist, but that we are able to quantify them. The important point here is that major hedge fund data vendors did not cover dissolved funds before 1994. Hedge fund data before 1994 are thus not reliable and should not be used in research. For this reason, Capocci and Hübner (2004) decided to exclude the largest part of their hedge fund data from 1984 to 2000 in their study of hedge fund performance. Liang and Park (2007) also start their analysis in 1995 due to that problem. The unreliability of data before 1994 is also discussed by Fung and Hsieh (2000), Liang (2000), and Li and Kazemi (2007).

markets.² After measuring persistence, we empirically analyse reasons for hedge fund performance persistence – the third contribution of this paper. And, finally, based on the given theoretical and empirical evidence, our fourth contribution is the development of a rationale for choosing between different performance persistence methodologies.

The empirical section of the paper brings to light several findings of interest. First, there are differences in results depending on the methodology used to assess performance persistence. Correlation-, contingency-table-, and regression-based tests provide evidence of performance persistence, whereas the Kolmogorov-Smirnov test reveals very little persistence. We thus conclude that the use of different methodologies is one of the key reasons for the varied results found in the literature. Second, there are large differences in results depending on the hedge fund strategy. With Convertible Arbitrage and Emerging Markets hedge funds we find very high levels of persistence, but Equity Long Only hedge funds provide low levels of persistence. Therefore, we conclude that persistence is related to the type of hedge fund strategy followed. Finally, comparing measures used to assess performance persistence like, e.g., raw returns, alphas, and appraisal ratios generates very similar results. The use of different performance measures is thus not the reason for the conflicting results found in performance persistence literature.

The analysis of reasons for persistence reveals three important insights into hedge fund performance measurement. (1) The high degree of performance persistence found with some methodologies cannot be explained by the fact that many hedge funds use option-like strategies, but (2) it can be at least partially explained by survivorship and backfilling bias. (3) Smoothing of returns also biases upward the levels of persistence of many hedge fund strategies. Our final result is that we recommend using the multi-period Kolmogorov-Smirnov test as employed by Agarwal and Naik (2000a) when evaluating performance persistence of hedge funds because, among other things, it avoids the assumption of normality of returns and is the most robust method.

The remainder of this paper is organised as follows. Section 2 is an overview of 25 studies on hedge fund performance persistence. The empirical examination of hedge fund performance persistence as well as the analysis of reasons for persistence is performed in Section 3. In Section 4 we develop a rationale for choosing between different performance persistence methodologies. We conclude in Section 5.

2. Literature Review on Hedge Fund Performance Persistence

In this section we provide an overview of 25 academic studies on hedge fund performance persistence. Table 1 summarises the main features of these studies.³ The first column gives the authors. The second and third columns display the database and the number of funds considered. The fourth and fifth columns show the investigation period and

² See, e.g., Amenc *et al.* (2003), Baquero *et al.* (2005), and Brown *et al.* (1999). Although many hedge funds do not use trend-following strategies, Capocci *et al.* (2005) found that the market phase may influence the results. For that reason it seems appropriate to have bullish as well as bearish market phases in the study. Ding and Shawky (2007) also emphasise the importance of considering different market cycles when analysing hedge fund performance.

³ See Géhin (2004), Henn and Meier (2004), Schneeweis *et al.* (2002), Gregoriou *et al.* (2002), and Gupta *et al.* (2003), which all provide smaller overviews on a subset of these studies.

Table 1
Studies on hedge fund performance persistence

This table reports the main characteristics and results for 25 studies on hedge fund performance persistence between 1998 and 2005. Column 1 gives the authors, Column 2 the considered database, Column 3 the number of funds, Column 4 the investigation period, Column 5 the time horizon, Column 6 the performance measures, Column 7 the statistical methodology, and Column 8 the results. Abbreviations: CISDM: Center for International Securities and Derivatives Markets, CSFB: Credit Suisse First Boston, HFR: Hedge Fund Research, LSE: Long/Short-Equity, MAR: Managed Account Reports, MSCI: Morgan Stanley Capital International, TASS: Tremont Advisory Shareholders Services.

Authors	Database	Number of Funds	Investigation Period	Time Horizon (Months)	Performance Measure	Statistical Methodology	Result
Agarwal <i>et al.</i> (2007)	CISDM, HFR, MSCI, TASS	7535	1994–2002	12	return	chi square, regression	persistence at yearly horizon
Agarwal and Naik (2000a)	HFR	746	1982–1998	3, 6, 12	alpha, appraisal ratio	cross product ratio, chi square, regression, Kolmogorov-Smirnov	persistence at quarterly horizon
Agarwal and Naik (2000b)	HFR	167	1995–1998	3	alpha, appraisal ratio	cross product ratio, regression	persistence at quarterly horizon
Amenc <i>et al.</i> (2003)	CSFB/Tremont indices	9	1994–2000	1	return	regression	persistence at monthly horizon
Baquero <i>et al.</i> (2005)	TASS	1797	1994–2000	3, 12, 24	return, alpha	none (descriptive comparison of rankings)	persistence at quarterly and yearly horizon, but not at two-year horizon
Barès <i>et al.</i> (2003)	Financial Risk Management	4934	1992–2000	1, 3, 6, 12	return, alpha	descriptive comparison of rankings, regression	persistence at monthly and quarterly horizon
Boyson and Cooper (2004)	TASS	1659	1994–2000	3	alpha	regression	persistence at quarterly horizon

Table 1
Continued.

Authors	Database	Number of Funds	Investigation Period	Time Horizon (Months)	Performance Measure	Statistical Methodology	Result
Brown and Goetzmann (2003)	TASS	1295	1992–1998	12	return	regression	no persistence at yearly horizon
Brown <i>et al.</i> (1999)	US Offshore Funds Directory	399	1989–1995	12	return, alpha, appraisal ratio	regression	no persistence at yearly horizon
Capocci <i>et al.</i> (2005)	CISDM, HFR, TASS	2894	1994–2002	12	alpha	regression	no persistence at yearly horizon
Capocci and Hübner (2004)	HFR, MAR	2796	1988–1995	12	alpha	regression	no persistence at yearly horizon
Chen and Passow (2003)	TASS, HFR	76 (only LSE)	1990–2002	12	alpha	regression	no persistence at yearly horizon
De Souza and Gokcan (2004)	HFR	314	1997–2002	24, 36	return, standard deviation, Sharpe ratio	cross product ratio, regression,	no persistence at two-and three year horizon with returns, but with risk
Edwards and Caglayan (2001)	MAR	1665	1990–1998	12, 24	alpha	cross product ratio, regression	persistence at one-year and two-year horizons
Gregoriou and Rouah (2001)	Zurich/LaPorte	(n/a)	1988–1999	12	alpha	descriptive comparison of rankings	no persistence at yearly horizon

Harri and Brorsen (2004)	LaPorte	1209	1977–1998	1, 2, 3, ... to 24	return, information ratio, Sharpe ratio, alpha	Spearman, regression	persistence at all horizons
Henn and Meier (2004)	EurekaHedge	1217	1994–2004	1, 3, 12	return	cross product ratio	persistence at monthly, quarterly and yearly horizon
Herzberg and Mozes (2003)	HedgeFund.net, Altvest, Spring Mountain Capital	3300	1995–2001	12	return, Sharpe ratio, max. draw. standard deviation, correlation	rank information coefficient	no persistence at yearly horizon with returns, but with risk
Jagannathan <i>et al.</i> (2006)	HFR	2141	1996–2003	36	alpha	regression	persistence at three-year horizon
Kat and Menexe (2003)	TASS	324	1994–2001	36	return, standard deviation, skewness, kurtosis, correlation	cross product ratio, regression	no persistence at three-year horizon with returns, but with the higher mo.
Koh <i>et al.</i> (2003)	EurekaHedge, AsiaHedge	3810 (only Asian funds)	1999–2003	1, 2, 3, 6, 9, 12	return, alpha	cross product ratio, chi square, Kolmogorov-Smirnov	persistence at monthly and quarterly but not at yearly horizon
Kosowski <i>et al.</i> (2007)	TASS, HFR, CISDM, MSCI	9338	1990–2002	12	alpha	regression, bootstrap approach, Bayesian approach	persistence at yearly horizon
Kouwenberg (2003)	Zurich (MAR)	2614	1995–2000	36	return, alpha, Sharpe ratio	chi square	persistence at three-year horizon
Malkiel and Saha (2005)	TASS	2065	1996–2003	12	return	chi square	no persistence at yearly horizon
Park and Staum (1998)	TASS	n/a	1986–1997	12	appraisal ratio	chi square, Spearman	persistence at yearly horizon

the time horizon. Columns 6 and 7 present the performance measures and the statistical methodology. Finally, in Column 8 the results of the studies are summarised. The rest of this section is subdivided by the columns listed in Table 1.

2.1. Database and number of funds

It is documented in literature that there are differences between hedge fund databases.⁴ For that reason, it is relevant to ask which database to consider in the empirical study. There are three main hedge fund database providers that are usually used in academic studies (see Liang and Park, 2007): Center for International Securities and Derivatives Markets (CISDM, formerly Managed Account Reports, contains 4,200 active and 2,000 inactive funds at present), Hedge Fund Research (HFR, 6,000 active and 3,500 inactive funds), and Tremont Advisory Shareholders Services (TASS, 3,900 active and 2,400 inactive funds). These three databases have also been used in combination: e.g., Ackermann *et al.* (1999) and Capocci and Hübner (2004) used a combination of HFR and CISDM, whereas Chen and Passow (2003) used a combination of TASS and HFR. Moreover, some smaller databases have been the subject of performance persistence studies. For example, Brown *et al.* (1999) collect hedge fund data from the US Offshore Funds Directory and Koh *et al.* (2003) consider the EurekaHedge and AsiaHedge database.

The number of funds contained in the hedge fund databases and analysed within performance persistence studies has risen rapidly over the last few years. Most older studies, that is prior to the year 2000, do not analyse more than 1,000 funds; the recent study by Kosowski *et al.* (2007) combines CISDM, HFR, and TASS with the new MSCI database, resulting in the largest dataset used for academic research to date – it contains 6,392 live and 2,946 dead funds.

2.2. Investigation period and time horizon

There is no clear answer in the literature to the question of which investigation period to choose in measuring performance persistence. There are studies with very short investigation periods of only three years (see Agarwal and Naik, 2000b), but also studies with time periods up to 21 years (see Harri and Brorsen, 2004). The 25 studies considered in this literature overview have a mean investigation period of 8.5 years with a median of 7 years.

There are some important considerations to keep in mind when choosing the investigation period. First, as mentioned in the introduction, the consideration of returns prior to 1994 may not be too meaningful due to the survivorship bias in hedge fund returns prior to that year (see Liang, 2000). Second, it is important not to use too long time periods, as hedge fund managers typically do not work for more than one decade with the same hedge fund (see Boyson and Cooper, 2004). Most studies measure the fund performance but not that of the underlying manager since they cannot control for a change in fund management. This is problematic as performance persistence is mostly associated with the special skills of a fund manager. Therefore, we recommend choosing

⁴ E.g., Liang (2000) finds significant differences in fund returns, attrition rate, and survivorship bias in the TASS and HFR databases. For an overview of different hedge fund databases see Goltz *et al.* (2007).

a time period which is not extremely long (e.g., not longer than 10 years) and using returns not older than 1994.

Another important decision involves the choice of time horizon. It clearly makes a difference whether yearly, quarterly, or monthly returns are considered. For example, Harri and Brorsen (2004) compare persistence for a horizon of 1 month up to 24 months and find large differences in significance levels of persistence. The same result is found by Henn and Meier (2004) and Koh *et al.* (2003). We will reconsider this aspect in the discussion of the main results (see Section 2.5).

2.3. Measure

A wide range of measures is used to analyse hedge fund performance persistence. These can be broken down into five groups: return, risk, higher moments, correlation, and risk-adjusted performance (an overview of the measures analysed in the 25 studies on hedge fund performance persistence is given later in Table 10).

The first group are raw-return-based measures. While most studies concentrate on post-fee returns, Brown *et al.* (1999) and Koh *et al.* (2003) also analyse pre-fee returns. Consideration of fees can provide additional information about the fund manager's performance because there is a difference between a fund that has a gross return of 10% and a net return of 5%, and a fund that has a gross return of 20% and a net return of 5%. Furthermore, risk measures (standard deviation and maximum drawdown), higher moments (skewness and kurtosis), and correlations with stock and bond markets were considered in persistence studies.

The most important measures to analyse performance persistence are risk-adjusted performance measures. Among these are the information ratio, the Sharpe ratio, alpha, and the appraisal ratio. The information ratio measures the relationship between the fund's return and its standard deviation (see Goodwin, 1998). The Sharpe ratio considers the relationship between the excess return (return minus the risk-free interest rate) and the standard deviation of the returns (see Sharpe, 1966). Alpha is the intercept of the regression of several market factors on the hedge fund excess returns. The appraisal ratio is the relationship between alpha and the residuals standard deviation of the above-mentioned regression.

Alpha was introduced by Jensen (1968) in the context of a single index model as a regression of the market excess return on the fund excess return. This single factor modelling can be extended to a multi-factor framework in order to improve the portion of variance explained by the regression. One example is the Fama and French (1993) model, which has two additional factors, one for size (SMB, small minus big) and the second being the ratio of book-to-market (HML, high minus low book-to-market price ratio). A number of researchers have stressed the importance of considering hedge-fund-specific style factors when studying hedge fund performance (see, e.g., Fung and Hsieh, 1997; Brown *et al.*, 1999, Gibson and Gyger, 2007). Thus, many models include hedge fund style factors along with common risk factors. The style factors are usually hedge fund indices (e.g., the Tremont indices; see Brown *et al.*, 1999) or an average return of all the funds using the same strategy in a database (Agarwal and Naik, 2000a) (see Table A1 in the Appendix for the different definitions of alpha).

2.4. Statistical methodology

Agarwal and Naik (2000a) distinguish between two-period and multi-period statistical approaches that can be used to examine performance persistence. In the first case,

two consecutive time units, e.g., months, are compared to each other while in the multi-period case more than two consecutive time units are considered. The statistical methodologies that build upon the two-period framework can be further distinguished into nonparametric and parametric approaches. To the nonparametric approaches belong the contingency-table-based cross-product ratio test and chi-square test, the correlation-based rank information coefficient test, and Spearman's rank correlation test. The parametric approach is a linear regression. In the multi-period framework, a Kolmogorov-Smirnov test can be applied.⁵

The contingency-table-based methods are based on the construction of tables of winners and losers. Winners are funds whose performance is higher than the median return of all funds following the same strategy over the chosen period, and losers are funds whose performance is lower than the median performance of all funds following the same strategy. Persistent are those funds that are winners (WW) and losers (LL) in both periods. Winners during the first period that are losers during the second period are denoted WL or LW in the opposite case. The *cross-product ratio (CPR) test* (also called log-odds ratio test; see, e.g., Agarwal and Naik, 2000a) is the ratio of the funds which persistent to the funds that did not persist:

$$CPR = (WW \cdot LL)/(WL \cdot LW). \quad (1)$$

CPR is equal to 1 in the null hypothesis of no persistence, i.e., each of the four categories WW, LL, WL, and LW represent 25% of all funds. The statistical significance of CPR can be tested using the standard error $\alpha_{\ln(CPR)}$ of the natural logarithm of CPR. The resulting Z-statistic is the ratio of the natural logarithm of the CPR to the standard error of the natural logarithm. Corresponding to the standard normal distribution, a value greater than 1.96 (2.58) indicates significant persistence at the 5% (1%) confidence level:

$$Z = \ln(CPR)/\alpha_{\ln(CPR)} \ln(CPR) / \sqrt{\frac{1}{WW} + \frac{1}{WL} + \frac{1}{LW} + \frac{1}{LL}}. \quad (2)$$

In the *chi-square test* (see, e.g., Park and Staum, 1998), the observed frequency distribution of WW, WL, LW, and LL is compared with the expected frequency distribution:

$$\chi^2 = (WW - D1)^2/D1 + (WL - D2)^2/D2 + (LW - D3)^2/D3 + (LL - D4)^2/D4, \quad (3)$$

with $D1 = (WW + WL) \cdot (WW + LW)/I$, $D2 = (WW + WL) \cdot (WL + LL)/I$, $D3 = (LW + LL) \cdot (WW + LW)/I$, and $D4 = (LW + LL) \cdot (WL + LL)/I$; I is the number of all funds. Following the chi-square distribution with one degree of freedom, a value of χ^2 greater than 3.84 (6.64) indicates significant persistence at the 5% (1%) confidence level.

The *rank information coefficient* (RIC; used by Herzberg and Mozes, 2003) measures the correlation between the value of a given variable for period 1 (e.g., the prior month)

⁵ An earlier draft of this paper also contained the Hurst exponent as a methodology for measurement of persistence. However, the application of this methodology with hedge funds is problematic, as only a very few monthly returns are available. Against it, in literature it is emphasised that the Hurst exponent needs much more data. See Lo (1991) and Couillard and Davison (2005) for more details on the Hurst exponent.

and its value for period 2 (e.g., the subsequent month). The statistical significance of the rank information coefficient can be tested using the Fisher T-statistic:

$$T_{RIC_i} = \sqrt{N_i - 2} \left(RIC_i / \sqrt{1 - RIC_i^2} \right), \quad (4)$$

with N as the number of returns of Fund i . Corresponding to the T-distribution, a value greater than 1.96 (2.58) indicates significant persistence at the 5% (1%) confidence level.

Under *Spearman's rank correlation test* (see, e.g., Park and Staum, 1998), performance rankings are compared for different time periods. In case of persistence, the correlation between the rankings of two consecutive periods should be relatively high, while a correlation coefficient of 0 indicates the absence of persistence. The statistical significance of the Spearman rank correlation coefficient can be tested using the Fisher T-statistic (see Equation (4)).

For the *regression-based parametric method* (see, e.g., Brown *et al.*, 1999), the measurement value (return, alpha, or another measure; see Section 2.3.) during the current period is regressed on the measurement value of the previous period. A positive significant slope coefficient indicates performance persistence. The statistical significance of the slope can be tested using the T-statistic. Corresponding to the standard normal distribution, a t-value greater than 1.96 (2.58) indicates significant persistence at the 5% (1%) confidence level:

$$r_t = \alpha + \beta \cdot r_{t-1}. \quad (5)$$

With the *Kolmogorov-Smirnov goodness-of-fit test* (used, e.g., by Agarwal and Naik, 2000a), the traditional two-period framework is extended to a multi-period approach in an effort to achieve more robust results. A series of wins and losses for each fund is constructed and the observed frequency distribution is compared with the theoretical frequency distribution of two or more consecutive wins and losses. For example, under the null hypothesis of no persistence, the theoretical probability of WWW and LLL is one-eighth and that of WWWW and LLLL is one-sixteenth. Using the two-sample Kolmogorov-Smirnov test, we check whether the observed distribution is statistically different from the theoretical distribution.

2.5. Results

Table 2 summarises the results of the 25 studies. '↑' indicates that performance persistence was found; '↓' indicates no performance persistence. '-' means that the time horizon was not analysed.

The main results of hedge fund performance persistence studies can be summarised as follows. First, short-term persistence for horizons of up to six months is reported by nearly all authors. Second, evidence for longer horizons is mixed, as the studies come to conflicting conclusions. For example, at the annual horizon, there are eight studies finding performance persistence, whereas ten studies reject the hypotheses of persistence in hedge fund performance. Agarwal and Naik (2000a) and Harri and Brorsen (2004) report persistence for both short- and long-term horizons. However, both studies mention that the return persistence significance levels weakens as the measurement horizon is extended.

Table 2
Hedge Fund Performance Persistence

This table reports the results for 25 studies on hedge fund performance persistence. The first column gives the authors and the second the results. We distinguish between six time horizons (from 1 to 36 months). ‘↑’ indicates that performance persistence was found; ‘↓’ indicates no performance persistence. ‘-’ means that the time horizon was not analysed.

Authors	Time horizon (month)					
	1	3	6	12	24	36
Agarwal <i>et al.</i> (2007)	-	-	-	↑	-	-
Agarwal and Naik (2000a)	-	↑	↑	↑	-	-
Agarwal and Naik (2000b)	-	↑	-	-	-	-
Amenc <i>et al.</i> (2003)	↑	-	-	-	-	-
Baquero <i>et al.</i> (2005)	-	↑	-	↑	↓	-
Barès <i>et al.</i> (2003)	↑	↑	↓	↓	-	-
Boyson and Cooper (2004)	-	↑	-	-	-	-
Brown and Goetzmann (2003)	-	-	-	↓	-	-
Brown <i>et al.</i> (1999)	-	-	-	↓	-	-
Capocci <i>et al.</i> (2005)	-	-	-	↓	-	-
Capocci and Hübner (2004)	-	-	-	↓	-	-
Chen and Passow (2003)	-	-	-	↓	-	-
De Souza and Gokcan (2004)	-	-	-	-	↓	↓
Edwards and Caglayan (2001)	-	-	-	↑	↑	-
Gregoriou and Rouah (2001)	-	-	-	↓	-	-
Harri and Brorsen (2004)	↑	↑	↑	↑	↑	-
Henn and Meier (2004)	↑	↑	-	↑	-	-
Herzberg and Mozes (2003)	-	-	-	↓	-	-
Jagannathan <i>et al.</i> (2006)	-	-	-	-	-	↑
Kat and Menexe (2003)	-	-	-	-	-	↓
Koh <i>et al.</i> (2003)	↑	↑	↑	↓	-	-
Kosowski <i>et al.</i> (2007)	-	-	-	↑	-	-
Kouwenberg (2003)	-	-	-	-	-	↑
Malkiel and Saha (2005)	-	-	-	↓	-	-
Park and Staum (1998)	-	-	-	↑	-	-
Total	5 ↑	8 ↑	3 ↑ 1 ↓	8 ↑ 10 ↓	2 ↑ 2 ↓	2 ↑ 2 ↓

We conclude that there is persistence in hedge fund performance at short horizons of up to six months, but that the longer the time horizon, the lower is the significance of performance persistence. The following are several additional insights revealed by comparing the 25 studies:

- It was analysed whether winners or the losers persist. Agarwal and Naik (2000a) find that persistence is driven mostly by losers. However, the level of persistence found by Edwards and Caglayan (2001) holds across both winners and losers.
- There is no consensus in the literature about whether fund strategy is a driver of persistence. Agarwal and Naik (2000a) find that persistence is not related to the type of strategy followed. However, Brown and Goetzmann (2003) and Harri and Brorsen

(2004) find that persistence of fund returns has a great deal to do with the style of fund management.

- Regarding whether survivorship bias might influence results, Malkiel and Saha (2005) find no persistence if all funds are considered, but slightly more persistence if dead funds are dropped from the database. This is in line with Capocci and Hübner (2004). They assume that the small degree of performance persistence they found from 1985 to 1993 is due to the absence of dissolved funds.
- The two-period framework was compared to the multi-period framework. Agarwal and Naik (2000a) find that the level of persistence observed in a multi-period framework is considerably smaller than that observed under the traditional two-period framework.
- Some fund and managerial characteristics were identified as drivers of persistence. Agarwal *et al.* (2007) find that hedge funds with greater managerial incentives (e.g., larger incentive fee) and a higher degree of managerial flexibility (e.g., longer lockup period) have superior performance. Boyson and Cooper (2004) show that young, previously good managers outperform old, previously poor managers.
- Finally, many authors discussed reasons for persistence. One possible reason for short-term performance persistence could be that monthly returns are smoothed out, either due to holding illiquid securities or managed returns (see Henn and Meier, 2004; Kosowski *et al.*, 2007). Barès *et al.* (2003) and Jagannathan *et al.* (2006) believe that short-term persistence is related to the hot-hands effect documented in mutual fund literature (see Hendricks *et al.*, 1993). The hot-hands effect occurs when securities held by funds that had better performance during one year realise superior returns the following year.

We thus can identify several consistent general themes in the literature in regard to short-term persistence, but nothing even this clear can be discovered regarding long term-persistence and other characteristics. The following empirical study will shed light on these issues.

3. Empirical Evidence on Hedge Fund Performance Persistence

The literature study reveals a heterogeneous picture of hedge fund performance persistence. The different results may be due to the use of different databases, investigation periods, performance measures, and statistical methodologies. To obtain a more accurate picture of hedge fund performance persistence, we will use the whole framework discussed in the overview and present new empirical evidence on all tests and measures.

3.1. Data

We obtained data on 6,186 funds for the period between January 1996 and December 2005 from the Center for International Securities and Derivatives Markets (CISDM). The CISDM database has been subject of many academic studies (for the properties of this database, see, e.g., Edwards and Caglayan, 2001; Kouwenberg, 2003; Capocci and Hübner, 2004; Ding and Shawky, 2007). The database contains 4,165 hedge funds and 2,021 funds of hedge funds. Depending on the strategy, the database can be broken down into 22 hedge fund strategy and 7 funds of funds strategy groups.⁶

⁶ Due to an insufficient number of funds, we combined the hedge fund strategies Capital Structure Arbitrage, Market Timing, Option Arbitrage, Other Relative Value, and Regulation

Twenty-seven funds that appeared twice in the database and one fund that only reports returns on a quarterly basis were deleted. This reduces our sample to 4,143 hedge funds and 2,015 funds of hedge funds. We require all funds to have at least 24 monthly returns because this is the minimum for calculating meaningful performance measures (see Ackermann *et al.*, 1999; Gregoriou, 2002; Capocci and Hübner, 2004; Liang and Park, 2007).⁷ Eliminating those 1,844 funds with less than 24 monthly returns reduces our sample to 2,936 hedge funds and 1,378 funds of hedge funds.

Like other hedge fund databases, the CISDM database suffers from survivorship bias. There are two common definitions of survivorship bias: the difference in fund returns between the surviving funds and the dissolved funds (see Ackermann *et al.*, 1999) or the difference between the returns of the surviving funds and all funds (see Liang, 2000). We use Liang's (2000) definition and find a survivorship bias of 0.08% per month with hedge funds (detailed calculations are available upon request), which is comparable to other values found in the literature (see, e.g., Ackermann *et al.*, 1999; Liang, 2000). The fact that compared to hedge funds, the attrition rate and the survivorship bias are lower with funds of hedge funds is well documented in literature (see Liang, 2000). In our sample, survivorship bias for funds of hedge funds only amounts to 0.02%.

If new funds are added to a database, historical returns are backfilled, which may also cause an upward bias in performance measurement results. We follow Brown *et al.* (1999), Fung and Hsieh (2000), and Capocci and Hübner (2004) and calculate backfilling bias by stepwise deleting the first 12, 24, 36, 48, and 60 months of returns (detailed calculations are available upon request). The monthly return of the portfolio that invests in all funds is 1.03% for hedge funds and 0.68% for funds of funds. Eliminating the first 12 (24, 36, 48, 60) months of returns for each fund reduces the return about 0.18% (0.38%, 0.38%, 0.40%, 0.31%) for hedge funds and 0.03% (0.02%, 0.06%, 0.08%, 0.08%) for funds of funds. Again, these values are comparable to other values in the literature. For example, Fung and Hsieh (2000) find that backfilling bias is noticeably lower with funds of funds than with hedge funds. However, we cannot confirm the finding of Capocci and Hübner (2004) that backfilling bias is larger the longer the estimation period is.

Hedge funds and funds of hedge funds are compared with the passive benchmark indices used as market factors in alpha measurement literature (see Table A1 in the Appendix). The equity market proxy is the value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks used in Fama and French (1993) and Carhart (1997). Furthermore, the MSCI World excluding the USA, the MSCI Emerging Markets, Fama and French's (1993) factors for size (SMB) and book-to-market (HML), and Carhart's (1997) momentum factor are used as passive equity benchmark indices. Bonds are compared by way of the Lehman US Aggregate Bond, the JP Morgan Global Government Bond, the JP Morgan Emerging Market Bond, and the Lehman BAA Corporate Bond indices. Finally, we use the JPM US Cash 1 Month Index and the Goldman Sachs

D into one new hedge fund strategy – Other – and the funds of funds strategies Conservative, Invest Funds in Parent Companies, and Opportunistic into one new funds of funds strategy – Other. In the empirical section of the paper, we thus consider 18 hedge fund strategy and 5 funds of funds strategy groups.

⁷ Fung and Hsieh (1997) and Liang (2000) eliminated funds with less than 36 monthly returns, which would reduce our sample to 2,319 hedge funds and 1,058 funds of hedge funds. As a robustness test, we conducted the performance measurement for those funds with at least 36 monthly returns and found robust results.

Commodity Index as passive benchmark indices for currencies and commodities. The data on the passive benchmark indices were collected from the Kenneth R. French website and the Thomson Financial Datastream database.

3.2. Descriptive statistics and performance measurement

Descriptive statistics and performance measurement results for the 4,314 funds and the passive strategy indices are presented in Table 3. The funds are subdivided by strategy group in the first column. The second, third, and fourth columns display the number of funds, subdivided in all, living, and dead funds. Columns 5 through 8 show the first four moments of the return distribution (mean value, standard deviation, skewness and excess kurtosis). Column 9 shows the results of the Jarque-Bera test, which are displayed as the portion of funds for which the assumption of normally distributed returns must be rejected at the 5% significance level. Mean excess returns are calculated in Column 10 using the one-month Treasury bill rate provided by Ibbotson Associates. The Sharpe ratio (Column 11) is computed as the mean excess return divided by the standard deviation.

In addition to looking at raw returns and the Sharpe ratio, we analyse two performance measurement models and the resulting alpha values. The first version of alpha (α_m) is market adjusted and very similar to the combined model presented by Capocci and Hübner (2004). It is calculated for each fund as the intercept of a regression of the returns of the 12 benchmark indices displayed in Table 3 on the fund's excess returns. The second version of alpha (α_{ms}) is market and hedge fund style adjusted. The market factors for the alpha calculation are again the 12 benchmark indices. For calculation of the hedge fund style factor, we follow Agarwal and Naik (2000a) and use the average return of the funds following the same strategy. The two versions of alpha and the associated R^2 are displayed in Columns 12 to 15 of Table 3.⁸

The performance measurement results provide significant evidence of superior hedge fund performance over long periods of time. The highest mean return of all strategies was achieved by Sector (1.38%), followed by funds that have no strategy description (1.23%) and Emerging Markets (1.21%). Strategies that offer the lowest mean return are Short Bias (0.43%), Fixed Income Arbitrage (0.51%), and Equity Market Neutral (0.61%). Compared to the benchmark indices both hedge funds and funds of hedge funds provide relatively high returns.

Taking investment risk into account through the Sharpe ratio, several strategies offer a good tradeoff between risk and return. Fixed Income – MBS (0.32), Other (0.32), and Relative Value Multi Strategy (0.31) have the highest Sharpe ratios. The lowest Sharpe ratio is obtained by Short Bias (0.02). Most funds provide a very high Sharpe ratio compared to the benchmark indices.

Some investors might be concerned with central tendencies of the return distribution (mean value, standard deviation); others may care more about the extreme values. For these investors, it is interesting to consider skewness, excess kurtosis, and the results of the Jarque-Bera test. The returns of most hedge funds and funds of funds are not normally distributed. The rejection rate for the Jarque-Bera test varies between 27% for

⁸ The full results are available upon request. These show that especially the Fama and French factors for size (SMB) and book-to-market (HML) are highly significant for most hedge funds and funds of funds.

Table 3
Descriptive statistics and performance measurement

This table shows the number of funds (subdivided in all, living, and dead funds, Columns 2, 3, and 4), the first four moments of the return distribution (mean value, standard deviation, skewness, and excess kurtosis, Columns 5–8), the portion of funds for which the assumption of normally distributed returns must be rejected at the 5% significance level (Jarque-Bera test; Column 9), mean excess returns (Column 10) using the one-month Treasury bill rate, the Sharpe ratio (Column 11), average age (Column 12), average annual management fee (Column 13), average annual incentive fee (Column 14), and average minimum investment (Column 15) for the funds in our database of 2,936 hedge funds (Panel A, broken down into 22 strategies) and 1,378 funds of funds (Panel B, broken down into 7 strategies). The table also reports the first for moments and the Jarque-Bera test results for the passive strategy indices (Panel C). * (**, ***) indicates significance at 10% (5%, 1%) level.

	Number of Funds				Returns					Market Adjusted Model		Market and Style Adjusted Model		
	Total	Living	Dead	Mean (%)	St. Dev. (%)	Skew.	Exc. Kurt.	Jarque-Bera (%)	Mean Excess Return	Sharpe Ratio	α_m	R^2_m	α_{ms}	R^2_{ms}
A. Hedge funds (2,936 funds)														
Convertible Arbitrage	159	82	77	0.82	2.13	-0.17	2.86	0.58	0.56	0.26	0.16***	0.37	-0.30***	0.44
Distressed Securities	113	66	47	1.14	3.64	0.11	3.14	0.64	0.87	0.24	0.79***	0.49	-0.29	0.56
Emerging Markets	257	137	120	1.21	7.05	-0.32	5.20	0.61	0.92	0.13	0.97***	0.56	-0.77***	0.60
Equity Long Only	57	33	24	1.19	6.43	-0.04	0.84	0.37	0.94	0.15	0.48***	0.68	-0.45	0.70
Equity Long/Short	1109	512	597	1.14	5.09	0.25	2.40	0.52	0.87	0.17	0.34***	0.53	-0.47***	0.57
Equity Market Neutral	164	92	72	0.61	2.74	0.13	2.72	0.44	0.38	0.14	0.16**	0.43	-0.04	0.47
Event Driven Multi Strategy	120	85	35	1.04	3.14	-0.11	3.11	0.63	0.81	0.26	0.68***	0.48	0.09	0.55
Fixed Income	65	55	10	0.69	1.71	-0.13	1.37	0.34	0.49	0.29	0.24***	0.56	0.07	0.59
Fixed Income—MBS	54	32	22	0.91	2.10	-1.51	15.69	0.87	0.67	0.32	0.76***	0.25	0.64***	0.29
Fixed Income Arbitrage	111	51	60	0.51	2.13	-0.84	6.53	0.67	0.25	0.12	0.56***	0.41	0.03	0.45
Global Macro	142	67	75	0.86	4.80	0.16	2.34	0.49	0.57	0.12	0.16	0.45	-0.25	0.48
Merger Arbitrage	104	37	67	0.70	2.00	-0.32	3.72	0.72	0.41	0.20	0.14	0.38	-0.11	0.41
Multi Strategy	47	43	4	1.00	2.89	0.25	2.43	0.51	0.80	0.28	0.51***	0.37	0.03	0.46
No Strategy	47	15	32	1.23	3.69	0.12	2.05	0.64	0.95	0.26	0.69***	0.46	-0.01	0.52

Other	36	28	8	1.01	2.48	0.78	3.06	0.75	0.79	0.32	0.14	0.42	0.12	0.45
Relative Value Multi Strategy	61	51	10	0.90	1.71	0.02	5.46	0.61	0.80	0.31	0.41***	0.37	0.09	0.44
Sector	257	108	149	1.38	7.06	0.44	3.35	0.56	1.10	0.16	0.51***	0.57	-0.36***	0.61
Short Bias	33	11	22	0.43	6.84	0.11	1.68	0.39	0.12	0.02	0.81***	0.66	0.57**	0.67
All funds	2936	1505	1431	1.03	4.53	0.05	3.28	0.55	0.77	0.17	0.44***	0.50	-0.30***	0.54
B. Funds of funds (1,378 funds)														
Market Neutral	41	41	0	0.64	1.51	-0.48	2.28	0.46	0.43	0.29	0.07	0.61	-0.20	0.70
Multi Strategy	832	779	53	0.70	1.70	-0.19	2.19	0.43	0.50	0.29	-0.01	0.49	-0.16*	0.62
No Strategy	336	50	286	0.64	2.86	-0.34	3.81	0.55	0.32	0.11	-0.02	0.59	-0.18***	0.70
Other	33	33	0	0.68	1.78	-0.33	0.54	0.27	0.51	0.29	0.09	0.63	-0.13	0.71
Single Strategy	136	130	6	0.71	1.85	-0.16	2.62	0.45	0.53	0.28	0.25	0.59	-0.06	0.69
All funds	1378	1033	345	0.68	2.00	-0.24	2.59	0.46	0.46	0.23	0.08	0.59	-0.10	0.68
C. Passive strategy indices														
<i>Equity</i>														
Market proxy				0.86	4.66	-0.67	0.43		0.56	0.12				
MSCI World ex U.S.				0.62	4.31	-0.52	0.37		0.33	0.08				
MSCI Emerging Markets				0.81	6.87	-0.91	2.06		0.51	0.07				
Fama & French SMB				0.29	4.35	0.73	5.79		-0.01	0.00				
Fama & French HML				0.46	3.96	0.02	1.84		0.16	0.04				
Momentum				0.86	5.75	-0.63	3.85		0.57	0.10				
<i>Bonds</i>														
Lehman U.S. Aggregate				-0.03	1.09	-0.63	0.65		-0.33	-0.30				
JPM Global Gov. Bond				0.44	1.87	0.36	0.14		0.14	0.07				
JPM Emerging Markets Bond				1.11	4.11	-2.90	18.52		0.82	0.20				
Lehman BAA Corporate				-0.01	1.59	-0.12	0.28		-0.31	-0.20				
<i>Currency</i>														
JPM U.S. Cash 1Month				0.34	0.16	-0.26	-1.54		0.04	0.24				
<i>Commodity</i>														
Goldman Sachs Commodity				0.95	6.17	0.10	-0.27		0.65	0.11				

Other Funds of Funds and 87% for Fixed Income – MBS. For the whole database, the rejection rate is 55% for hedge funds and 46% for funds of funds.

Considering alpha, there is strong evidence of superior hedge fund performance for the market adjusted model, as 15 of 18 strategies achieve significant positive alphas. We find no evidence of superior funds of funds performance. The average R^2 varies between 0.25 for Fixed Income – MBS and 0.68 for Equity Long Only, making the model quite powerful for some hedge fund strategies. R^2 is even higher looking at the market and strategy adjusted model. It varies between 0.29 for Fixed Income – MBS and 0.71 for Other Funds of Funds. Alpha is significantly lower in this model. There are only two hedge fund strategies that achieve significant positive alphas: Fixed Income – MBS and Short Bias.

3.3. *Measurement of performance persistence*

In the performance persistence study, we analyse six time horizons (monthly, bimonthly, quarterly, half-yearly, yearly, and two-yearly horizon), six performance measures (raw returns, Sharpe ratio, two versions of alpha, and the two associated appraisal ratios), and six statistical methodologies (cross-product ratio test (CPR), chi-square test (CS), rank information coefficient (RIC), Spearman rank correlation test (SRC), cross-sectional regression (CSR), and Kolmogorov-Smirnov test (KS)).

The results are presented at different levels of aggregation so as to focus on different aspects of performance persistence. The first focus is on the different methodologies (see Table 4), the second focus is on differences in hedge fund strategies (see Table 5), and the third focus is on the different performance measures used to assess performance persistence (see Table 6).

Table 4 compares the methodologies used in performance persistence analysis. We focus on raw returns as a performance measure and compare the results of the six methodologies for different time horizons (the results for the other performance measures are available upon request). As done in Agarwal and Naik (2000a), this and the following tables show the percentage of cases exhibiting significant performance persistence, which refers to the number of periods (with CPR, CS, RIC, SRC, and KS) or to the number of funds (with CSR) that exhibit persistence.

In Table 4, most of the tests show high levels of persistence for horizons of up to six months. The persistence significance levels slightly weaken as the time horizons are lengthened, which basically confirms the findings presented in Section 2.5. Consider the cross-sectional regression (CSR) with hedge funds as an example. At the monthly horizon, almost one out of three cases (30.73%) exhibits statistically significant performance persistence. However, at the annual and biannual horizons, only 5.77% and 5.65% of all cases show persistence. An exception is the Kolmogorov-Smirnov test with hedge funds where the relatively low level of persistence remains at 8% to 11%. However, with funds of funds, persistence also declines with time and is only 6.67% at the two-year horizon. Comparable results can be found for the Sharpe ratio and the alpha-based measures.

Comparing the results we find that the levels of significance vary widely depending on the methodology. Considering the panel of hedge funds at the monthly horizon, the correlation-based tests (RIC and SRC), the contingency-table-based tests (CPR and CHI), and the regression-based tests (CSR) provide evidence of performance persistence; the portion of significant cases varies between 28.29% (CPR) and 37.35% (RIC). With

Table 4
Comparison of methodologies

This table shows the percentage of cases exhibiting statistically significant performance persistence (at 5% significance level) for the sample of 2,936 hedge funds (Panel A) and 1,378 funds of funds (Panel B). We analyse performance persistence on a monthly, bi-monthly, quarterly, half-yearly, yearly, and bi-yearly basis. The tests employed include cross-product ratio (CPR), chi-square (CHI), the rank information coefficient (RIC), Spearman rank correlation (SRC), cross-sectional regression (CSR), and Kolmogorov-Smirnov (KS).

Time horizon (months)	1	2	3	6	12	24
A. Hedge funds (2,936 funds)						
CPR	28.29	28.81	29.06	31.58	29.63	18.06
CHI	35.90	34.93	34.76	34.80	36.42	30.56
RIC	37.35	40.49	36.75	41.23	41.36	33.33
SRC	36.83	39.83	37.46	45.32	44.44	33.33
CSR	30.73	21.28	10.48	9.05	5.77	5.65
KS	9.09	8.05	7.75	8.64	8.33	11.11
Average	29.70	28.90	26.04	28.44	27.66	22.01
B. Funds of funds (1,378 funds)						
CPR	29.41	30.51	31.28	40.00	33.33	25.00
CHI	40.34	40.34	38.97	42.11	42.22	30.00
RIC	41.68	44.07	39.49	47.37	46.67	40.00
SRC	38.15	38.98	37.44	51.58	46.67	35.00
CSR	35.98	18.17	6.42	5.19	3.57	7.24
KS	21.02	20.00	15.79	20.00	20.00	6.67
Average	34.43	32.01	28.23	34.37	32.08	23.98

the multi-period Kolmogorov-Smirnov, there is hardly any performance persistence; the percentage of cases exhibiting statistically significant performance persistence is only 9.09%. We thus conclude that the use of different methodologies is one of the main drivers for the mixed results found in the literature. We also confirm the findings of Agarwal and Naik (2000a) that the level of persistence observed in a multi-period framework is considerably smaller than that observed in a two-period framework.

Our second step is to analyse differences in hedge fund strategies. Table 5 focuses on raw returns and shows the percentage of cases exhibiting statistically significant performance persistence for 23 strategy groups (the results for the other performance measures are available upon request). In this table the results are aggregated above the different methodologies presented in Table 4.

There are large differences in results depending on the hedge fund strategy. Hedge funds following the Convertible Arbitrage or the Emerging Markets strategy have very high levels of persistence. Other Funds of Funds and Equity Long Only hedge funds have considerable smaller levels of significance. We thus follow Brown and Goetzmann (2003) and Harri and Brorsen (2004) and conclude that persistence is related to the type of hedge fund strategy followed. It is also interesting that Merger Arbitrage and Sector hedge funds retain their high levels of significance across time horizons, whereas with most of the other strategies the significance level decreases as the measurement horizon lengthens.

Table 5
Comparison of hedge fund strategies

This table shows the percentage of cases exhibiting statistically significant performance persistence (at 5% significance level) for the sample of 2,936 hedge funds (Panel A, subdivided into 18 strategies) and 1,378 funds of funds (Panel B, subdivided into 5 strategies). We analyse performance persistence on a monthly, bi-monthly, quarterly, half-yearly, yearly, and bi-yearly basis.

Time horizon (months)	1	2	3	6	12	24
A. Hedge funds (2,936 funds)						
Convertible Arbitrage	53.23	42.05	33.43	43.10	50.59	21.60
Distressed Securities	32.17	28.01	21.97	29.00	12.96	8.33
Emerging Markets	45.05	40.71	42.42	36.82	43.64	44.95
Equity Long Only	13.22	16.70	10.70	16.37	10.07	26.75
Equity Long/Short	33.75	33.82	30.28	28.26	34.29	26.34
Equity Market Neutral	24.45	26.14	21.56	25.11	6.45	0.39
Event Driven Multi Strategy	36.15	28.48	25.38	30.89	14.81	17.53
Fixed Income	20.70	23.79	23.50	22.70	19.06	4.17
Fixed Income—MBS	37.20	41.24	36.31	38.58	26.27	29.72
Fixed Income Arbitrage	38.39	39.30	36.35	31.35	35.11	1.23
Global Macro	19.43	21.96	18.87	20.12	31.81	21.59
Merger Arbitrage	23.94	26.16	24.37	35.82	33.77	40.70
Multi Strategy	19.41	20.79	17.15	10.54	19.42	31.80
No Strategy	17.92	24.39	18.09	20.01	29.06	14.46
Other	25.30	24.95	25.94	15.94	23.06	17.59
Relative Value Multi Strategy	36.58	26.37	26.56	33.98	22.55	17.39
Sector	36.45	40.07	38.82	51.07	59.04	59.06
Short Bias	21.23	15.26	17.11	22.20	25.92	12.50
All funds	29.70	28.90	26.04	28.44	27.66	22.01
B. Funds of funds (1,378 funds)						
Market Neutral	28.19	26.23	27.29	32.94	34.61	6.25
Multi Strategy	45.09	43.12	35.45	52.11	45.35	35.61
No Strategy	46.14	45.28	38.42	46.85	49.18	27.05
Other	17.71	13.94	10.34	10.62	11.57	25.00
Single Strategy	35.02	31.49	29.67	29.34	19.67	26.01
All funds	34.43	32.01	28.23	34.37	32.08	23.98

Finally, we compare the performance measures used to assess performance persistence. The results presented in Table 6 are aggregated for all the methodologies presented in Table 4 and all the hedge fund strategies presented in Table 5.⁹

Comparing the performance measures reveals relatively small differences in the levels of significance. Considering hedge fund returns at the monthly horizon as an example,

⁹ For the persistence measurement we follow Agarwal and Naik (2000a) and define α_m as the fund return minus the market return and α_{ms} as the fund return minus the average return of the funds using the same strategy. The appraisal ratio (AR) is defined as α divided by the residual standard deviation resulting from the regression of the hedge fund return on the market return/the average return of all funds using the same strategy.

Table 6
Comparison of measures

This table shows the percentage of cases exhibiting statistically significant performance persistence (at 5% significance level) for the sample of 2,936 hedge funds (Panel A) and 1,378 funds of funds (Panel B). We analyse performance persistence on a monthly, bi-monthly, quarterly, half-yearly, yearly, and bi-yearly basis using six different performance measures: raw returns, the Sharpe ratio (SR), two versions of alpha, and the two associated versions of appraisal ratio (AR).

Time horizon (months)	1	2	3	6	12	24
A. Hedge funds (2,936 funds)						
Return	29.70	28.90	26.04	28.44	27.66	22.01
Excess return	29.66	28.64	25.87	28.47	28.19	24.54
Sharpe ratio	36.08	36.87	34.03	39.07	38.39	36.59
α_m	25.22	26.31	24.26	28.67	25.45	22.31
AR_m	35.89	36.64	33.01	40.30	39.67	36.98
α_{ms}	28.47	27.99	26.17	29.34	28.05	20.99
AR_{ms}	36.92	36.97	34.08	40.60	37.85	39.22
Average	31.71	31.76	29.07	33.55	32.18	28.95
B. Funds of funds (1,378 funds)						
Return	34.43	32.01	28.23	34.37	32.08	23.98
Excess return	34.67	31.35	27.85	32.73	31.38	26.73
Sharpe ratio	39.95	37.90	38.02	43.44	45.78	38.71
α_m	28.74	28.91	26.72	37.38	25.88	26.94
AR_m	37.73	38.30	36.74	45.21	46.29	38.71
α_{ms}	35.42	30.92	28.50	34.20	33.60	25.33
AR_{ms}	38.13	37.79	34.75	43.59	41.24	42.32
Average	35.58	33.88	31.55	38.70	36.61	31.82

significance levels vary between 25.22% with the market adjusted alpha and 36.92% with the market and strategy adjusted appraisal ratio. Again, the persistence significance levels weaken as the measurement horizon lengthens. The only exception is that the level of significance remains very stable with the appraisal ratios. The finding that the level of persistence is a bit higher with appraisal ratios compared to alphas is consistent with the findings of Agarwal and Naik (2000a). However, overall it seems that the level of hedge fund performance persistence is not related to the choice of performance measure.

3.4. Reasons for performance persistence

In this section we further explore the economic insights that can be derived from the interplay between the differences in statistical methodologies and the results found in the literature. We take a look at three issues that are discussed in literature as reasons for persistence: (1) the use of option-like strategies, (2) return smoothing, and (3) survivorship and backfilling bias. We show that the second and third reasons at least partially explain the high degree of persistence found in the last section of this paper, whereas the first one does not.

Measurement of performance persistence for a short out-of-the-money strategy. Some authors argue that persistence of returns signals a strategy bearing tail risk, for example,

returns from writing an out-of-the-money (OTM) put option will be positive until the occurrence of a tail event brings losses when the option pays out (see Agarwal and Naik, 2004; Mitchell and Pulvino, 2001; Fung and Hsieh, 2001; Chan *et al.*, 2007). How will the different statistical methodologies perform with data generated from such a strategy?

We use two different simulation approaches to generate data for this strategy. In the first approach, we calculate all performance persistence measures for 250 funds that write OTM put options with a maturity of one month on different US stocks. We approximate the empirical return distribution of this strategy using Black and Scholes (1973) option pricing and simulation. For each of the 250 funds, we simulate 120 returns of the underlying stocks using a normal distribution with a mean (μ) of 0.86% per month and a standard deviation (σ) of 4.66% per month (these are the values of the equity market proxy; see Table 3). We assume that the option will get into the money when the return of the stock s in time period t (r_{st}) is one standard deviation below the mean ($\mu - \sigma$, i.e., $r_{st} < -3.80\%$; the exercise price X is thus $S_0 \cdot (1 + (\mu - \sigma))$, with S_0 as current stock price of stock s). As this option is far out of the money, the option premium is relatively small; we use the Black and Scholes (1973) model for puts to derive the option price and as a result assume that the premium P_t is lognormally distributed with a mean of 0.46% and a standard deviation of 0.50%.¹⁰ The return of the option writing fund i in time period t is then given as $r_{it} = \mu - \sigma - r_{st} + P_t$, if $r_{st} < \mu - \sigma$, and P_t otherwise.¹¹ Although the resulting simulated return distribution is only an approximation of the real returns that an OTM put option writing strategy would produce, it leads to 250 return distributions that exhibit negative skewness and positive excess kurtosis and thus fit the desired tail risk of the option strategy. To illustrate our simulation approach, a plot of a simulated fund is presented in the left part of Figure 1, together with a plot of one of the hedge funds that we analyse in Section 3.3 (right part of Figure 1).

When comparing the two parts of this figure, one can see that the simulated fund fits the tail risk of the hedge fund quite well. Many of the 2,936 hedge funds and 1,378 funds of hedge funds in our database show negative skewness and positive kurtosis similar to that depicted in Figure 1 (see Table 3), which confirms the empirical finding that many hedge funds exhibit significant tail risk (see, e.g., Ding and Shawky, 2007; Liang and Park, 2007).

¹⁰ We use the input parameters $\sigma_{\text{ann}} = \sigma_{\text{monthly}} \cdot \sqrt{12} = 4.86\% \cdot \sqrt{12} = 16.14\%$, $r_{\text{fann}} = r_{\text{fmonthly}} \cdot 12 = 0.299\% \cdot 12 = 3.58\%$ (annualised one-month Treasury bill rate provided by Ibbotson Associates; see Table 3), $\tau = 1/12$ (maturity one month = 1/12 years) and the Black and Scholes (1973) option pricing formula for puts: $\text{Price} = Xe^{-r_{\text{fann}}\tau}(1 - N(d_2)) - S_0(1 - N(d_1))$, with $d_1 = (\ln(S_0/X) + (r_{\text{fann}} + \sigma_{\text{ann}}^2/2)\tau)/\sigma_{\text{ann}}\sqrt{\tau}$, and $d_2 = d_1 - \sigma_{\text{ann}}\sqrt{\tau}$. A current stock price of $S_0 = 100$ would then result in $X = S_0 \cdot (1 + (\mu - \sigma)) = 96.20$ and an option price of 0.46. If Black and Scholes (1973) prices represent the fair price of the option, then the option premium should be 0.46% of the current stock price. Given that empirical observed option prices deviate from the values derived by the Black and Scholes (1973) model (see, e.g., Macbeth and Merville, 1980), we assume that the option prices vary around this estimator of 0.46% with a standard deviation of 0.50%. As the premium cannot be negative, we assume a lognormal distribution of the premium.

¹¹ To obtain this return formula, we assume that the capital basis is the stock price S_0 . Assuming a lower capital basis would leverage up the risk and return of this option strategy. However, leverage is delimited by the margin requirement that depends on Regulation T and the broker's individual policies.

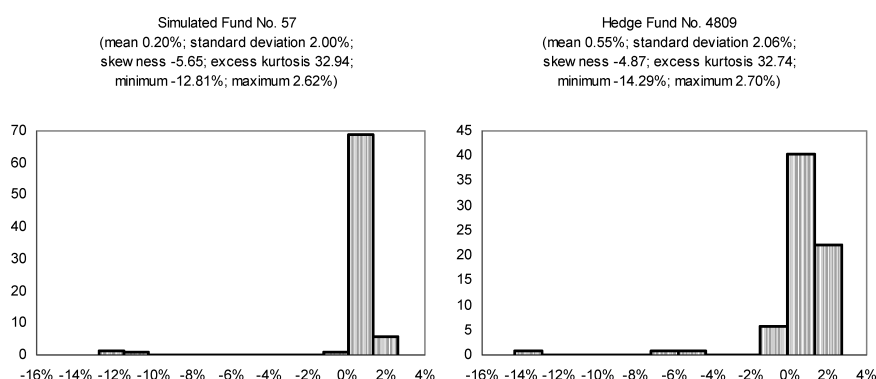


Fig. 1. Return distribution for simulated fund no. 57 and hedge fund no. 4809

There is one empirical estimator of an OTM put option strategy available in literature – the S&P 500 OTM put factor developed by Agarwal and Naik (2004). We use this factor as the basis for our second simulation approach. We first estimate the distribution of shorting the OTM put factor using the distribution fitting software BestFit and find a generalised beta distribution with the continuous shape parameters $\alpha_1 = 2.22$ and $\alpha_2 = 2.82$, and a minimum (maximum) of -1.75 (3.82). Then, we simulate 120 returns for 250 funds under this distributional assumption. Finally, we use these simulated returns as the basis for the performance persistence analysis.¹²

Panels A and B of Table 7 set forth the results of the performance persistence measurement for both simulation approaches using returns as performance measures.

For the simulated funds that follow the short OTM strategy, very little persistence is found. Correlation-based tests (RIC and SRC), contingency-table-based tests (CPR and CHI), and regression-based tests (CSR) show only very low levels of persistence; the multi-period Kolmogorov-Smirnov test shows no persistence at all. It thus seems that the level of persistence as found in the last section cannot be explained by the use of option-like strategies.¹³

To discover whether the option strategy makes any difference at all, we added a performance persistence analysis for a random walk without drift (Panel C of Table 7, using the standard deviation of the equity market proxy; see Table 3) and a random walk with drift (Panel D of Table 7, using the mean of the equity market proxy as drift).

¹² We are grateful to Vikas Agarwal and Narayan Y. Naik for providing the data for this analysis. We use the negative natural logarithm of the returns (from January 1996 to December 2005) of the OTM put factor to obtain the short OTM strategy, which Agarwal and Naik (2004) observe for many hedge fund strategies. Indeed, most of the simulated returns are positive, apart from the tail events, when the option pays out.

¹³ We also simulated the returns under different correlation assumptions (across funds as well as across time) and found very little persistence for low levels of correlation (up to 0.2). However, as shown in Brown *et al.* (1992) and Carpenter and Lynch (1999), increasing the level of correlations to higher values also increases performance persistence. The cross-sectional correlation among the hedge funds in this sample is about 0.16 (also see Eling, 2008); cross-sectional correlation should thus not be a main driver of persistence. However, as shown in Table 8, serial correlation can be high for some hedge fund strategies. Serial correlation is discussed in more detail in the next subsection.

Table 7
Performance persistence for simulated returns of 250 funds

This table shows the percentage of cases exhibiting statistically significant performance persistence (at 5% significance level) for 250 simulated funds. We analyse performance persistence on a monthly, bi-monthly, quarterly, half-yearly, yearly, and bi-yearly basis. The tests employed include cross-product ratio (CPR), chi-square (CHI), the rank information coefficient (RIC), Spearman rank correlation (SRC), cross-sectional regression (CSR), and Kolmogorov-Smirnov (KS).

Measure	Returns					
	1	2	3	6	12	24
Panel A: Short Out-of-the-money Strategy (own simulation)						
CPR	0.84	0.00	2.56	5.26	0.00	0.00
CHI	2.52	1.69	10.26	5.26	0.00	0.00
RIC	5.04	5.08	0.00	5.26	0.00	0.00
SRC	4.20	1.69	2.56	0.00	0.00	0.00
CSR	2.40	4.00	1.20	2.00	1.60	2.00
KS	0.00	0.00	0.00	0.00	0.00	0.00
Average	2.50	2.08	2.76	2.96	0.27	0.33
Panel B: Short Out-of-the-money Strategy (simulation using Agarwal and Naik's (2004) OTM put factor)						
CPR	4.20	1.69	0.00	0.00	0.00	0.00
CHI	5.88	8.47	5.13	5.26	0.00	0.00
RIC	2.52	3.39	5.13	5.26	0.00	0.00
SRC	2.52	1.69	2.56	0.00	0.00	0.00
CSR	4.80	4.80	2.00	1.60	4.00	11.24
KS	0.00	0.00	0.00	0.00	0.00	0.00
Average	3.32	3.34	2.47	2.02	0.67	1.87
Panel C: Random walk without drift (standard deviation 4.86%)						
CPR	0.84	3.39	0.00	5.26	0.00	0.00
CHI	3.36	3.39	2.56	10.53	0.00	0.00
RIC	2.52	1.69	2.56	5.26	0.00	0.00
SRC	2.52	3.39	2.56	5.26	0.00	0.00
CSR	2.40	3.20	3.20	4.00	4.40	8.40
KS	0.00	0.00	0.00	0.00	0.00	0.00
Average	1.94	2.51	1.82	5.05	0.73	1.40
Panel D: Random walk with drift (standard deviation 4.86%, drift 0.86%)						
CPR	0.00	3.39	5.13	0.00	0.00	0.00
CHI	2.52	3.39	5.13	5.26	0.00	0.00
RIC	1.68	8.47	2.56	0.00	0.00	0.00
SRC	0.84	6.78	2.56	0.00	0.00	0.00
CSR	2.40	0.80	1.20	0.80	1.20	2.80
KS	0.00	0.00	0.00	0.00	0.00	0.00
Average	1.24	3.81	2.76	1.01	0.20	0.47

The results are very similar to those discussed above – very little persistence in any simulation. Obviously, other explanations of persistence need to be found.

Transparency of valuations and smoothing of returns. As mentioned in Section 2.5, the smoothing of returns might explain short-term performance persistence. In particular, hedge fund managers engaged in nonmarketable securities and illiquid assets exercise wide discretion in pricing, which results in smoothing of returns (see Asness *et al.*, 2001). As a consequence, the level of performance persistence might be very high for funds acting in illiquid markets (e.g., convertibles or mortgage-backed securities), whereas funds engaged in markets with transparent prices (e.g., equity markets) show less persistence. Artificial smoothing of returns becomes visible in hedge fund data in the form of serial correlation (see Getmansky *et al.*, 2004). Table 8 shows serial correlation for lags 1, 2, and 3, as well as performance persistence for different hedge fund strategies.

Strong serial correlation is observable for the sample of hedge funds as well as for the sample of funds of funds. The highest serial correlation can be found with Convertible Arbitrage (serial correlation 0.34 with lag 1), Relative Value Multi Strategy (0.27), and Fixed Income MBS (0.26). All these strategies act in relative illiquid markets (e.g., none of the three are primarily active in the relatively liquid equity markets). The lowest serial correlations are found with Global Macro (0.03), Equity Long Only (0.04), and Short Bias (0.05) – three strategies that are mainly active in equity and other liquid markets. Convertible Arbitrage, Relative Value Multi Strategy, and Fixed Income MBS are among the strategies with the highest persistence, whereas Global Macro, Equity Long Only, and Short Bias show relatively low levels of persistence. The high levels of persistence found for the first group of funds might thus be due to the fact that these hedge funds smooth their returns. We thus confirm the finding that smoothing of returns might explain the high levels of short-term persistence found with some strategies (see, e.g., Kosowski *et al.*, 2007).

Survivorship bias and backfilling bias. Brown *et al.* (1992) show that survivorship gives rise to biases in the first and second moments and cross-moments of returns and performance persistence when there is dispersion of risk among the population of managers (see also Getmansky *et al.*, 2004). To see how survivorship influences our results, we dropped the dead funds from the database and repeated the analysis for the living funds. Part 1 of Table 9 shows the results for returns as a measure of persistence.

There are differences in the levels of returns and standard deviations between the total sample and the sample of living funds, confirming the findings of Brown *et al.* (1992). Living funds tend to have higher returns and lower standard deviations. The level of performance persistence is a bit lower when only surviving funds are analysed (on average 4.17% (5.58%) for hedge funds (funds of hedge funds)); we also analysed dead funds and found even lower levels of persistence for these funds; the results are available upon request). We thus confirm the finding of Malkiel and Saha (2005) that survivorship bias influences the level of performance persistence with hedge funds.¹⁴ However, we

¹⁴ Note that our result is different from Malkiel and Saha (2005) in that these authors find slightly more persistence when dead funds are dropped from the database. However, the downward bias that we found is consistent with findings in the mutual funds literature

Table 8
Serial correlation and performance persistence for different strategies

This table shows serial correlation for lags 1, 2, and 3 as well as the percentage of cases exhibiting statistically significant performance persistence (at 5% significance level) for the sample of 2,936 hedge funds (Panel A, subdivided into 18 strategies) and 1,378 funds of funds (Panel B, subdivided into 5 strategies). Performance persistence is analysed at a monthly, bi-monthly, quarterly, half-yearly, yearly, and bi-yearly basis.

	No. of Funds	Serial correlation			Performance Persistence						
		Lag 1	Lag 2	Lag 3	1	2	3	6	12	24	Average
A. Hedge funds (2,936 funds)											
Convertible Arbitrage	159	0.34	0.14	0.10	43.63	37.42	46.24	50.71	23.40	43.63	37.42
Distressed Securities	113	0.22	0.06	0.01	33.03	28.17	34.60	17.13	15.74	33.03	28.17
Emerging Markets	257	0.16	0.04	0.00	43.32	45.71	39.76	50.28	38.63	43.32	45.71
Equity Long Only	57	0.04	0.00	−0.01	20.50	15.96	18.13	12.51	25.88	20.50	15.96
Equity Long/Short	1109	0.10	0.03	0.00	36.55	34.11	31.25	36.67	27.49	36.55	34.11
Equity Market Neutral	164	0.08	0.02	0.01	30.43	26.64	28.16	9.91	3.88	30.43	26.64
Event Driven Multi Strategy	120	0.21	0.05	0.02	34.31	34.13	36.73	20.21	22.99	34.31	34.13
Fixed Income	65	0.17	0.02	0.01	29.43	30.68	27.57	23.36	7.29	29.43	30.68
Fixed Income—MBS	54	0.26	0.14	0.09	44.95	40.94	41.05	32.73	34.72	44.95	40.94
Fixed Income Arbitrage	111	0.22	0.10	0.07	41.87	39.38	34.65	37.33	3.09	41.87	39.38
Global Macro	142	0.03	0.03	0.05	26.45	23.91	23.88	35.77	24.24	26.45	23.91

Merger Arbitrage	104	0.14	0.01	-0.01	29.85	28.22	35.50	35.65	40.70	29.85	28.22
Multi Strategy	47	0.17	0.05	0.00	26.82	22.47	16.93	23.02	36.18	26.82	22.47
No Strategy	47	0.07	-0.02	-0.02	31.48	25.89	23.91	31.62	13.48	31.48	25.89
Other	36	0.23	0.01	0.00	25.91	30.10	18.71	25.75	21.30	25.91	30.10
Relative Value Multi Strategy	61	0.27	0.00	0.01	31.29	35.03	42.72	28.43	18.84	31.29	35.03
Sector	257	0.11	0.08	0.11	43.37	42.06	52.37	60.87	61.41	43.37	42.06
Short Bias	33	0.05	0.05	0.05	17.28	19.13	21.69	25.38	12.50	17.28	19.13
B. Funds of funds (1,378 funds)											
Market Neutral	41	0.25	0.15	0.15	32.25	33.14	36.64	42.29	39.29	11.46	32.25
Multi Strategy	832	0.21	0.11	0.03	51.31	52.72	46.92	62.68	53.76	39.14	51.31
No Strategy	336	0.19	0.04	0.00	47.32	47.75	42.65	49.96	52.21	27.97	47.32
Other	33	0.14	-0.08	-0.06	23.27	20.00	19.93	19.21	19.07	28.33	23.27
Single Strategy	136	0.20	0.04	0.01	39.67	38.77	37.14	34.74	25.28	29.55	39.67

Table 9
Comparison of hedge fund strategies – survivorship bias and backfilling bias

This table shows the percentage of cases exhibiting statistically significant performance persistence (at 5% significance level) for the sample of 2,936 hedge funds (Panel A, subdivided into 18 strategies) and 1,378 funds of funds (Panel B, subdivided into 5 strategies). We analyse performance persistence on a monthly, bi-monthly, quarterly, half-yearly, yearly, and bi-yearly basis. Note that for Funds of funds-Market neutral and Funds of funds-Other there are no dead funds in the database (see Table 3). The results for all funds and the living funds are thus equal for these two strategies. ‘w/o 24’ stands for ‘without first 24 months of return.’

Time horizon (months)	1. Survivorship Bias					2. Backfilling Bias										
	Mean (%)	St. D. (%)	1	2	3	6	12	24	1	2	3	6	12	24		
A. Hedge funds (2,936 funds)																
Convertible Arbitrage	All	0.82	2.13	53.23	42.05	33.43	43.10	50.59	21.60	All	53.23	42.05	33.43	43.10	50.59	21.60
	Living	0.75	1.89	49.03	37.40	28.34	41.70	36.01	17.52	w/o 24	35.86	31.15	23.08	31.58	26.67	15.00
Distressed Securities	All	1.14	3.64	32.17	28.01	21.97	29.00	12.96	8.33	All	32.17	28.01	21.97	29.00	12.96	8.33
	Living	1.27	1.89	31.34	26.07	22.63	32.46	12.96	8.33	w/o 24	16.08	16.52	15.38	14.91	20.00	12.50
Emerging Markets	All	1.21	7.05	45.05	40.71	42.42	36.82	43.64	44.95	All	45.05	40.71	42.42	36.82	43.64	44.95
	Living	1.79	5.70	39.10	32.28	28.97	26.02	30.02	43.82	w/o 24	31.10	29.08	25.13	23.77	35.11	50.00
Equity Long Only	All	1.19	6.43	13.22	16.70	10.70	16.37	10.07	26.75	All	13.22	16.70	10.70	16.37	10.07	26.75
	Living	1.31	4.50	11.57	17.68	9.83	14.17	12.96	18.33	w/o 24	11.13	11.86	12.82	14.47	5.56	25.00
Equity Long/Short	All	1.14	5.07	33.75	33.82	30.28	28.26	34.29	26.34	All	33.75	33.82	30.28	28.26	34.29	26.34
	Living	1.22	4.07	25.79	23.67	19.90	17.44	41.20	25.00	w/o 24	28.24	30.08	22.72	24.41	19.44	33.33
Equity Market Neutral	All	0.61	2.74	24.45	26.14	21.56	25.11	6.45	0.39	All	24.45	26.14	21.56	25.11	6.45	0.39
	Living	0.59	2.26	19.11	23.12	15.94	15.67	10.16	0.00	w/o 24	19.20	15.69	14.36	16.84	24.44	10.00
Event D. M. Strategy	All	1.04	3.14	36.15	28.48	25.38	30.89	14.81	17.53	All	36.15	28.48	25.38	30.89	14.81	17.53
	Living	1.03	2.97	35.50	30.08	30.77	36.84	16.67	25.00	w/o 24	41.49	36.64	35.03	41.85	18.35	15.17

Fixed Income	All	0.69	1.71	20.70	23.79	23.50	22.70	19.06	4.17	All	20.70	23.79	23.50	22.70	19.06	4.17
	Living	0.75	1.53	21.85	23.66	19.06	16.44	25.93	20.83	w/o 24	10.59	16.71	19.64	7.37	15.56	0.00
Fixed Income—MBS	All	0.91	2.10	37.20	41.24	36.31	38.58	26.27	29.72	All	37.20	41.24	36.31	38.58	26.27	29.72
	Living	0.89	1.68	38.76	33.01	32.22	31.55	17.24	21.81	w/o 24	25.35	29.59	28.36	31.95	24.44	5.00
Fixed Income Arbitrage	All	0.51	2.13	38.39	39.30	36.35	31.35	35.11	1.23	All	38.39	39.30	36.35	31.35	35.11	1.23
	Living	0.71	1.41	23.67	26.77	19.45	21.20	34.34	1.11	w/o 24	23.20	22.13	24.92	16.84	22.22	10.00
Global Macro	All	0.86	4.80	19.43	21.96	18.87	20.12	31.81	21.59	All	19.43	21.96	18.87	20.12	31.81	21.59
	Living	1.19	4.31	17.68	16.83	18.20	12.53	17.24	12.50	w/o 24	16.73	14.33	16.41	13.68	35.56	20.00
Merger Arbitrage	All	0.70	2.00	23.94	26.16	24.37	35.82	33.77	40.70	All	23.94	26.16	24.37	35.82	33.77	40.70
	Living	0.65	1.34	20.28	21.74	18.64	21.65	18.06	17.42	w/o 24	19.97	18.06	19.64	26.69	22.22	30.00
Multi Strategy	All	1.00	2.89	19.41	20.79	17.15	10.54	19.42	31.80	All	19.41	20.79	17.15	10.54	19.42	31.80
	Living	1.05	2.99	19.82	19.46	17.55	12.95	20.85	15.28	w/o 24	12.82	12.88	11.79	11.58	8.89	45.00
No Strategy	All	1.23	3.69	17.92	24.39	18.09	20.01	29.06	14.46	All	17.92	24.39	18.09	20.01	29.06	14.46
	Living	0.55	1.08	14.42	11.30	9.69	9.01	9.49	18.06	w/o 24	13.03	13.22	9.23	10.53	20.00	15.00
Other	All	1.01	2.48	25.30	24.95	25.94	15.94	23.06	17.59	All	25.30	24.95	25.94	15.94	23.06	17.59
	Living	0.88	2.23	22.83	21.21	21.52	10.28	13.11	17.86	w/o 24	17.90	13.22	18.46	12.63	8.89	15.00
Relative V. M. Strategy	All	0.90	1.71	36.58	26.37	26.56	33.98	22.55	17.39	All	36.58	26.37	26.56	33.98	22.55	17.39
	Living	0.87	1.59	36.28	25.82	27.52	36.44	17.07	4.92	w/o 24	21.18	16.95	20.15	16.84	24.44	20.00
Sector	All	1.38	7.06	36.45	40.07	38.82	51.07	59.04	59.06	All	36.45	40.07	38.82	51.07	59.04	59.06
	Living	1.35	5.25	35.50	35.59	37.82	47.37	55.56	50.00	w/o 24	31.51	37.71	37.18	42.11	61.11	50.00
Short Bias	All	0.43	6.84	21.23	15.26	17.11	22.20	25.92	12.50	All	21.23	15.26	17.11	22.20	25.92	12.50
	Living	0.31	7.62	16.53	14.13	17.12	16.19	22.22	16.67	w/o 24	18.91	17.14	15.54	9.47	27.78	0.00
All funds	All	0.88	3.56	29.70	28.90	26.04	28.44	27.66	22.01	All	29.70	28.90	26.04	28.44	27.66	22.01
	Living	0.90	2.86	26.61	24.43	21.95	23.33	22.84	18.58	w/o 24	21.91	21.28	20.55	20.42	23.37	20.61

Table 9
Continued.

1. Survivorship Bias																	2. Backfilling Bias							
		Mean	St. D.																					
Time horizon (months)				(%)	(%)	1	2	3	6	12	24	24	1	2	3	6	12	24						
B. Funds of funds (1,378 funds)																								
Market Neutral	All	0.64	1.51	28.19	26.23	27.29	32.94	34.61	34.61	6.25	All	28.19	26.23	27.29	32.94	34.61	34.61	6.25						
	Living	0.64	1.51	28.19	26.23	27.29	32.94	34.61	34.61	6.25	w/o 24	16.22	11.19	22.46	31.28	27.78	27.78	5.00						
Multi Strategy	All	0.71	1.66	45.09	43.12	35.45	52.11	45.35	45.35	35.61	All	45.09	43.12	35.45	52.11	45.35	45.35	35.61						
	Living	0.71	1.62	43.91	41.32	34.30	49.38	36.05	36.05	31.56	w/o 24	31.81	37.81	28.21	40.79	51.56	35.00	35.00						
No Strategy	All	0.61	2.81	46.14	45.28	38.42	46.85	49.18	49.18	27.05	All	46.14	45.28	38.42	46.85	49.18	49.18	27.05						
	Living	0.76	1.64	24.94	21.25	21.86	15.79	14.81	14.81	12.50	w/o 24	30.85	30.78	28.72	36.58	40.89	20.00	20.00						
Other	All	0.68	1.78	17.71	13.94	10.34	10.62	11.57	11.57	25.00	All	17.71	13.94	10.34	10.62	11.57	25.00	25.00						
	Living	0.68	1.78	17.71	13.94	10.34	10.62	11.57	11.57	25.00	w/o 24	13.48	8.70	12.00	19.02	6.67	10.00	10.00						
Single Strategy	All	0.71	1.85	35.02	31.49	29.67	29.34	19.67	19.67	26.01	All	35.02	31.49	29.67	29.34	19.67	26.01	26.01						
	Living	0.72	1.85	33.51	30.90	27.58	29.27	19.75	19.75	25.56	w/o 24	27.58	26.15	25.44	28.80	22.22	20.00	20.00						
All funds	All	0.56	1.60	34.43	32.01	28.23	34.37	32.08	32.08	23.98	All	34.43	32.01	28.23	34.37	32.08	23.98	23.98						
	Living	0.59	1.40	29.68	26.78	24.36	27.43	24.84	24.84	18.51	w/o 24	23.99	22.93	23.36	31.29	29.82	18.00	18.00						

cannot identify a systematic connection between the extent of survivorship bias (both for risk and return) and the level of persistence for different hedge fund strategies, i.e., strategies with high survivorship bias do not systematically show higher or lower levels of persistence compared to other strategies. We thus conclude that survivorship bias might influence the level of persistence, but it cannot explain differences among hedge fund strategies.

As shown in Section 3.1, backfilling of historical returns causes an upward bias in performance measurement results. To analyse whether backfilling bias also influences levels of persistence, we repeated the analysis after dropping the first 24 months of returns for each fund from the database (see Part 2 of Table 9). Again we found lower levels of persistence (on average 5.77% (5.95%) for hedge funds (funds of hedge funds)), which seems reasonable given that usually successful funds with good track records are backfilled. Removing these winners from the database thus eliminated this bias and, in this new environment, we find that persistence decreases. It thus seems that survivorship bias as well as backfilling bias can partly explain the persistence found in Section 3.3.

4. A Rationale for Choosing between Methodologies

We have concluded that the use of different methodologies is one driver for the mixed results reported in the literature overview and in the new empirical evidence. Considering the wide range of methodologies available to measure performance persistence, we now develop a rationale for choosing the most appropriate methodology to measure hedge fund performance persistence. It is important to recognise that all theoretical models – including the six methodologies analysed here – build upon certain assumptions that might be more or less appropriate to different investigations. The suitability of the different methods must thus be evaluated in the context of the data or the conjecture being explored. For example, it is well known that the returns of hedge funds are not normally distributed (see Section 3.2) and that hedge fund data suffers from survivorship bias (see Section 3.1); these characteristics will now be important when comparing the suitability of different methodologies. Table 10 sets out a comparison of the methodologies and provides the foundation for the following discussion.

The fourth column in Table 10 shows that the cross-product ratio test and the regression are most commonly employed methodologies in the literature. However, given the above discussions, three things need to be kept in mind when selecting a methodology: (1) the statistical characteristics of the methods, (2) the statistical characteristics of the hedge fund returns, and (3) the robustness of results.

Statistical characteristics of the methods. Contingency-tables-based tests such as the CPR provide an easy way to understand and communicate persistence. However, one problem with this sort of test is that they incorporate large differences in the evaluation for nearly identical funds at the thresholds, e.g., comparing the worst fund of the upper half with the best fund of the lower half (see Blake and Timmermann, 2003). It would thus be useful to supplement contingency-tables-based tests with a statistical analysis that does not depend on forming quartiles of funds. This critique is especially relevant for the CPR and CHI tests. The multi-period KS test also relies on thresholds, but not

(Hendricks *et al.*, 1993; Carhart *et al.*, 2002) that persistence is stronger in full samples than in survivor-only samples.

Table 10
Methodologies for testing performance persistence

This table shows the statistical methodologies used within 25 studies on hedge fund performance persistence. The first column subdivides the methodologies into two-period and multi-period measures. Within the two-period framework we can further distinguish between nonparametric and parametric approaches (second column). The third column displays six tests for performance persistence. The fourth column gives the authors of the studies in which the methodology was applied. The last column contains an assessment of the methodology including implicit assumptions, data requirements, and implementation.

Period	Methodological Basis	Test (Statistic)	Used in	Assessment
Two Period	contingency table-based (nonparametric)	cross-product ratio test (Z-statistic)	1. Agarwal and Naik (2000a)	+ Data requirement low
			2. Agarwal and Naik (2000b)	+ Simple to calculate (implemented in standard software such as Excel or SPSS)
			3. Brown <i>et al.</i> (1999)	+ Easy to communicate
			4. DeSouza and Gokcan (2004)	– Large differences in the evaluation of nearly identical funds at the thresholds (see Blake and Timmermann, 2003)
			5. Edwards and Caglayan (2001)	
			6. Henn and Meier (2004)	
			7. Kat and Menexe (2003)	
			8. Koh <i>et al.</i> (2003)	
			1. Agarwal <i>et al.</i> (2007)	+ Data requirement low
			2. Agarwal and Naik (2000a)	+ More robust in the presence of survivorship bias (see Carpenter and Lynch, 1999)
			3. Koh <i>et al.</i> (2003)	– Large differences in the evaluation of nearly identical funds at the threshold (see Blake and Timmermann, 2003)
			4. Kouwenberg (2003)	– More complicated to calculate and to communicate than the CPR test
			5. Malkiel and Saha (2005)	
			6. Park and Staum (1998)	
		chi-square test (X^2 -statistic)		

correlation-based (nonparametric)	rank information coefficient (Fisher T-statistic)	1. Herzberg and Mozes (2003)	<ul style="list-style-type: none"> + Data requirement low + Simple to calculate (implemented in standard software such as Excel or SPSS) + Easy to communicate - Displays serial correlation
	Spearman rank correlation test (Fisher T-statistic)	1. Harri and Brorsen (2004) 2. Park and Staum (1998)	<ul style="list-style-type: none"> + Data requirement low + No requirement that the variables must be measured on interval scales + No assumption that the relationship between the variables is linear + Simple to calculate (implemented in standard software such as Excel or SPSS) - Displays serial correlation
regression-based (parametric)	cross-sectional regression (T-statistic)	1. Agarwal <i>et al.</i> (2007) 2. Agarwal and Naik (2000a) 3. Agarwal and Naik (2000b) 4. Amenc <i>et al.</i> (2003) 5. Barès <i>et al.</i> (2003) 6. Boyson and Cooper (2004) 7. Brown <i>et al.</i> (1999) 8. Brown and Goetzmann (2003)	<ul style="list-style-type: none"> + Simple to calculate (implemented in standard software such as Excel or SPSS) - Assumption of normally distributed residuals is critical - Assumption of uncorrelated residuals is critical, but modified test statistics for autocorrelated data are available (see Newey and West, 1987)

Table 10
Continued.

Period	Methodological Basis	Test (Statistic)	Used in	Assessment
Multi Period	(nonparametric)	Kolmogorov-Smirnov test (KS-statistic)	9. Capocci <i>et al.</i> (2005) 10. Capocci and Hübner (2004) 11. Chen and Passow (2003) 12. De Souza and Gokcan (2004) 13. Edwards and Caglayan (2001) 14. Harri and Brorsen (2004) 15. Jagannathan <i>et al.</i> (2006) 16. Kat and Menexe (2003) 17. Kosowski <i>et al.</i> (2007)	– Too few data, especially at the annual and the biannual horizon, to run a sound regression analysis
			1. Agarwal and Naik (2000a) 2. Koh <i>et al.</i> (2003)	+ Data requirement low + Very robust compared to other methodologies (see Agarwal and Naik, 2000a) + More efficient than chi-square test (see Géhin, 2004) + Differences in the evaluation of nearly identical funds at the threshold (see Blake and Timmermann, 2003) is less severe problem compared to the CPR and chi-square test (evaluation based on more than one threshold) – Test is designed to use with independent data, but modified test statistics for autocorrelated data are available (see Weiss, 1978)

just on one, but on many more depending on the number of periods considered. Thus, under this test, funds are not discriminated based on only a single threshold. Therefore, the KS test is most preferable of these three tests considering this point of critique.

In the presence of very few data points, nonparametric test might be more appropriate than parametric-regression-based tests (see Lee *et al.*, 2000, p. 759; McNabb, 2004, p. 267). This data requirement is especially relevant when analysing hedge fund returns where only monthly returns are available. For example, in this investigation, 120 monthly returns are examined, which are then reduced even further to 10 yearly and 5 biyearly data points; the data availability regarding hedge funds is thus a clear argument against using parametric-regression-based tests, a problem that becomes even more severe if the returns do not follow a normal distribution (see Tanizaki, 2004, p. 446).

Statistical characteristics of the hedge fund returns. As shown by the results of the Jarque-Bera test (see Table 3), the returns of most hedge funds and funds of funds are not normally distributed. Nonparametric tests avoid imposing assumptions of normality on hedge fund returns and thus might be more appropriate for evaluating the persistence of hedge funds. The parametric cross-sectional regression builds on the classical axioms of linear regression analysis, such as uncorrelated and normally distributed residuals (see, e.g., McClave *et al.*, 2001, p. 473). Therefore, this methodology might be more suitable for evaluating the performance persistence of mutual funds, where the returns are more likely to be normally distributed than are those of hedge funds (see Eling, 2008).

Furthermore, smoothing of hedge fund returns leads to high levels of serial correlation and increases the levels of persistence. In such a situation, the use of the correlation-based measures becomes questionable because these will not reflect manager skill but, instead, the higher serial correlation of those funds that smooth their returns. This is also a problem for the regression-based tests and the KS test, which are both designed for use with independent data; however, for both, modified test statistics for autocorrelated data are available (see Newey and West, 1987, for the regression test, and Weiss, 1978, for the KS test).

Robustness of results. Carpenter and Lynch (1999) study the specification and power of various persistence tests, finding that the CHI test is well specified, powerful, and more robust to the presence of survivorship bias compared to other methodologies. As shown in Section 3.4, survivorship bias is relevant when analysing hedge fund performance persistence – a clear argument in favour of the chi-square test.

The multi-period framework is better able to discriminate between persistence due to chance and persistence due to manager skill than is the traditional two-period framework (see Agarwal and Naik, 2000a). The multi-period KS test might thus produce even more robust results than the CHI test. Furthermore, this test is generally more efficient than the CHI test for small samples (see Géhin, 2004).

In summary, we find that there are several problems with using parametric-regression-based tests, especially due to the nature of hedge fund data. It thus seems that nonparametric tests are more reliable for analysing hedge fund performance persistence. Among the nonparametric tests, the multi-period KS test as employed by Agarwal and Naik (2000a) has some advantages compared to other methodologies, among them that it is more robust and efficient.

We thus conclude that the Kolmogorov-Smirnov test is the most useful method of evaluating persistence of hedge funds. However, given the variability of results presented in the empirical section of this paper, we also recommend considering more than one methodology when analysing persistence. We thus suggest that the results of the

multi-period Kolmogorov-Smirnov test might be supplemented with the two-period chi-square test, as this test is very robust in the presence of survivorship bias (see Carpenter and Lynch, 1999).

5. Conclusion

The contribution of this paper was to review recent studies on hedge fund performance persistence and to provide new empirical evidence on this widely discussed and controversial topic. The extant literature varies widely in the results, which is possibly due to the use of different databases, performance measures, and statistical methodologies. Most authors find short-term persistence for horizons of up to six months, but the return persistence significance levels weaken as the measurement horizon lengthens.

To obtain a more accurate picture of performance persistence, we investigated 4,314 hedge funds from the CISDM database. Our empirical study reveals that there is short-term persistence for horizons of up to six months. We also find that the persistence significance levels are lower the longer the time horizon is. Thus our findings confirm the general consensus in the literature. Additionally, we identified one of the key drivers behind the varied results found in the literature; the large differences are related to the methodology used to assess performance persistence. Using correlation-, contingency-table-, and regression-based tests, we obtained evidence of performance persistence; however, the multi-period Kolmogorov-Smirnov test showed hardly any persistence. We thus conclude that the use of different methodologies is one of the main drivers behind the mixed results found in the literature.

Different hedge fund strategies also produce very different results. Convertible Arbitrage and Emerging Markets hedge funds provide very high levels of persistence, but Equity Long Only hedge funds have low levels of persistence. We thus conclude that persistence is related to the type of hedge fund strategy. However, an important result empirically observed is that the high levels of persistence found with some strategies might be due to return smoothing, e.g., Convertible Arbitrage hedge funds exhibit significant serial correlation as well as high levels of persistence. Finally, the results comparing different measures used to assess performance persistence, such as raw returns, alphas, and appraisal ratios, were very similar. The use of different performance measures is thus not the reason for the conflicting results found in performance persistence literature.

The current state of literature and our empirical findings indicate that there is short-term, but not long-term persistence in hedge fund performance. However, the effects that survivorship and backfilling bias as well as return smoothing have on persistence levels casts serious doubt on whether the observed short-term persistence is at all related to superior manager skill. And even if there is real short-term persistence in hedge fund returns due to manager skill, the problem is that it cannot be profitably exploited by hedge fund investors due to significant lockup periods, entry costs, and exit costs. Future research should thus concentrate on new methodologies for analysing long-term performance persistence in hedge fund returns. As shown in our literature overview, there are many studies that concentrate on short-term performance persistence of up to one year, but there has been very little study of two-year and three-year horizons. However, some recent research, such as Jagannathan *et al.* (2006) and Kosowski *et al.* (2007), has provided interesting new insights into long-term persistence by using new, sophisticated methodologies. Therefore, long-term persistence in hedge fund performance might be a promising area of research in the coming years.

References

- Ackermann, C., McEnally, R. and Ravenscraft, D., 'The performance of hedge funds: risk, return, and incentives', *Journal of Finance*, Vol. 54, No. 3, 1999, pp. 833–74.
- Agarwal, V. and Naik, N. Y., 'Multi-period performance persistence analysis of hedge funds', *Journal of Financial and Quantitative Analysis*, Vol. 35, No. 3, 2000a, pp. 327–42.
- Agarwal, V. and Naik, N. Y., 'On taking the alternative route: risks, rewards, and performance persistence of hedge funds', *Journal of Alternative Investments*, Vol. 2, No. 4, 2000b, pp. 6–23.
- Agarwal, V. and Naik, N. Y., 'Risk and portfolio decisions involving hedge funds', *Review of Financial Studies*, Vol. 17, No. 1, 2004, pp. 63–98.
- Agarwal, V., Daniel, N. D. and Naik, N. Y., 'Role of managerial incentives and discretion in hedge fund performance' *Working Paper* (Georgia State University and London Business School, October 2007).
- Amenc, N., El Bied, S. and Martellini, L., 'Predictability in hedge fund returns', *Financial Analysts Journal*, Vol. 59, No. 5, 2003, pp. 32–46.
- Asness, C., Krail, R. and Liew, J., 'Do hedge funds hedge?' *Journal of Portfolio Management*, Vol. 28, No. 1, Fall 2001, pp. 6–19.
- Baquero, G., ter Horst, J. and Verbeek, M., 'Survival, look-ahead bias and the persistence in hedge fund performance', *Journal of Financial and Quantitative Analysis*, Vol. 40, No. 3, 2005, pp. 493–517.
- Barès, P.-A., Gibson, R. and Gyger, S., 'Performance in the Hedge Funds Industry: An Analysis of Short and Long-Term Persistence', *Journal of Alternative Investments*, Vol. 6, No. 3, 2003, pp. 25–41.
- Black, F. and Scholes, M., 'Pricing of options and corporate liabilities', *Journal of Political Economy*, Vol. 81, No. 3, 1973, pp. 637–54.
- Blake, D. and Timmermann, A., 'Performance persistence in mutual funds', Report prepared for the Financial Services Authority (April 2003).
- Boyson, N. M. and Cooper, M. J., 'Do hedge funds exhibit performance persistence? A new approach' *Working Paper* (Northeastern University and Purdue University, November 2004).
- Brown, S. J. and Goetzmann, W. N., 'Performance persistence', *Journal of Finance*, Vol. 50, No. 2, 1995, pp. 679–98.
- Brown, S. J. and Goetzmann, W. N., 'Hedge funds with style', *Journal of Portfolio Management*, Vol. 29, No. 2, Winter 2003, pp. 101–12.
- Brown, S. J., Goetzmann, W. N. and Ibbotson, R. G., 'Offshore hedge funds: survival and performance 1989–1995', *Journal of Business*, Vol. 72, No. 1, 1999, pp. 91–117.
- Brown, S. J., Goetzmann, W. N., Ibbotson, R. G. and Ross, S., 'Survivorship bias in performance studies', *Review of Financial Studies*, Vol. 5, No. 4, 1992, pp. 553–80.
- Capocci, D., Corhay, A. and Hübner, G., 'Hedge fund performance and persistence in bull and bear markets', *European Journal of Finance*, Vol. 11, No. 5, 2005, pp. 361–92.
- Capocci, D. and Hübner, G., 'Analysis of hedge fund performance', *Journal of Empirical Finance*, Vol. 11, No. 1, 2004, pp. 55–89.
- Capon, N., Fitzsimons, G. J. and Prince, R. A., 'An individual level analysis of the mutual fund investment decision', *Journal of Financial Services Research*, Vol. 10, No. 1, 1996, pp. 59–82.
- Carhart, M., 'On persistence in mutual fund performance', *Journal of Finance*, Vol. 52, No. 1, 1997, pp. 57–82.
- Carhart, M., Carpenter, J. N., Lynch, A. W. and Musto, D. K., 'Mutual fund survivorship', *The Review of Financial Studies*, Vol. 15, No. 5, 2002, pp. 1439–63.
- Carpenter, J. N. and Lynch, A. W., 'Survivorship bias and attrition effects in measures of performance persistence', *Journal of Financial Economics*, Vol. 54, No. 3, 1999, pp. 337–74.
- Chan, N., Getmansky, M., Haas, S. M. and Lo, A. W. (2007), Systemic Risk and Hedge Funds, in: Carey, M., and R. Stulz (eds.): *The Risks of Financial Institutions and the Financial Sector*, University of Chicago Press, Chicago, IL, pp. 49–80.
- Chen, K. and Passow, A., *Quantitative Selection of Long-Short Hedge Funds*, FAME Research Paper No. 94 (2003).
- Couillard, M. and Davison, M., 'A comment on measuring the Hurst exponent of financial time series', *Physica A*, Vol. 348, 2005, pp. 404–18.

- De Souza, C. and Gokcan, S., 'Hedge fund investing: a quantitative approach to hedge fund manager selection and de-selection', *Journal of Wealth Management*, Vol. 6, No. 4, 2004, pp. 52–73.
- Ding, B. and Shawky, H. A., 'The performance of hedge fund strategies and the asymmetry of return distributions', *European Financial Management*, Vol. 13, No. 2, 2007, pp. 309–31.
- Droms, W. G., 'Does past performance predict future returns?' *Journal of Financial Planning*, Vol. 19, No. 5, 2006, pp. 60–69.
- Edwards, F. and Caglayan, M., 'Hedge fund performance and manager skill', *Journal of Futures Markets*, Vol. 21, No. 11, 2001, pp. 1003–28.
- Eling, M., 'Does the measure matter in the mutual fund industry?' *Financial Analysts Journal*, Vol. 64, No. 3, 2008, pp. 54–66.
- Fama, E. F. and French, K. R., 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics*, Vol. 33, No. 1, 1993, pp. 3–56.
- Fung, W. and Hsieh, D. A., 'Empirical characteristics of dynamic trading strategies: the case of hedge funds', *Review of Financial Studies*, Vol. 10, No. 2, 1997, pp. 275–302.
- Fung, W., and Hsieh, A., 'Performance characteristics of hedge funds and commodity funds: natural vs. spurious biases', *Journal of Financial and Quantitative Analysis*, Vol. 35, No. 3, 2000, pp. 291–307.
- Fung, W. and Hsieh, D. A., 'The risk in hedge fund strategies: theory and evidence from trend followers', *Review of Financial Studies*, Vol. 14, No. 2, 2001, pp. 313–41.
- Fung, W. and Hsieh, D. A., 'Hedge fund benchmarks: a risk based approach', *Financial Analyst Journal*, Vol. 60, No. 5, 2004, pp. 65–80.
- Géhin, W., 'A survey on the literature on hedge fund performance', *Working Paper* (Edhec Risk and Asset Management Research Center, October 2004).
- Getmansky, M., Lo, A. W. and Makarov, I., 'An econometric model of serial correlation and illiquidity in hedge fund returns', *Journal of Financial Economics*, Vol. 74, No. 3, 2004, pp. 529–609.
- Gibson, R. and Gyger, S., 'The style consistency of hedge funds', *European Financial Management*, Vol. 13, No. 2, 2007, pp. 287–308.
- Goltz, F., Martellini, L. and Vaissié, M., 'Hedge fund indices: reconciling investability and representativity', *European Financial Management*, Vol. 13, No. 2, 2007, pp. 257–86.
- Goodwin, T. H., 'The information ratio', *Financial Analysts Journal*, Vol. 54, No. 4, 1998, pp. 34–43.
- Gregoriou, G. N., 'Hedge fund survival lifetimes', *Journal of Asset Management*, Vol. 3, No. 3, 2002, pp. 237–52.
- Gregoriou, G. N. and Rouah, F., 'Last year's winning hedge fund as this year's selection: a simple trading strategy', *Derivatives Use, Trading & Regulation*, Vol. 7, No. 3, 2001, pp. 269–74.
- Gregoriou, G. N., Rouah, F. and Sedzro, K., 'Market timing and security selection: the case of hedge funds', *Derivatives Use, Trading & Regulation*, Vol. 8, No. 2, 2002, pp. 140–58.
- Grinblatt, S. and Titman, M., 'The persistence of mutual fund performance', *Journal of Finance*, Vol. 47, No. 5, 1992, pp. 1977–84.
- Gupta, B., Cerrahoglu, B. and Daglioglu, A., 'Evaluating hedge fund performance: traditional versus conditional approaches', *Journal of Alternative Investments*, Vol. 6, No. 3, 2003, pp. 7–24.
- Harri, A. and Brorsen, B. W., 'Performance persistence and the source of returns for hedge funds', *Applied Financial Economics*, Vol. 14, No. 2, 2004, pp. 131–41.
- Hendricks, D., Patel, J. and Zeckhauser, R., 'Hot hands in mutual funds: short-run persistence of relative performance', *Journal of Finance*, Vol. 48, No. 1, 1993, pp. 93–130.
- Henn, J. and Meier, I., 'Performance analysis of hedge funds', in: Dichtl, H., Kleeberg, J. M. and Schlenger, C., eds, *Handbuch Hedge Funds* (Uhlenbruch, Bad Soden/Ts, 2004), pp. 435–66.
- Herzberg, M. M. and Mozes, H. A., 'The persistence of hedge fund risk: evidence and implications for investors', *Journal of Alternative Investments*, Vol. 6, No. 2, 2003, pp. 22–42.
- Jagannathan, R., Malakhov, A. and Novikov, D., 'Do hot hands persist among hedge fund managers? An empirical evaluation' *NBER-Working Paper* 12015 (January 2006).
- Jensen, M. C., 'The performance of mutual funds in the period 1945–1968', *Journal of Finance*, Vol. 23, No. 2, 1968, pp. 389–416.
- Kat, H. M. and Menexe, F., 'Persistence in hedge fund performance: the true value of a track record', *Journal of Alternative Investments*, Vol. 5, No. 4, 2003, pp. 66–72.

- Koh, F., Koh, W. T. H. and Teo, M., 'Asian hedge funds: return persistence, style, and fund characteristics', *Working Paper* (Singapore Management University, June 2003).
- Kosowski, R., Naik, N. Y. and Teo, M., 'Do hedge funds deliver alpha? A Bayesian and bootstrap analysis', *Journal of Financial Economics*, Vol. 84, No. 1, 2007, pp. 229–64.
- Kouwenberg, R., 'Do hedge funds add value to a passive portfolio?' *Journal of Asset Management*, Vol. 3, No. 4, 2003, pp. 361–82.
- Lee, C.-F., Lee, J. C. and Lee, A. C., *Statistics for Business and Financial Economics*, 2nd ed. (World Scientific Publishing Company, Singapore, 2000).
- Li, Y. and Kazemi, H., 'Conditional properties of hedge funds: evidence from daily returns', *European Financial Management*, Vol. 13, No. 2, 2007, pp. 211–38.
- Liang, B., 'Hedge funds: the living and the dead', *Journal of Financial and Quantitative Analysis*, Vol. 35, No. 3, 2000, pp. 309–26.
- Liang, B. and Park, H., 'Risk measures for hedge funds: a cross-sectional approach', *European Financial Management*, Vol. 13, No. 2, 2007, pp. 333–70.
- Lo, A. W., 'Long-term memory in stock prices', *Econometrica*, Vol. 59, No. 5, 1991, pp. 1279–1313.
- Macbeth, J. D. and Merville, L. J., 'Tests of the Black-Scholes and Cox call option valuation models', *Journal of Finance*, Vol. 35, No. 2, 1980, pp. 285–301.
- Malkiel, B. G. and Saha, A., 'Hedge funds: risk and return', *Financial Analysts Journal*, Vol. 61, No. 6, 2005, pp. 80–88.
- McClave, J. T., Benson, P. G. and Sincich, T., *Statistics*, 8th ed. (Prentice Hall, Upper Saddle River, NJ, 2001).
- McNabb, D. E., *Research Methods for Political Science* (M.E. Sharpe, Armonk, NY, 2004).
- Mitchell, M. and Pulvino, T., 'Characteristics of risk and return in risk arbitrage', *Journal of Finance*, Vol. 56, No. 6, 2001, pp. 2135–75.
- Newey, W. and West, K., 'A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix', *Econometrica*, Vol. 55, No. 3, 1987, pp. 703–8.
- Park, J. M. and Staum, J. C., 'Performance persistence in the alternative investment industry', *Working Paper* (Long Island University and Columbia University, December 1998).
- Schneeweis, T., Kazemi, H. and Martin, G., 'Understanding hedge fund performance: research issues revisited – part 1', *Journal of Alternative Investments*, Vol. 5, No. 3, 2002, pp. 6–22.
- Sharpe, W. F., 'Mutual fund performance', *Journal of Business*, Vol. 39, No. 1, 1966, pp. 119–38.
- Sharpe, W. F., 'Asset allocation: management style and performance measurement', *Journal of Portfolio Management*, Vol. 18, No. 2, Winter 1992, pp. 7–19.
- Tanizaki, H., *Computational Methods in Statistics and Econometrics* (Dekker, New York, NY, 2004).
- Weiss, M. S., 'Modification of the Kolmogorov-Smirnov statistic for use with correlated data', *Journal of the American Statistical Association*, Vol. 73, No. 364, 1978, pp. 872–75.

Appendix

Table A1
Definition of alpha

This table reports the definition of alpha in 16 studies on hedge fund performance persistence. Column 1 gives the authors, Column 2 the number of factors considered, and Column 3 the factors themselves. Abbreviations: AMEX: American Stock and Options Exchange, CRSP: Center for Research in Security Prices, ER: excess return, HML: high minus low book to price ratio, IHML: international high minus low book to price ratio, LB: Lehman Brothers, NYSE: New York Stock Exchange, MER: market excess return, SMB: small minus big.

Authors	Number of factors	Factors
Agarwal and Naik (2000a)		average return of all the funds using the same strategy

Table A1
Continued.

Authors	Number of factors	Factors
Agarwal and Naik (2000b)	1	average return of all the funds using the same strategy
Baquero <i>et al.</i> (2005)		Tremont hedge fund style indices
Barès <i>et al.</i> (2003)	8	hedge fund style factors obtained by a principal component analysis
Boyson and Cooper (2004)	19	<ul style="list-style-type: none"> ■ 6 traditional indices: US Dollar, Gold, Commodities, CRSP Value Weighted, LB Aggregate Bond, LB 30 Year US Treasury ■ Fama and French (1993) + Carhart (1997): HML, SMB, Momentum ■ 10 CSFB/Tremont indices
Brown <i>et al.</i> (1999)	1	Tremont hedge fund style indices
Capocci <i>et al.</i> (2005)	11	<ul style="list-style-type: none"> ■ Jensen (1968) + Fama and French (1993) + Carhart (1997): MER (all NYSE, AMEX, and Nasdaq stocks), SMB, HML, IHML, Momentum ■ Agarwal and Naik (2004): Lehman BAA Corp. Bond, MSCI World excluding U.S., LB U.S. Aggregate Bond, Salomon World Government, Goldman Sachs Commodity ■ JP Morgan Emerging Market Bond
Capocci and Hübner (2004)	11	<ul style="list-style-type: none"> ■ Jensen (1968) + Fama and French (1993) + Carhart (1997): MER (all NYSE, AMEX, and Nasdaq stocks), SMB, HML, IHML, Momentum ■ Agarwal and Naik (2004): Lehman BAA Corp. Bond, MSCI World excluding US, LB U.S. Aggregate Bond, Salomon World Government, Goldman Sachs Commodity ■ JP Morgan Emerging Market Bond
Chen and Passow (2003)		<ul style="list-style-type: none"> ■ Jensen (1968) + Fama and French (1993): ER (Russel 3000), HML, SMB ■ Agarwal and Naik (2004): Goldman Sachs Commodity Index
Edwards and Caglayan (2001)	6	<p>model similar to that of Fama and French (1993)</p> <ul style="list-style-type: none"> ■ Jensen (1968) + Fama and French (1993): MER (S&P 500) HML, SMB ■ WML (winners minus losers), TERM (a long-term government bond portfolio minus the 1-month-lagged 30-day T-bill return), DEF (monthly return on a portfolio of long-term corporate bonds minus the monthly return on a portfolio of long-term government bonds)

Table A1
Continued.

Authors	Number of factors	Factors
Gregoriou and Rouah (2001)		Jensen (1968): MER (S&P 500 and MSCI, separately)
Harri and Brorsen (2004)	8	<ul style="list-style-type: none"> style analysis similar to Sharpe (1992) and Fung and Hsieh (1997) ■ three equity classes: S&P500, MSCI World excluding U.S., and MSCI Emerging Markets ■ two bond indices: a government bond index and a corporate bond index ■ Cash (1-month eurodollar deposit), gold, currency
Jagannathan <i>et al.</i> (2006)	3	<ul style="list-style-type: none"> ■ Jensen (1968): MER (CRSP) ■ Self-reported style index J from HFR ■ Additional style index K from HFR
Koh <i>et al.</i> (2003)	7	<ul style="list-style-type: none"> ■ Asian equity factor (broken down into an Asia ex Japan factor and a Japan factor) ■ Asian bond factor ■ U.S. equity factor ■ Fama and French (1993) + Carhart (1997): SMB, HML, momentum
Kosowski <i>et al.</i> (2007)		<p>seven-factor model developed by Fung and Hsieh (2004)</p> <ul style="list-style-type: none"> ■ S&P 500 return minus risk-free rate ■ Wilshire small cap minus large cap return ■ change in the constant maturity yield of the 10-year Treasury ■ change in the spread of Moody's Baa minus the 10-year Treasury ■ bond primitive trend following strategy ■ currency and commodities
Kouwenberg (2003)	3	style adjusted: portfolio of S&P 500, Nasdaq, and Option Selling Strategies