Hedge-Fund Benchmarks: Information Content and Biases

William Fung and David A. Hsieh

We discuss the information content and potential measurement biases in hedge fund benchmarks. Hedge-fund indexes built from databases of individual hedge funds inherit the measurement biases in the databases. In addition, broad-based indexes mask the diversity of individual hedge-fund return characteristics. Consequently, these indexes provide incomplete information to investors seeking diversification from traditional asset classes through the use of hedge funds. The approach to constructing hedge-fund benchmarks we propose is based on the simple idea that the most direct way to measure hedge-fund performance is to observe the investment experience of hedge-fund investors themselves—the funds of hedge funds. In terms of measurement biases, returns of FOFs can deliver a cleaner estimate of the investment experience of hedge-fund investors than the traditional approach. In terms of risk characteristics, indexes of FOFs are more indicative of the demand-side dynamics driven by hedge-fund investors' preferences than are broad-based indexes. Therefore, indexes of FOFs can provide valuable information for assessing the hedge-fund industry's performance.

We analyze the problems in creating or choosing benchmarks for assessing the performance characteristics of hedge funds. We begin by discussing potential measurement biases embedded in the historical returns of hedge funds. A complete record of every single hedge fund simply does not exist. The lack of data arises from three reasons. First, hedge-fund participation in any database is voluntary. Organized as private investment vehicles, hedge funds generally do not disclose their activities to the public. Second, most commercially available hedge-fund databases only came into existence in the mid-1990s. Third, different databases have different criteria for including funds. We focus on two important biases arising from the data themselves that affect the analysis of hedge-fund data: survivorship bias and selection bias.

William Fung is co-chief executive officer at PI Asset Management, LLC, Delaware and visiting research professor at the Centre for Hedge Fund Research and Education, London Business School, London, United Kingdom.
David A. Hsieh is professor of finance at the Fuqua School of Business, Duke University, Durham, North Carolina.

An important attribute of hedge-fund investing is the diversity of the funds' performance characteristics. Nevertheless, a number of vendors have constructed composites of hedge-fund performance. In particular, two organizations have made serious attempts to create hedge-fund indexes that are comprehensive and transparent. Hedge Fund Research (HFR) and CSFB/Tremont (CT) both attempt to rectify some of the measurement biases we described. Nonetheless, measurement bias is unavoidable. We thus discuss measurement biases that may arise in constructing hedge-fund benchmarks based on historical returns.

Hedge-fund managers naturally focus their efforts on liquid markets, where trading opportunities and leverage are readily available. Thus, as the dynamics in the global markets change, the nature of hedge funds in operation change over time—through the birth and death rate of funds and changes in the trading styles of existing funds. Benchmarking such a dynamic industry is in itself a difficult task, and the difficulty is compounded by the fact that the hedge funds that compose the benchmarks are drawn from a population of funds managed by nimble managers with diverse investment styles. We address how well existing indexes reflect the risk characteristics of hedge funds.

Finally, we propose a new approach to assessing the performance characteristics of hedge funds. It is based on the simple idea of looking directly at the investment experience of hedge-fund investors, namely, the funds of hedge funds (FOF). We argue for the use of FOF returns, rather than returns of individual hedge funds, to construct hedge-fund indexes. One reason is that the performance characteristics of FOFs are driven not only by the opportunities in the global markets but also by investor preferences. To stay in business, FOFs have to respond to what investors demand. Another reason is that data from the demand side of hedge funds, the FOF, are less susceptible to the measurement biases we describe.

Database Biases and Information Content

Advances in information technology in the past decade have led to dramatic improvements in the investment arena. Nowadays, U.S. investors can readily access historical data with unquestioned quality and consistency on almost any security or mutual fund. The same cannot be said, however, about hedge funds. Although hedge-fund database services have expanded dramatically since the 1990s, a generally accepted provider of standardized information on hedge funds has yet to emerge because of the absence of a centralized depository of performance records similar to the Investment Company Institute. Both the scope and the quality of the data vary among hedge-fund database vendors. Therefore, *caveat emptor* is still very much the case for users of available hedge-fund performance data. We consider here some problems frequently encountered involving the information content of a sample of hedge-fund returns.

Consider an investor interested in assessing the general performance characteristics of hedge funds. The natural way to go about it is to obtain a sufficiently broad "sample portfolio" of hedge funds and construct its pro forma return statistics. In assessing these statistics, what kind of measurement errors should investors be aware of? The answer to this question is especially important for the benchmarking of hedge-fund performance. To gain insight into this question, we begin by examining the way the hedge-fund industry is organized and its impact on data collection.

Organized as private and frequently offshore investment vehicles, hedge funds generally do not disclose their activities to the public.[1] A complete record of every single hedge fund simply does not exist. Available information comes as samples of hedge funds in the form of databases. To sharpen the discussion, we will use the term "universe" (or "population") to denote the collection (or set) of all hedge funds that have operated, past or present, dead or alive, and we will use the term "database" to refer to a subset of the population of hedge funds collected by data vendors.[2]

The incompleteness of hedge-fund data has several reasons. First, hedge-fund participation in any database is voluntary, and a well-known result in statistical sampling theory is that voluntary participation can lead to sampling biases. Voluntary participation means that only a portion of the universe of hedge funds is observable.

Second, most commercially available hedge-fund databases came into existence in the mid-1990s; vendors began collecting hedge fund data in earnest in 1993 and 1994. Inevitably, pre-1994 observable data on hedge funds contain measurement biases as a natural consequence of the way the hedge fund industry evolved. Information on funds that ceased operation before they could be included in databases may have been lost forever.

Third, different databases have different criteria for including funds and different data-collection methods. Post-1994 hedge fund data are less susceptible to measurement biases than the pre-1994 data, but these differences in data collection and criteria can lead to other forms of measurement biases.

These three reasons can lead to important differences between the hedge funds in a database and those in the population. We focus on two main biases that arise in analyzing hedge-fund data—survivorship bias and selection bias. We further distinguish between biases that are consequences of sampling from an unobservable universe of hedge funds, which we call "natural biases," and those that arise from the way data vendors collect hedge-fund information, which we call "spurious biases."

---

[1] See Fung and Hsieh (1999) for an overview of how hedge funds are organized and their economic rationale.

[2] This distinction does not arise in the mutual fund industry, where public disclosure enforces the convergence of the universe and the database of all mutual funds. With the hedge-fund industry, the very fact that the population is not observable means that a single database (or for that matter, the set of all databases) need not coincide with the universe.

Survivorship Bias. Survivorship bias arises when a sample of hedge funds includes only funds that are operating at the end of the sampling period and excludes funds that have ceased operations during the period. Presumably, funds cease operation because of poor performance. Therefore, the historical return performance of the sample is biased upward and the historical risk is biased downward relative to the universe of all funds. Survivorship bias is a natural consequence of the way the hedge-fund industry evolved.[3] Therefore, in the context of analyzing hedge-fund data, survivorship bias cannot be completely mitigated.

The effect of survivorship bias is well documented in the mutual fund literature.[4] The standard procedure, as in Malkiel (1995), is to obtain the population of all mutual funds that operated during a given time period. The average return of all funds is compared with that of the surviving funds at the end of the period. The return difference is survivorship bias.

Unlike the case for mutual funds, survivorship bias in hedge funds cannot be measured directly because the universe of hedge funds is not observable. Survivorship bias can only be estimated by using hedge funds in a database. This limitation creates a new set of problems that do not arise for mutual funds.

The first problem concerns information on hedge funds that ceased to exist before database vendors started their data collection. Because of the lack of public disclosure, database vendors have only sketchy information on hedge funds that ceased operation (i.e., died) prior to the mid-1990s.[5] Thus, hedge-fund databases, no matter how broad, are vulnerable to survivorship bias,

---

[3] Technically, over any sample period, if a complete record of defunct funds is available, survivorship bias can be mitigated through tedious data manipulation. The problem is in verifying the completeness of historical records on defunct hedge funds.

[4] See Grinblatt and Titman (1989), Brown, Goetzmann, Ibbotson, and Ross (1992), and Malkiel (1995).

[5] The reason is that these funds predated the existence of most hedge-fund databases.

especially prior to the mid-1990s, and analysts cannot assess survivorship bias prior to the mid-1990s.

The second problem arises from the difference between funds that simply exited a database (termed “defunct funds”) and funds that ceased operation (termed “dead funds”). A defunct fund is a fund that was in a database but ceases to report information to the database vendor; a dead fund is one that is known to have terminated operations. Of course, a dead fund must also be a defunct fund, but a defunct fund need not be dead. For example, a fund delisted by the database vendor is a “defunct” but not “dead” fund.[6] Presumably, vendors delist funds they believe are likely to harm their reputations for providing reliable information to their customers. In that case, delisted funds are likely to have a less accurate—and, in most cases, a worse—performance history than the typical hedge fund.

Another type of fund that is defunct but not dead is one that voluntarily stops reporting information to a database vendor because it has reached the optimal size for its style of trading. The diminished appetite for new capital, coupled with a preference for privacy, often means that the fund no longer wants to provide its performance statistics to database vendors.[7] This type of defunct fund may actually have a higher return and lower risk than the typical hedge fund in the universe or in the database.[8]

In short, defunct funds in a database are not necessarily dead funds in the universe of hedge funds. Defunct funds may include dead funds, delisted funds (that may or may not be dead), and operating funds that reached capacity constraints. With this caveat in mind, we used both surviving and

---

[6] Generally, database vendors have listing requirements that a hedge fund must meet to be included in their databases. Such requirements typically involve a minimum amount of assets under management, timely reporting of information, and the ability of the database vendor to verify the fund’s performance record.

[7] Fung and Hsieh (1997a) cited anecdotal evidence that some managers with superior performance have refused to participate in databases because they have reached capacity constraints and are no longer looking for investors.

defunct funds from a database to estimate the survivorship bias as the difference between the returns of the “observable portfolio” and those of the “surviving portfolio.”[9]

Given a set of hedge funds for a sampling period, the observable portfolio consists of an equally weighted investment in *all funds* in the portfolio rolling forward from the beginning of the period. The observable portfolio is rebalanced when a new fund is added to the portfolio or when a fund becomes defunct.[10] The surviving portfolio consists of an equally weighted investment *only in* those funds that survived until the end of the sampling period. Going forward in time, this portfolio is rebalanced only when a new fund is added to the sample, but by construction, it never has to be rebalanced when a fund becomes defunct.

Following this approach, Malkiel estimated the survivorship bias in mutual funds to be 0.5 percentage points a year in returns. Fung and Hsieh (2000b) estimated the survivorship bias in hedge funds in the TASS database to average roughly 3 pps a year. This figure is consistent with Brown, Goetzmann, and Ibbotson (1999), who studied offshore hedge funds. We refer to this 3 pps figure as an estimate because we used a sample, not the population, of hedge funds in this study and because we used defunct funds in a database to proxy for dead funds in the population.

Selection and “Instant History” Biases. The combination of the voluntary nature of hedge-fund information in databases and the different inclusion processes of database vendors can lead to differences between the performance of funds in a database and that of funds in the universe of hedge funds—that is, selection bias.

---

[8] Of course, if the database from which a sample of hedge funds is extracted has survivorship bias, then the smaller sample portfolio is likely with even greater reason to exhibit survivorship bias.

[9] Here, we are following a methodology first used by Malkiel.

[10] This calculation requires that data vendors retained records of defunct funds.

Selection bias manifests itself in two basic ways. Hedge funds that satisfy the inclusion criteria of a vendor may enter a database on the basis of their track record and assets under management. On the one hand, presumably, only those funds that have “good” performance and are looking to attract new investors want to be included in a database. Therefore, hedge funds in a database tend to have better performance than those that were excluded. On the other hand, hedge funds may not be participating in a database because they are not looking to attract new investors. These self-excluded funds may have better performance than the average hedge fund. Thus, the net effect of selection bias on the returns of hedge funds in a database is ambiguous.

In addition to the biases arising from the voluntary nature of fund participation in a database, the database vendors themselves may introduce sampling biases through their inclusion criteria. For example, of the three major hedge-fund database vendors, one (HFR) excludes managed futures programs but two (TASS and Managed Account Research, MAR) include them.

The magnitude of the selection bias in a database is difficult to determine empirically because one cannot compare the observed hedge funds in the database with the unobservable hedge funds in the population. Differences in the number and the identity of hedge funds among databases, however, are indicative of selection bias. (We will return to this issue in a later section.)

A problem related to selection bias has come to be known as the “instant history bias.”[11] When a data vendor adds a fund into a database, the vendor often backfills the fund’s historical returns into the database. Thus, funds enter a database with, in the words of Park (1995), instant history. It occurs because hedge funds usually undergo an incubation period. The fund manager starts the fund with a small amount of seed capital (often from friends and relatives in addition to the manager’s personal capital). When the fund’s track record is satisfactory, the fund manager markets the fund to investors, which often include asking to be included in a hedge-fund database. Because the fund manager can decide when to “reveal” the fund’s track record, the returns from the incubation period can reasonably be presumed to be high.

To estimate the magnitude of instant history bias, Fung and Hsieh (2000b) studied the hedge funds in the TASS database, which reports the inception date of each fund as well as the date the fund entered the database. Fung and Hsieh measured the instant history bias as the average difference between two portfolios. The first was the observable portfolio, as defined previously; the second was the “adjusted” observable portfolio, which was constructed in the same manner as the observable portfolio but after dropping the first 12 monthly returns of every fund. (On average, the incubation period—from a fund’s inception to its entry into TASS—is one year.) The adjusted observable portfolio’s return was found to be the lower, on average, by 1.4 pps year.

Fung and Hsieh (2000b) considered selection bias and instant history bias to be spurious biases because both the causality and magnitude of these biases are inherent in the data-collection process. Most spurious biases can be remedied, but only with careful and tedious data manipulation: Selection bias can be eliminated if hedge-fund databases eventually converge to the universe of hedge funds. Instant history bias can be remedied by dropping the returns of a fund prior to its entry into a database. In contrast, natural biases (such as survivorship) generally cannot be rectified.

Later, we propose a simple remedy for both types of performance measurement biases. Next, however, we examine the impact of measurement biases on an important application of hedge-fund data—the benchmarking of hedge-fund performance.

Measurement Errors and Index Differences. Vendors have created two broadly based hedge-fund indexes to benchmark the performance of the hedge-fund industry. They are the Hedge Fund Research Performance Index (HFRI)

---

[11] Instant history bias was first analyzed by Park (1995).

and the CSFB/Tremont Hedge Fund Index (CTI).[12] The HFRI is an equally weighted index of more than 1,000 hedge funds tracked by HFR, and the CTI is a value-weighted index (with assets under management as "value" in the weighting scheme) based on a sample of approximately 300 funds extracted from the TASS database. The CTI was constructed to be an "investable" index, whereas the HFRI was designed to be a proxy for the hedge-fund industry.[13] These two indexes are typical of hedge-fund portfolio benchmarks, so understanding their potential measurement errors—and thus their information content—is important.

Hedge-fund databases, including HFR and TASS, inevitably suffer from natural and spurious biases. According to Ackerman, McNally, and Ravenscraft (1999) and Liang (2000), both databases have limited records of funds that became defunct before 1994. Hence, both HFR and TASS suffer from survivorship bias for pre-1994 data. In addition, the absence of any defunct funds before 1994 indicates that HFR and TASS began their data collection around 1994. Thus, their historical pre-1994 data suffer from selection bias because data prior to 1994 had to be compiled and backfilled.[14] For these reasons, in our study, we ignored the HFRI data prior to 1994. (Note that the CTI data start in 1994.)

For the post-1994 period, both databases have information on defunct as well as operating funds. Therefore, the pro forma returns of the two hedge-fund indexes should not suffer from survivorship bias related to defunct funds so long as the vendors adhered to proper adjustment procedures when computing the time series of index returns. *Relative to the universe of all hedge funds,*

---

[12] Of the three most widely known hedge-fund database vendors—HFR, TASS, and MAR—MAR does not report a performance index of hedge funds as a group. MAR cites the diversity of trading styles as the reason. We analyze this question in the section "Problems with Broad Benchmarks".

[13] At the time we were writing this article, a new value-weighted index of hedge funds was added to the family of HFR indexes, but pre-2000 returns for this index had not been released.

[14] Although the MAR existed much before 1994, its focus was historically on Commodity Trading Advisors.

however, the indexes may suffer from various forms of natural and spurious biases.

This possibility rests on four reasons. First, no realistic way exists of verifying that complete records of defunct funds were used to adjust the index returns for survivorship bias, especially prior to the mid-1990s. Second, differences in data-collection methodologies could result in different degrees of selection bias and instant history bias. That is, “missing funds” could be a consequence of data-collection methodologies. Third, different approaches to index construction can result in performance differences. Fourth, the CTI can be interpreted as a tracking portfolio of the TASS universe of hedge funds, so it must naturally contain tracking errors, whereas the HFRI is supposed to be an average of *all* hedge funds tracked by HFR.

Inherited survivorship bias. Indexes and benchmarks inherit survivorship bias from the databases on which they are built. Currently, observable hedge funds in databases do not fully reflect the universe of all hedge funds. In time, observable funds may converge to the universe of all hedge funds, and from that point forward, analysts can remedy survivorship bias by analytical methods. Until convergence occurs, however, performance statistics derived from the observable funds remain biased estimators of the population statistics.[15] And time series of returns prior to the “point of convergence” will remain vulnerable to survivorship bias.[16] These problems in the data affect indexes and benchmarks based solely on samples of observable funds.

The HFR and TASS databases yield different estimates of survivorship bias. Ackerman et al., Fung and Hsieh (2000b), and Liang found that the attrition rate (i.e., the percentage of funds that become defunct each year) is much higher in TASS. As a result, the measured survivorship bias (i.e., the performance difference between all funds and surviving funds) is also higher in

---

[15] This issue is separate from the question of stationarity of the return time series.

TASS. Interestingly, TASS's attrition rate and survivorship bias are comparable to those reported by Brown, Goetzmann, and Ibbotson (1999), who used the *U.S. Offshore Hedge Funds Directory*. What is not clear is why the attrition rate is so low in HFR.

Inherited instant history and selection biases. Indexes and benchmarks also inherit instant history and selection biases from the databases on which they are built. To begin with, HFR and TASS backfill the return history of funds that enter their databases. Thus, they both suffer from instant history bias.

In addition, both HFR and TASS are likely to have selection bias because of idiosyncrasies in their data-collection methods. For the purposes of this article, we focus only on those methods that are relevant to our analysis.[17] Most notable of the differences in data collection is that HFR excludes funds that are generally referred to as "commodity funds" whereas CTI includes "managed futures," an important subset of commodity funds.[18]

In addition to inclusion criteria, a large number of hedge funds appear to report to only one database vendor. For instance, Liang found only 465 common funds out of the 1,162 funds in HFR and the 1,627 funds in TASS.

Thus, the potential for selection bias clearly exists in the hedge-fund databases. Its impact on benchmark indexes created from the databases, however, is difficult to assess directly.

Different weighting schemes. The weighting schemes used by index and benchmark creators can generate differences in the returns reported. Specifically, the HFR indexes use equal weights, whereas the CT indexes use value weights. First, consider the historical returns of the two indexes as

---

[16] In much the same way, historical returns on mutual funds before the 1960s are vulnerable to survivorship bias because of sketchy records on defunct mutual funds.
[17] See Liang for a more detailed comparison of the two databases.
[18] See Fung and Hsieh (1999) for a description of the differences, or lack of differences, between commodity funds and hedge funds.

reported in **Table 1**.[19] In the two years with the lowest returns and the two years with the highest returns for the CTI, returns for the HFRI and the CTI are significantly different. In the other years, the differences are small.

Such sizable return differences can result from the different weighting schemes of the two indexes. The HFRI represents the returns to a “contrarian” asset allocation strategy because an equally weighted portfolio is rebalanced by selling “winners” and buying “losers” every month. To maintain equal weighting, assets have to be diverted from performing funds to “underperforming” funds. (We later discuss the problems of implementing such a strategy.) In contrast, the CTI represents the returns to a momentum-driven asset allocation strategy. A value-weighted portfolio allows winners to naturally increase their weight in the index and losers to naturally reduce their weight. The differences in the two strategies generate path-dependent divergence in performance, especially in such a diverse universe of assets as hedge-fund investments.

The difference in weighting schemes can account for the 7.4 pps difference between the HFRI and the CTI in 1999 in the following way. According to both data sources, during the 1998 turmoil, emerging market hedge funds lost well over 30 percent of their value. This loss was followed by a dramatic rebound in 1999 in which the returns to emerging market hedge funds were in the region of 50 percent. To illustrate the effect of the weighting schemes on index performance, Panel A in **Table 2** provides returns to the subindexes of emerging market funds for the two indexes and (based on figures reported by HFRI) the portfolio weights of emerging market hedge funds in an equally weighted portfolio (as in HFRI) versus the weights in a value-weighted portfolio (as in CTI).

After the August 1998 debacle, the HFRI would have diverted assets into emerging market hedge funds (thereby holding their weight in the index to

---

[19] Monthly performance data on the CTI are available through the CT Web site for January 1994 to date.

about 12 percent).[20] During their rebound in 1999, these funds contributed a gain of 6.57 percent (55.22 percent × 11.9 percent) toward the overall HFRI return. In contrast, the CTI approach would have allowed the weight of emerging market hedge funds to fall as their assets shrank (from 9.50 percent in 1998 to 7.10 percent in 1999), so these funds contributed only 3.18 percent (44.82 percent × 7.10 percent) in 1999 toward the overall CTI return. The differential contribution of emerging market hedge funds thus would have accounted for nearly half of the 7.4 pps difference in 1999 between the two overall indexes.[21]

One other noteworthy aspect is that the substantial return difference between the HFR and CTI subindexes of emerging market hedge funds in this 1994–99 period is consistent with the existence of selection bias—that different databases can contain different samples of the universe of hedge funds.

Panel B of Table 2 shows that different weights of “global/macro” funds may explain the 9.1 pps performance discrepancy between the HFRI and CTI in 1997. The global/macro subindex contributed a small gain of 1.13 percent (6.0 percent × 18.82 percent) in 1997 toward the overall HFRI return. In contrast, it contributed a large gain of 8.83 percent (23.8 percent × 37.11 percent) in 1997 toward the overall CTI return because of its higher value weight and the fact that the CTI macro/global subindex recorded a much higher return than its counterpart in the HFRI. The difference of 7.7 pps accounts for much of the 9.1 pps difference between returns to the two overall indexes in 1997.

These examples point to spurious biases generated by the different index construction methods. We turn next to an analysis of the risk characteristics of these two indexes.

---

[20] Technically, the weights are adjusted on a monthly basis. Here, we have used annual figures to illustrate our point.

[21] A value-weighting method in the CTI would have implied less exposure in emerging market funds, and the emerging market funds in the CTI also returned less than their counterparts in the HFRI. A crude estimate is 7.1 percent × 44.82 percent = 3.18 percent, compared with the 6.57 percent of the HFRI, which is almost half the difference between the performance of the two indexes.

Problems with Broad Benchmarks

Broadly based benchmarks can mask interesting performance characteristics of hedge funds. An important issue for hedge-fund benchmark is whether it captures the key performance characteristics of individual hedge funds—characteristics that would attract institutional investors looking for alternatives to traditional asset classes and investors who are simply looking for good performance on a risk-adjusted basis. An important attribute of hedge-fund investing is the diversity of styles used by hedge-fund managers, but diversity comes at a cost. Finding a benchmark that reflects the overall performance characteristics of the hedge-fund industry as well as its diversity may be quite difficult. To address this issue, we analyzed the risk characteristics of the HFRI and CTI.

The correlation coefficients for the HFRI and CTI and nine standard asset class indexes over the sample period of 1994–1999 are given in Panel A in **Table 3**. The nine standard classes are one-month Eurodollar deposits; the Goldman Sachs Commodity Index (GSCI); for equities, the Morgan Stanley U.S. and World ex-U.S. indexes; for bonds, the J.P. Morgan U.S. and world ex-U.S. government bond indexes; the U.S. Federal Reserve's U.S. dollar index against major currencies; the International Finance Corporation's Investable Emerging Market Equities Index; and the Merrill Lynch High Yield bond index. Both the HFRI and CTI are strongly positively correlated with U.S. equities, non-U.S. equities, emerging market equities, and high-yield bonds.

The high degree of correlation we found between the hedge-fund indexes and certain standard asset classes contrasts to earlier empirical findings in Fung and Hsieh (1997a), Schneeweis and Spurgin (1998), Brown et al. (1999), and Ackerman et al., who reported low correlations between indexes of standard asset classes and *individual* hedge funds. To standardize the results found when researchers looked at individual hedge funds and the results we found for the two hedge-fund indexes, we used the methodology in Fung and

Hsieh (1997a) to examine a sample consisting of 1,129 hedge funds in the TASS database that had at least 36 monthly returns between 1994–99. We regressed their monthly returns on the nine broad-based market indexes. Table 4 reports the distribution of the adjusted R^2 from these regressions. More than 50 percent of these funds regressions have an adjusted R^2 below 0.3. Our results for individual hedge funds are thus very similar to the result in Fung and Hsieh (1997a).

In contrast, the HFRI and CTI have much higher R^2s when regressed against the standard asset indexes. Panel B in Table 3 reports the results of the regressions for the 1994–98 period. For the HFRI, the adjusted R^2 of the regression is quite high, at 0.76. The HFRI is positively related to the U.S. stock market, emerging market equities, high-yield U.S. bonds, and the GSCI. For the CTI, the adjusted R^2 of the regression is somewhat lower but is still substantially higher than for the average hedge fund in the sample. A direct comparison of the regression coefficients in Panel B of Table 3 with the simple correlation coefficients reported in Panel A is difficult. The regressors clearly exhibit colinearity, but scatterplots of the indexes against each of the standard asset indexes reveal that the relationship is basically linear.[22]

The evidence indicates that hedge-fund benchmarks have much greater exposure to traditional asset categories than the typical hedge fund has—which leaves us with a puzzle: *By diversifying among hedge funds, an investor may be exchanging idiosyncratic hedge-fund risk for systematic exposure to traditional risk factors*. The implication is most uncomfortable for investors looking to hedge-fund investments as a mean of diversifying a portfolio of traditional asset classes.

A natural explanation that apparently reconciles the low correlation of individual hedge funds with traditional asset markets and the high correlation of broad-based hedge-fund indexes with those assets is that individual hedge

---

[22] These scatterplots are available from the authors upon request.

funds have significant common risk factors.[23] With traditional equities, returns of individual equities contain a large component of idiosyncratic (nonsystematic) risk. Modern portfolio theory posits that a sufficiently large portfolio of equities can diversify away these idiosyncratic risks, leaving only systematic risk as the dominant risk in the portfolio. It would be tempting to apply the same argument to portfolios of hedge funds to explain the emergence of traditional risk factors at the index level. However, a closer examination of the evidence on individual hedge-fund returns does not favor this explanation.

Fung and Hsieh (1997a) showed that substantial style diversity characterizes hedge funds. The authors found five principal components (common risks) in 409 hedge funds and commodity funds. These principal components themselves have low correlations with the indexes of standard assets. This evidence is not consistent with the view that individual hedge funds have substantial idiosyncratic risk. It is consistent, however, with the conjecture that hedge funds have common risk characteristics that are not "systematic" in the traditional sense of being highly correlated with standard asset indexes. Unlike idiosyncratic risks commonly found in traditional equities, these risk characteristics are common among groups of hedge funds, referred to by the authors as Style Groups, and cannot be easily diversified away. We demonstrate this assertion empirically as follows.

A standard method of determining a portfolio's ability to diversify away idiosyncratic risks is to examine how the standard deviation of the portfolio changes with the number of assets in the portfolio. If the portfolio standard deviation drops quickly when more assets are added, then it contains a substantial amount of idiosyncratic risk.

To determine this relationship for hedge funds, we estimated the average monthly standard deviation of 1,000 randomly created portfolios of hedge funds in the TASS database. For example, for a portfolio of 20 funds, we

---

[23] We use the term "common risk factors" to avoid the confusing usage of the term "systematic risks."

randomly drew (without replacement) 20 funds from the sample of TASS funds that had returns at any time during a given time period—1994–1998. We then formed an equally weighted portfolio of these funds and computed the monthly standard deviation of the portfolio. We repeated this procedure 1,000 times and recorded the average monthly standard deviation of these 1,000 portfolios.[24]

The line “Individual Hedge Funds” in **Figure 1** shows how the average standard deviation of the hedge-fund portfolios changes with the number of funds in the portfolio. (We will discuss the line “Funds of Hedge Funds” in the section on “Tracking Error” later.) Although the standard deviation declines rapidly, it continues to decline without stabilizing to a fixed number. Indeed, it takes a portfolio of at least 120 hedge funds to have a standard deviation within 10 percent of that of an equally weighted portfolio of all TASS hedge funds. This pattern is quite different from the pattern for equity portfolios, where it is well known that only a few dozen stocks are needed in a portfolio to achieve the standard deviation of the market portfolio.

This evidence is consistent with the presence of significant common, but not easily diversifiable, risk factors among the hedge funds. In such a case, on average, a large number of funds are needed before a portfolio converges to a stable standard deviation.

The study by Fung and Hsieh (1997) also found that groups of hedge funds have strong correlations with each other, but their average returns have low correlation with standard benchmark returns. A single hedge-fund benchmark will not be able to reflect this heterogeneity in hedge funds. Therefore, the question is: What risk characteristics does a broad-based index of hedge funds reflect?

To answer this question, we need to explore the interaction between global market dynamics and the growth of the hedge-fund industry. Hedge-fund managers naturally focus their efforts on liquid markets where trading

[24] Because we did not require each fund to have performance information for the entire sample, we adjusted the portfolio weights to allow for entry and exit.

opportunities and leverage are readily available. In the past few years, the global equity markets certainly have been most conducive to these managers' inclinations. Therefore, as **Table 5** shows, both the number and amount of capital managed by "equity-oriented" funds have increased dramatically since 1994. A broad-based index of hedge funds is likely to reflect this trend. Consequently, broad-based indexes of hedge funds are more likely to reflect the risk characteristics inherent in the recent "popular bets" among hedge-fund managers, which bets have a significant degree of equity content.

Because of this concentration, the hedge-fund indexes understate the diversity of hedge-fund trading styles in general and overstate the risk of style convergence. When a large number of hedge funds in a portfolio converge into a similar set of bets, portfolio diversification implodes. For example, a large number of hedge-fund managers took big bets on U.S. and European bonds in 1993 and were caught in the market turmoil of 1994 when the U.S. Federal Reserve unexpectedly raised interest rates (see Fung and Hsieh 2000a). Therefore, for investors seeking diversification from traditional asset classes, properly constructed subindexes of specific hedge-fund trading styles are more informative in terms of risk than a broad-based index.

Benchmarks Based on Hedge-Fund Investors' Experience

As we have discussed, when investors measure hedge-fund performance using pro forma returns of a portfolio of individual hedge funds extracted from a database, natural biases can arise from the way the hedge-fund industry is organized. These natural biases cannot be easily remedied. We now examine potential solutions to the problems that also emanate from the idiosyncrasies of the hedge-fund industry.

If an analyst wants to estimate the investment experience of hedge funds, why not look directly at the experience of the hedge-fund investors themselves?

FOFs as Proxies for Hedge-Fund Investment Experience. Unlike mutual funds, for which the concept of FOF never gained popularity, the structure of the hedge-fund industry has led to the demand for and the existence of FOFs.[25] The HFR database contains 224 FOFs, and the TASS database contains 322. Of the three major database vendors, two regularly report FOF composite performance—for HFR, a fund-of-funds index (FOFHFR), and for MAR, a fund-of-funds benchmark (FOFMAR). FOFHFR is an equally weighted index based on 112 FOFs in 1994 and 224 FOFs in 1999. FOFMAR is a historical series of the median FOF's returns. For completeness in this analysis, we constructed an equally weighted index for the FOFs in the TASS database (FOFTASS).

What do the FOF data tell us about the experience of hedge-fund investing over the last five years? Following the format of the earlier sections, we begin by discussing potential measurement biases in the FOF returns.

Biases in FOF Returns. The returns from FOFs are less susceptible to measurement bias than returns to the commercial hedge-fund indexes. The track records of FOFs avoid many of the idiosyncratic biases that are embedded in pro forma returns based on individual hedge funds extracted from databases. First, the majority of FOFs make audited performance reports to their investors that include investments in successful funds as well as "mistakes," so a successful investment in a hedge fund that reached capacity constraint and stopped reporting to database vendors will remain in the history

[25] Construction of a passively diversified portfolio of hedge funds is not a practical proposition for individual investors for many reasons. For instance, the minimal investment in a single hedge fund runs from $100,000 for small funds to several million dollars for the bigger macro-funds. Unless different hedge funds are efficiently blended together, Figure 1 tells us that more than 100 funds may be needed to passively reach the limit of diversification. Even if we assume a modest minimum investment of $1 million per fund, a substantial amount of capital would be required to passively diversify away idiosyncratic hedge-fund risk. In addition, the investors would face the daunting task of administering such a large portfolio of essentially private investment vehicles. Therefore, in contrast to mutual funds, where passive diversification is available from "low-cost" indexed funds, the reverse is true for hedge-fund investing. FOFs offer investors a simple way of accessing a diversified portfolio of hedge funds.

of the FOF. This investment experience will continue to be a part of the FOF's performance as long as the FOF stays invested in that hedge fund. Past investments in funds that ceased operation will also remain in the track record of the FOF. Consequently, the actual track record of a FOF has no survivorship bias. In addition, selection bias is not relevant. An individual hedge fund may choose not to participate in a database, but its return is fully embedded in the performance of any FOF that invests in it.[26] When a FOF adds a hedge fund to its portfolio, the portfolio's history is not affected, so the issue of instant history bias does not arise.

Finally, as noted in Fung and Hsieh (2000b), survivorship bias in FOF returns is less severe than in individual hedge funds. The reason is that FOFs, through the natural process of diversification, inadvertently minimize the measurement errors that may arise.[27]

Because the returns of each individual FOF reflect the actual decisions of the FOF manager, we can compare the return behavior of the FOFHFR and the HFRI during extreme market conditions to see whether the spurious bias generated by the weighting scheme of the HFRI exists in the FOFHFR. **Figure 2** tells the story for 1998. Prior to August 1998, the indexes moved in tandem. Then, the pattern changes. The HFRI appears to recover from the August 1998 debacle more rapidly than the FOFHFR. But this "rapid" recovery can be largely attributed to the artifact of the overall index's equal-weighting methodology, in which losers are bought and winners are sold. When applied to a diverse hedge-fund universe under extreme market conditions, such a strategy leads to unrealistic return patterns. The recovery of the FOFHFR was much more

---

[26] Even for a hedge fund that does not report performance to database vendors, some FOFs are no doubt among its investor base.

[27] There is, however, a peculiar form of self-selection bias that may occur with FOFs. Large institutional portfolios managed by FOF managers are generally kept confidential. In this case, a downward bias may occur in the recorded amount of assets managed by a FOF manager in databases. The impact of this bias on performance statistics is small because the FOF indexes are generally not weighted by assets under management. In addition, the management style of the unreported programs is likely to be similar to those programs disclosed to database vendors.

gradual, despite the fact that the FOFHFR is also an equally weighted index.[28] The implication is that few, if any, actual portfolio managers could have followed the contrarian asset allocation strategy implicit in the HFRI. In addition, effecting such a quick asset reallocation for more than 1,000 funds within the space of a month is almost impossible.

Based on this evidence, the FOF indexes apparently do not suffer from spurious biases arising from unrealistic asset allocation schemes.

Tracking Errors. In the case of the HFRI, the large number of funds (in excess of 1,000) makes replication of the index virtually impossible without significant tracking error. In the case of the CTI, which has roughly 300 funds, tracking error may still be quite large but for different reasons. First, the CTI may contain funds that are closed to new investments (and/or investors). This aspect is clearly noted in the "frequently asked questions" section on the CSFB/Tremont Web site. A second, and more interesting, reason is that the liquidity and redemption policies of hedge funds make defining an unambiguous rebalancing scheme impossible. Hedge funds frequently require advance notification—ranging from 10 days to six months—prior to redemption. Redemption intervals also vary from fund to fund between 30 days to a year. For an index to be rebalanced on the basis of "known" values of its constituents, the rebalancing date must coincide with the component hedge fund that requires the longest notification period and redemption interval. If this date turns out to be a quarter or longer, then it will not match the monthly rebalancing rule of the CTI, and more importantly, it imposes a buy-and-hold strategy on an index between rebalancing dates. This "rebalancing bias" is a natural consequence of the way the hedge-fund industry operates.

If the inherent level of such natural tracking error is high, then the index is no longer investable. Thus, a broad-based index is purely a synthetic hedge-fund portfolio whose objective is to deliver the risk–return characteristics of the

[28] We observed similar return behavior for FOFMAR and FOFTASS.

"popular bets" among hedge-fund managers. For a real world benchmark, we argue that using FOFs as building blocks for hedge-fund performance is a better alternative to using aggregated individual funds. Fewer FOFs are needed to deliver the same diversification benefit as individual hedge funds. The line "Funds of Hedge Funds" in Figure 1 is the average standard deviations of 1,000 randomly selected portfolios of FOFs, which declines as we increase the number of FOFs in the portfolio. It turns out we need a portfolio of only 33 FOFs to have a standard deviation within 10 percent of that of an equally weighted portfolio of all TASS hedge funds. In addition, FOFs are likely to have a more uniform redemption policy than individual funds. These observations suggest that portfolios of FOFs are likely to have a lower rebalancing bias than portfolios of individual funds. In addition, FOFs have smaller measurement biases than individual funds.

In interpreting the risk–return characteristics of FOFs, however, investors must make adjustments for the portfolio management costs of the FOFs. A procedure for approximating these costs can be found in Fung and Hsieh (2000b).

Finally, neither a broad-based index of FOFs nor a broad-based index of individual hedge funds can capture the diverse risk characteristics of different hedge-fund trading styles. Because FOFs reflect the actual portfolio decisions of FOF managers, however, their return characteristics undoubtedly reflect hedge-fund investors' preferences better than do portfolios of hedge funds defined by arbitrary index rules.

Conclusions

Pro forma portfolio returns from hedge-fund databases suffer from survivorship bias that is natural to the youth of the hedge-fund industry and the way it is organized. Fung and Hsieh (2000b) estimated the upward survivorship bias in hedge-fund returns to be about 3 pps a year, and analysts have no way to assess the impact of the unobserved defunct funds. In addition, the combination of the voluntary nature of hedge fund participation in a database and the varied inclusion processes of the database vendors leads to selection bias. In this case also, determining the size of the selection bias is not possible because analysts cannot compare the performance of the hedge funds in databases with the performance of unobserved hedge funds in the population. Moreover, when a hedge fund enters into a vendor's database, the fund's history is generally backfilled, which gives rise to an instant history bias estimated by Fung and Hsieh (2000b) to bias returns upward by, on average, 1.4 pps a year.

In post-1994 hedge-fund data, natural biases can be partially rectified because the difference between databases and the unobservable population data has converged over time. The extent to which survivorship bias can be mitigated depends, however, on the completeness of database vendors' collection of defunct funds information. Based on empirical studies (Ackerman et al. and Liang), we can conclude that survivorship bias continues to exist in post-1994 pro forma return series based on hedge-fund databases. Moreover, these earlier studies also reported differences in the composition of the HFR database and the TASS database. These differences support the conclusion that a spurious selection bias exists in pro forma return series from these databases.

Significant differences exist between the returns of the two broad-based hedge-fund indexes—the equally weighted HFRI and the value-weighted CTI because of the way the indexes are constructed and the different number of funds in the databases. The HFRI is an index of more than 1,000 hedge funds, whereas the CTI is an index of about 280 hedge funds. The impact of varied construction methods on returns is magnified by diversity in the trading styles and performance of the underlying hedge funds.

Broad-based indexes such as the HFRI and CTI mask the diversity of individual hedge-fund risk and return characteristics. For this reason, MAR publishes median hedge-fund returns for individual styles. A broad-based index is helpful for investors who wish exposure to the popular bets among hedge-fund managers but exposes investors to the risk of style convergence, which diminishes diversification.

A simple solution that mitigates some of these biases is an index based on the records of FOFs, which avoid many of the idiosyncratic biases in pro forma returns based on individual hedge funds extracted from databases. Through the natural process of diversification, FOFs minimize measurement errors. Selection bias is also muted because all individual hedge funds are likely to have some FOF investors irrespective of whether they report performance to any hedge-fund database. Instant history bias does not arise with FOFs, and empirical evidence suggests that indexes of FOFs do not suffer from spurious biases arising from unrealistic asset allocation schemes.

Fung and Hsieh (2000b) estimated the portfolio management costs from FOFs returns to be approximately 2 percent a year. After adjusting for these costs and for instant history bias, the broad-based indexes deliver roughly the same returns as the FOF indexes with comparable standard deviations. Investors interested in current growth trends may prefer equally weighted broad-based indexes as benchmarks because these indexes are more responsive to newer, smaller hedge funds. Investors interested in diversifying into more established, larger hedge funds would benefit from using indexes of FOFs. This distinction is similar to that between small-capitalization and large-capitalization stock indexes. For investors looking for diversifying alternatives to traditional asset classes, neither the broad-based indexes nor the subindexes of hedge-fund styles are suitable building blocks for customized indexes that reflect those investment objectives.

The authors acknowledge financial support from the Duke Global Capital Markets Center. We benefited from comments from Bruce Johnson, Simon Ruddick, Joseph Sweeney, and Howard Wohl. All remaining errors are solely the responsibility of the authors. The

views expressed are those of the authors alone and may not reflect the positions of the affiliated institutions.

References

Ackerman, C., R. McNally, and D. Ravenscraft. 1999. “The Performance of Hedge Funds: Risk, Return, and Incentive.” *Journal of Finance,* vol. 54, no. 3 (June):833–874.

Brown, S.J., W. Goetzmann, and R. Ibbotson. 1999. “Offshore Hedge Funds: Survival & Performance, 1989–95.” *Journal of Business*, vol. 72, no. 1 (January):91–118.

Brown, S.J., W. Goetzmann, R. G. Ibbotson, and S.A. Ross. 1992. “Survivorship Bias in Performance Studies.” *Review of Financial Studies*, vol. 5, no. 4 (October):553–580.

Fung, W., and D.A. Hsieh. 1997a. “Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds.” *Review of Financial Studies*, vol. 10, no. 2 (April):275–302.

———. 1997b. “Survivorship Bias Investment Style In The Returns Of CTAs.” *Journal of Portfolio Management*, vol. 24, no. 1 (Fall):30–41.

———. 1999. “A Primer for Hedge Funds.” *Journal of Empirical Finance*, vol. 6, no. 3 (September):309–331.

———. 2000a. “Measuring the Market Impact of Hedge Funds.” *Journal of Empirical Finance*, vol. 7, no. 1 (May):1–36.

———. 2000b. “Performance Characteristics of Hedge Funds and Commodity Funds: Natural vs. Spurious Biases.” *Journal of Financial and Quantitative Analysis*, vol. 35, no. 3 (September):291–307.

———. 2001. “The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers.” *Review of Financial Studies*, vol. 41, no. 2 (June):313–341.

Grinblatt, M., and S. Titman. 1989. “Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings.” *Journal of Business*, vol. 62, no. 3 (July):393–416.

Liang, B. 2000. “Hedge Funds: The Living and the Dead.” *Journal of Financial and Quantitative Analysis*, vol. 35, no. 3 (September):309–336.

Malkiel, B. 1995. “Returns from Investing in Equity Mutual Funds, 1971 to 1991.” *Journal of Finance*, vol. 50, no. 2 (June):549–572.

Park, J. 1995. *Managed Futures as an Investment Set*. Doctoral dissertation, Columbia University.

Schneeweis, T., and T. Spurgin. 1998. "Multifactor Analysis of Hedge Fund, Managed Futures, and Mutual Fund Return and Risk Characteristics." *Journal of Alternative Investments*, vol. 1, no. 2 (Fall):1–24.

Sharpe, W. 1992. "Asset Allocation: Management Style and Performance Measurement." *Journal of Portfolio Management*, vol. 18, no. 2 (Winter):7–19.

Table 1. Annual Returns of Hedge-Fund Indexes

Year	HFRI	CTI	Difference (HFRI – CTI)
1994	4.1%	–4.4%	9.5 pps
1995	21.5	21.7	–0.2
1996	21.1	22.2	–1.1
1997	16.8	25.9	–9.1
1998	2.6	–0.4	2.2
1999	30.8	23.4	7.4
Mean 1994–99	16.2	14.7	1.5

Table 2. Statistics on Subindexes of Hedge Funds

	Annual Return		Weighting Scheme in Overall Index	
Year	HFR Subindex	CT Subindex	Equal	Value
A. Emerging market hedge-fund subindexes				
1994	3.38%	12.51%	12.5%	16.3%
1995	0.69	–16.91	13.9	18.0
1996	27.14	34.50	14.5	14.7
1997	16.57	26.59	12.7	12.1
1998	–32.96	–37.66	11.8	9.5
1999	55.22	44.82	11.9	7.1
B. Global/macro hedge-fund subindexes				
1994	–4.30%	–5.72%	4.9%	38.2%
1995	29.32	30.67	4.9	34.3
1996	9.32	25.58	4.7	24.8
1997	18.82	37.11	6.0	23.8
1998	6.19	–3.64	5.4	13.7
1999	17.62	5.81	4.8	13.2

Table 3. Correlations between Hedge-Fund Indexes and Market Indexes

Market Index	HFRI	CTI
A. Correlation coefficients		
Eurodollar one-month return	0.14	0.29
GSCI	0.26	0.15
U.S. equities	0.73	0.55
Non-U.S. equities	0.66	0.42
U.S. bonds	–0.05	0.17
Non-U.S. bonds	–0.14	–0.36
U.S. dollar	0.01	0.40
Emerging market equities	0.76	0.47
High-yield U.S. bonds	0.56	0.51
B. Regression of HFRI and CTI on market indexes		
Constant	–0.02	0.02
Eurodollar one-month return	6.77*	5.14*
GSCI	0.08*	0.05
U.S. equities	0.15*	0.10
Non-U.S. equities	0.04	0.10
U.S. bonds	–0.35	0.30
Non-U.S. bonds	–0.25	0.29
U.S. dollar	–0.24	0.37
Emerging market equities	0.11*	0.06
High-yield U.S. bonds	0.51*	0.24
Adjusted R^2	0.76	0.55

* An asterick indicates statistical significance at the 5% level.

Table 4. Frequency Distribution of Adjusted R^2s for Individual Hedge Funds Regressed on Market Indexes

Adjusted R^2		
From	To	Percent of Individual Hedge Funds
–0.2	–0.1	0.9
–0.1	0.0	8.1
0.0	0.1	15.3
0.1	0.2	14.5
0.2	0.3	15.5
0.3	0.4	11.9
0.4	0.5	11.2
0.5	0.6	11.8
0.6	0.7	5.4
0.7	0.8	3.2
0.8	0.9	1.9
0.9	1.0	0.4

Table 5. Size of HFR Equity Hedge-Fund Subindex Relative to the Overall HFRI

Year	By Number of Funds	By Assets of Funds
1994	16.8%	7.1%
1995	14.7	7.4
1996	14.1	12.8
1997	15.4	10.2
1998	17.9	14.2
1999	27.3	30.4

Figure 1. Portfolio Standard Deviation and Number of Funds: Groups of Individual Hedge Funds Vs Groups of Funds of Hedge Funds

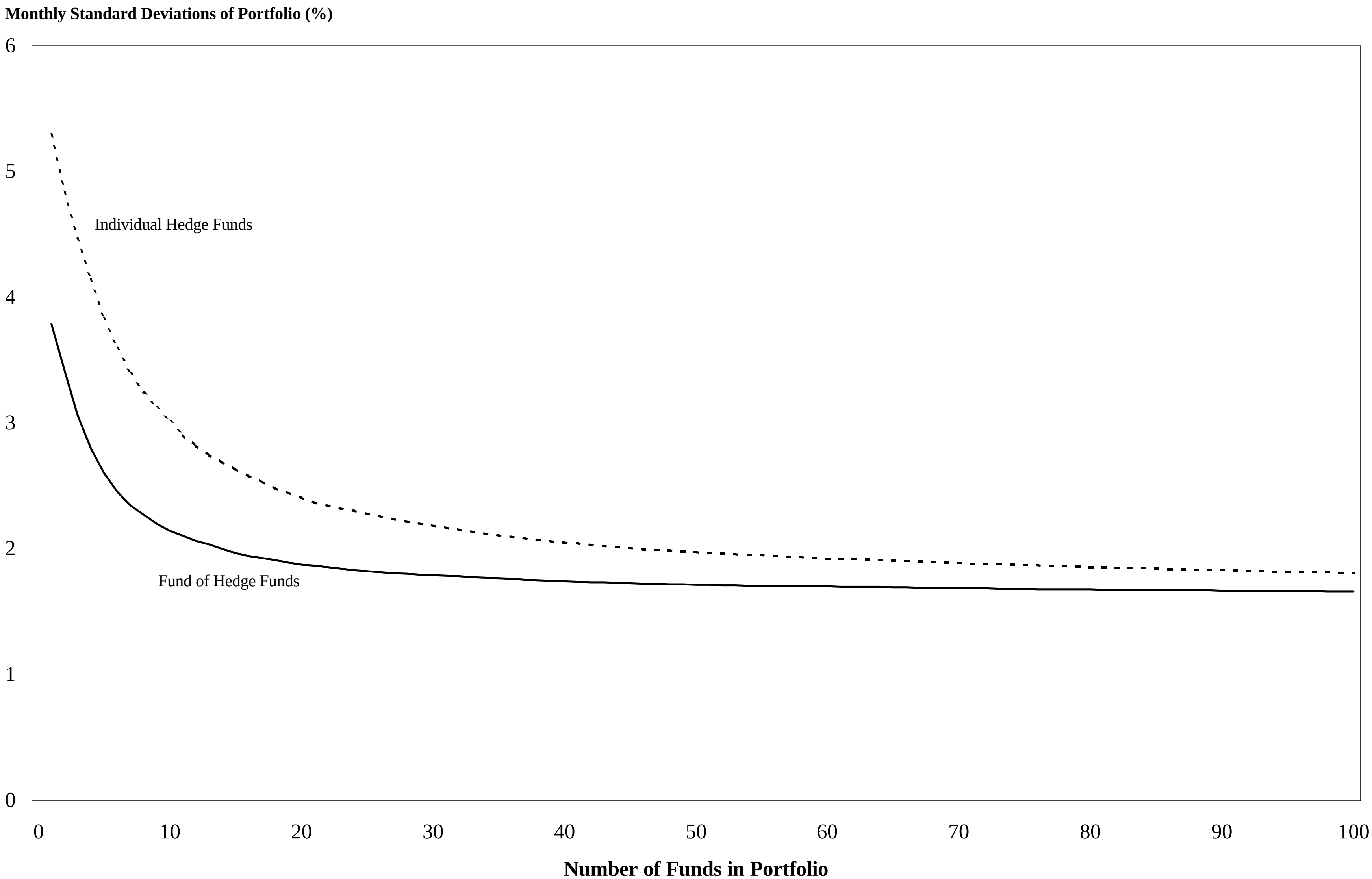

Figure 2. Monthly Returns of HFRI and FOFHFR