This chapter reviews some important innovations in quantitative strategies for investment management. We summarize the advantages and limitations of applying cointegration and principal component analysis to portfolio construction. Empirical examples show that by exploiting the presence of common factors, both models can significantly enhance traditional optimization strategies. More important, such strategies can be used successfully both for traditional investments such as equities and for optimizing more complex portfolios such as funds of hedge funds.

INTRODUCTION

In addition to offering considerable alpha opportunities, hedge funds present very attractive diversification opportunities due to their low correlation with traditional asset classes and also with their peer group. The benefits of adding hedge funds to traditional asset classes portfolios have long been illustrated through the use of hedge fund indexes as investable instruments (Schneeweis and Spurgin 1998a; Schneeweis, Karavas, and Georgiev 2002; Indjic 2002). To this end, investors have the choice of selecting one of the readily investable hedge funds indices (e.g., the ones provided by Standard & Poor’s (S&P), Hedge Fund Research (HFR), Credit Suisse First Boston (CSFB)/Tremont, or Morgan Stanley Capital International (MSCI)) or constructing a custom-made index tracking portfolio. The main advantage of the latter scenario is the possibility to control both the fund selection and the portfolio management as to include the investor’s preferences.

Hedge funds portfolio optimization is a somewhat controversial topic, with the majority of research on alternative investments still focusing on the
absolute performance portfolios (see Favre and Galeano 2002; Krokhmal, Uryasev, and Zrazhevsky 2002; Amenc and Martellini 2003c). Despite the significant enhancements brought to portfolios in traditional asset classes, Liew (2003) argues that hedge fund indexing might not be optimal, given that after adjusting for market exposure and nonsynchronous trading, the majority of hedge funds exhibit negative performance. However, hedge fund alpha estimation is prone to considerable specification and sampling errors, which makes the construction of absolute performance portfolios rather complex. As demonstrated recently (Amenc and Martellini 2003a; Alexander and Dimitriu 2004a), the absolute alpha benefit of hedge fund investing is very difficult to measure with any reasonable degree of accuracy. Due to the myriad of strategies employed by hedge funds, their highly dynamic nature, and the extensive use of derivatives and leverage inducing nonlinear relationships of their returns with the traditional asset classes, models for hedge fund returns are inherently complex (Fung and Hsieh 1997a; Schneeweis and Spurgin 1998b; Agarwal and Naik 2000c; Amenc and Martellini 2003a).

There seems to be considerable interest in hedge funds indexing models. Martellini, Vaissié, and Goltz (2004) review the issue of hedge fund indices in an attempt to reconcile investability with representativity. Their solution follows the methods of Fung and Hsieh (1997a) for constructing portfolios to replicate the principal components of hedge fund returns. However, the range of models available for indexing is considerably larger. In the equity world, many index tracking models have been developed (Roll 1992; Rudolf, Wolter, Zimmermann 1999; and Alexander and Dimitriu 2005), and practitioners now use them extensively. In the alternative investment universe, the portfolio management process has significant particularities, and no hedge fund indexation models have been studied in the extant research literature. In a realistic portfolio management setting, one should attempt to gain exposure to the hedge fund universe with a portfolio comprising a reasonably small number of funds and optimized to replicate a broad index without frequent rebalancing. There are at least two predominant considerations for any hedge fund indexation model.

1. There are significant operational restrictions to actively trading hedge funds, such as minimum investment limits, long lockup periods and advance notice, regular subscriptions and redemptions as rare as once per year, and sales and early redemption fees. To include these in the optimization model constraints on the general objective of tracking accuracy must be added.

2. The optimization model should result in a stable portfolio structure since we are aiming for a passive investment in an alternative asset class.
With these properties in mind, this chapter examines three popular indexing models that are commonly employed for standard equity strategies: the classic tracking error variance minimization model (Roll 1992), the cointegration-based index tracking model (Alexander 1999), and the common factor replication model (Alexander and Dimitriu 2004a). As the name suggests, the third model is more than an indexing model, focusing on replicating only the common trends affecting fund returns rather than tracking a traditional benchmark. If there is a strong common movement in fund returns, their first principal component will be highly correlated to traditional benchmarks, and explain them to a large extent.\(^1\) The use of principal component analysis for indexing purposes in alternative investments has been recently promoted by Amenc and Martellini (2003b), as providing the best one-dimensional summary of the information conveyed by several competing indexes. Fung and Hsieh (1997a) have examined the performance of portfolios of hedge funds constructed to replicate the first principal components of fund returns, noting the similarity between the first principal components portfolio (PC1) and a broad index.

Given the predominant considerations just mentioned, for hedge fund indexing, these models need to be implemented with constraints such as no short sales, and minimum investment limits, and the rebalancing frequency should be set to no less than six months. Our out-of-sample results show that it is possible to obtain fair replicas of hedge fund benchmarks, which preserve most of their features and comprise a relatively small number of funds. Of the models investigated, the standard tracking error variance minimization model produced a reasonable replica of the benchmark, with the lowest turnover. The cointegration portfolio was the most accurate tracker, but at the expense of additional transaction costs. Funds of funds seeking absolute outperformance of a hedge fund index would be well advised to employ a common factor replication framework. All these models produce portfolios that are highly correlated with the benchmark, and more attractive in terms of higher moments.

**MODELS FOR INDEX TRACKING**

**Classic Tracking Error Variance Minimization**

Fueled by the increased interest in indexing and the practice of evaluating managers’ performance relative to a benchmark, an extension of classical

\(^1\)In a perfectly correlated system, the first principal component will be in fact the equally weighted index of all the system components.
mean-variance analysis was made to accommodate a tracking error optimization of equity portfolios (Rudd 1980; Rudd and Rosenberg 1980; Roll 1992; Rudolf, Wolter, and Zimmermann 1999). In this setting, the problem faced by the investor is formulated in terms of expected tracking error and its volatility, rather than expected absolute return and its volatility. As emphasised by El-Hassan and Kofman (2003), the tracking error can be either the investment goal for a passive strategy that seeks to reproduce accurately a given benchmark or an investment constraint for an active strategy that seeks to outperform a benchmark while staying within given risk limits defined by the benchmark.

A general form of indexing model minimizes the tracking error variance for a given expected tracking error. (In the case of a pure index fund, this would be zero.) The intuition behind it is that a fund that meets its return target consistently has no volatility in the tracking error. The analytic solution derived by Roll (1992) is not applicable for portfolios comprising only a subset of the stocks in the benchmark. Instead, the following numerical optimization in equation 8.1 needs to be implemented:

\[
\min_{w_i} \sum_{t=1}^{T} \left( r_{index,t} - \sum_{k=1}^{N} w_k r_{k,t} \right)^2 \quad \text{s.t.} \\
\sum_{t=1}^{T} \left( r_{index,t} - \sum_{k=1}^{N} w_k r_{k,t} \right) = 0 \\
\sum_{k=1}^{N} w_k = 1 \\
L < w_k < U
\]

where \( r_{index,t} \) = the return on the index  \\
r_{k,t} = the return on the \( k \)th fund, both measured at time \( t \)  \\
w_k = the weight assigned to the \( k \)th fund

The model minimizes the variance of the tracking error subject to the constraint of zero expected tracking error, and unit sum of weights. Given the specifics of hedge funds investing, we need to impose a positive lower bound \( L \) on portfolio weights. The upper bound \( U \) is needed for diversification so that the tracking portfolio is not concentrated in just a few hedge funds.

\[\text{The tracking error is defined as the difference between portfolio returns and benchmark returns.}\]
Optimization models based on tracking error are known to have some drawbacks, which limit their applicability to a passive investment framework. One is that the attempt to minimize the in-sample tracking error with respect to an index that, as a linear combination of stock prices, comprises a significant amount of noise may result in large out-of-sample tracking errors. This is a result of the well-known trade-off between the in-sample fit and the out-of-sample performance of a model. An optimization based on tracking error will attempt to overfit the data in-sample, but this is done at the expense of additional out-of-sample tracking error. Moreover, the in-sample overfitting may result in an unstable portfolio structure that is unsuited to passive investments as it implies frequent rebalancing and significant transaction costs.

Classic tracking error models are optimized using a covariance matrix of asset or risk factor returns; these models have additional weaknesses generated by the very nature of correlation as a measure of dependency: It is a short-term statistic, which lacks stability; it is only applicable to stationary variables, such as asset returns, which requires prior detrending of level variables and has the disadvantage of losing valuable information (the common trends in level variables); and its estimation is very sensitive to the presence of outliers, nonstationarity, or volatility clustering, which limit the use of a long data history. All these exacerbate the general problems created by optimization and small sample overfitting.

We therefore examine two other models that are specifically designed to suit a passive investment framework: a variant of the cointegration-based index tracking (Alexander 1999) and the common factors replication (Alexander and Dimitriu 2004b). Both models have been shown to produce, in the equity universe, stable portfolios having strong relationships with either the benchmark itself or with only one of its components (i.e., the common factors affecting stock returns). Their enhanced stability results in a low amount of rebalancing and, consequently, reduced transaction costs.

### Cointegration-Based Index Tracking

The most general form of cointegration model allows the replication of any type of index. The rationale for constructing portfolios based on a cointegration relationship with the market index rests on two features of cointegration:

1. The value difference between the index and the portfolio is, by construction, stationary, and this implies that the tracking portfolio will be tied to the benchmark in the long run.
2. The portfolio weights are based on the history of prices rather than returns, and as a result they have an enhanced stability.
As pointed out in the introduction, the issue of transaction costs is central for passive investments. Along with the absolute tracking error and its variance as performance criteria, the amount of transaction costs incurred in managing the tracking portfolio also plays an important role. In these circumstances it seems sensible to include a proxy for the transaction costs in the optimisation model.

The most general form of the cointegration model for index tracking (Alexander 1999) is to minimize the variance of the log price spread between the tracking portfolio and the benchmark, subject to zero sum of price spreads (equivalent to zero in-sample tracking error), unit sum of weights that also lie within a lower bound and an upper bound, and, finally, stationary series of price spreads. Instead of minimizing the variance of the price spread, an alternative objective function minimizes the number of trades required to adjust the portfolio weights from one period to another, subject to the same set of constraints, of which the most important is the cointegration with the benchmark. Starting with an initial tracking portfolio cointegrated with the benchmark, the model identifies the new portfolio structure that is closest to the current one, thus involving a minimum number of rebalancing trades, and that preserves the feature of cointegration with the benchmark.

The new optimization problem can be written as:

$$\min_{w_{k,t}} \sum_{k=1}^{N} |w_{k,t} - w_{k,t-1}| P_{k,t} \text{ s.t.}$$

$$\sum_{k=1}^{N} w_{k,t} = 1$$

$$L < w_{k,t} < U$$

$$\text{ADF(ln(index))} - \sum_{k=1}^{N} w_{k,t} \ln(P_{k,t}) < \text{critical value}$$

where \(\text{index}_{t}\) = the value of the index,

\(P_{k,t}\) = the value of the \(k\)th fund, both measured at time \(t\)

\(w_{k}\) = the weight assigned to the \(k\)th fund

**Common Factors Replication**

The third indexing model investigated is a general portfolio construction model based on principal component analysis. From all possible portfolios containing all the assets in the benchmark and subject to the unit norm constraint on the weights, this model identifies the portfolio that accounts for
the largest amount of the total joint variation of the asset returns. Such a property makes it the optimal portfolio for capturing only the common factors driving asset returns, thus filtering out a significant amount of variation that can be ascribed to noise.

The $i$th principal component, where $i = 1, \ldots, k$, may be written as:

$$P_i = w_{i1}r_1 + w_{i2}r_2 + \ldots + w_{ik}r_k$$

where $r_1, \ldots, r_k$ = the returns on the hedge funds in the portfolio

$(w_{i1}, \ldots, w_{ki})'$ = the $i$th eigenvector of the returns covariance matrix.

In Alexander and Dimitriu (2004b) the portfolio replicating the first principal component is constructed directly from the normalized eigenvectors of the covariance matrix of asset returns. However, the sampling of the model in the hedge funds universe is an essential feature to preserve, so we will use the first eigenvector as a selection criterion. The higher the loading of a fund on the first principal component, the higher will be its contribution to the common factor. Given that the first eigenvector is determined so as to maximize the variance of the corresponding linear combination of fund returns, high factor loadings will be allocated to funds that have been highly correlated with their group over the calibration period. Such funds should be the most representative in their groups. Having selected the funds according to their loading on the first principal component, the portfolio is optimized to have maximum correlation with that principal component, subject to the usual constraints: nonnegativity and an upper bound of 15 percent on individual weights. The optimization problem can be written as:

$$\max (w_{k1}P_{k1}, PC1) \text{ s.t.}$$

$$\sum_{k=1}^{N} w_{kj} = 1$$

$$L < w_{kj} < U$$

Hedge Fund Data and Backtesting Procedure

Hedge fund data are subject to several measurement biases caused by the data collection process and by the nature of the industry: survivorship bias, when a database does not include the performance of funds that ceased operating during the sample period; selection or self-reporting bias, when the hedge funds in the database are not representative of the population of hedge funds; instant history bias, when the funds entering the database are
allowed to backfill their results; and multiperiod sampling bias, when the analysis is restricted to funds having a minimum amount of history available. Fung and Hsieh (2000c) provide an extensive analysis of biases in the TASS hedge fund database. They estimate a survivorship bias of approximately 3 percent per annum. Regarding the instant history bias, they found an average incubation (backfilled) period of one year with an associated bias of 1.4 percent per annum while the multiperiod sampling bias was negligible.

Our fund data comes from the Hedge Fund Research (HFR) dead and alive funds databases, from which we select the period December 1992 to May 2003. We restrict our analysis to U.S.-domiciled funds reporting net of all fees in U.S. dollars, having funds under management above $10 million, and not using leverage. Additionally, to minimize the sample bias of alpha estimates, we require that each fund has at least five years of reporting history. After imposing these selection criteria, our database comprises 282 funds.

To determine the impact of the instant history bias in our database, for each fund we examine the difference between the monthly average of the excess return (over Standard & Poor’s [S&P] 500 index) in the first year and the monthly average of the excess return in the first five years. The difference is equivalent to 3.97 percent, and the standard deviation of the difference is 1.01 percent per annum so there is a clear first-year bias in the reported fund performance. In order to eliminate the instant history bias on alpha, we use dummy variables for the first year of reporting in all factor models. The estimated multiperiod bias is negligible, at .33 percent per annum. Selection and survivorship biases are addressed by including dead funds that have sufficiently long reporting history. But this is still not sufficient to ensure that the portfolio performance is identical to the experience of an investor in these funds, because there is no information on the performance of individual funds after having ceased reporting. Statistics show that some funds stopped reporting to HFR because of extraordinarily good performance, but some also because of negative performance. If some funds were liquidated, their investors probably recovered only part of the net asset value last reported. To deal with all these potential biases, we construct an equally weighted index of all funds in our selected database.³ This

³A more realistic alternative would be to construct a value-weighted index of all funds. However, the net asset value data are missing for some funds and are discontinued for some others. In order to preserve the number of funds in the selected database, we can only construct the index based on an equal-weighting scheme. Still, as demonstrated by Larsen and Resnick (1998), equally weighted indexes are the most difficult to replicate, and our results can therefore be interpreted as minimal for the case of more commonly value-weighted indexes of hedge funds.
will be affected by the same biases. An indexing model needs to be evaluated on a relative basis; with this equally weighted benchmark its performance measurement is bias free, as both the tracking portfolios and their benchmark are affected by the same biases.

In order to test the out-of-sample performance of these models, we use a rolling sample of 60 months prior to the portfolio construction moment for calibration purposes. The first index tracking portfolios are set up in December 1997 and left unmanaged for the next six months as this is the typical lock-in period for hedge fund investments. The portfolios are then rebalanced every six months, reselecting funds based on the relevant criterion and optimizing them according to the indexing model used. We impose a nonnegativity constraint and a 15 percent upper bound constraint on portfolio weights.4

OUT-OF-SAMPLE PERFORMANCE ANALYSIS

In order to construct realistic hedge fund portfolios, we need to restrict the number of funds selected. Considering the evidence of maximum diversification benefits with around 30 funds, we use a relevant selection criterion to pick, at each rebalancing moment, approximately 30 funds that are the most likely to support the index tracking objective. Given that several fund allocations will not satisfy the lower bound constraint, the indexing portfolios will contain less than 30 funds. In fact, our indexing portfolios generally contain no more than 10 percent of the total number of funds in the benchmark. Figures 8.1 to 8.4 plot the evolution of each portfolio weights over the sample period. In general, the index tracking portfolios invest in no more than 10 percent of the funds in the universe. Note that the moments when the portfolio structure changes significantly are not the same for the four models.

Tracking Error Variance Minimization Model (TEV)

Since the objective of this model is to minimize the tracking error variance (TEV), a natural candidate for a selection criterion is the correlation of the fund returns with the equally weighted index returns. That is, at each reba-

---

4In order to accommodate the minimum investment constraint in our hedge funds portfolio, we eliminate all portfolio holdings of less than 0.5 percent and renormalize the remaining ones. Given the fact that this applies to very small holdings, the differences between the optimized portfolio and the renormalised version are minimal.
ancing moment, we select the funds that had the highest correlation with the index returns over the calibration period.

The out-of-sample performance results are summarized in Table 8.1. The TEV portfolio is very highly correlated with the equally weighted portfolio of all funds (correlation coefficient 0.94), but it has a slightly lower
information ratio (1.26 as compared to 1.48), mostly due to higher volatility. This comes as no surprise, given that the number of funds included in the TEV portfolio is less than 10 percent of the total number of funds, and because they are selected to have high correlation with the index they also have high correlation with each other. Apart from volatility, in terms of higher moments, the TEV portfolio is very similar to the equally weighted benchmark.

Figure 8.5 plots the evolution of a $100 investment in this portfolio, alongside an investment in the equally weighted index. The TEV portfolio outperforms the equally weighted benchmark until mid-2001. Since then

**FIGURE 8.3** Weights in ADF Cointegration Portfolio

**FIGURE 8.4** Weights on Correlation Selected Cointegration Portfolio
it has underperformed, but over the five-year horizon it still remains above the benchmark.

### Cointegration-Based Index Tracking Portfolio

As an alternative to the TEV model for constructing index-tracking hedge fund portfolios, we implement the cointegration tracking model described in the previous section. Under the constraint of a cointegration relationship between the value of the portfolio and the value of the benchmark, the opti-

### TABLE 8.1 Performance of Portfolios Designed to Replicate the Equally Weighted Portfolio of All Funds (Equally Weighted Portfolio of All Funds Included for Comparison)

<table>
<thead>
<tr>
<th></th>
<th>TEV Portfolio</th>
<th>ADF Cointegration Portfolio</th>
<th>Correlation Selected Cointegration Portfolio</th>
<th>Common Factors Portfolio</th>
<th>Equally Weighted Portfolio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual returns</td>
<td>9.72</td>
<td>10.24</td>
<td>9.34</td>
<td>12.10</td>
<td>10.16</td>
</tr>
<tr>
<td>Annual volatility</td>
<td>7.73</td>
<td>7.46</td>
<td>12.07</td>
<td>11.85</td>
<td>6.85</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.32</td>
<td>-0.38</td>
<td>0.04</td>
<td>-0.21</td>
<td>-0.18</td>
</tr>
<tr>
<td>Excess kurtosis</td>
<td>2.99</td>
<td>1.86</td>
<td>1.01</td>
<td>2.08</td>
<td>2.10</td>
</tr>
<tr>
<td>Information ratio</td>
<td>1.26</td>
<td>1.37</td>
<td>0.77</td>
<td>1.02</td>
<td>1.48</td>
</tr>
<tr>
<td>Turnover</td>
<td>5.83</td>
<td>10.74</td>
<td>8.91</td>
<td>7.10</td>
<td>4.92</td>
</tr>
<tr>
<td>Correlation EW</td>
<td>0.94</td>
<td>0.88</td>
<td>0.94</td>
<td>0.94</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*The information ratio is computed as average annual portfolio returns divided by annualized returns volatility.*
mization target now focuses on the stability of the portfolio structure, rather than the correlation with the benchmark. The weight constraints are the same as for the TEV model.

In order to control the number of funds in the tracking portfolio, we implement a selection criterion that is consistent with the cointegration constraint. We select the funds according to the degree of cointegration of their cumulative returns with the cumulative returns of the equally weighted index. We note that it is by no means necessary for each individual fund to be cointegrated with the benchmark in order to be able to find a linear combination of them that is cointegrated with the benchmark. However, selecting funds that are individually cointegrated with the benchmark is likely to facilitate the task of finding portfolios that are cointegrated with the benchmark.

As a proxy for the degree of cointegration, we use the augmented Dickey-Fuller (ADF) statistic for the residuals of univariate regressions of index cumulative returns on each individual fund cumulative returns. At each rebalancing moment, we select the 33 funds having the lowest ADFs (highest degree of cointegration with the benchmark) and optimize the portfolio so as to generate the minimum amount of trades subject to the cointegration constraint on the portfolio-index values relationship.

The out-of-sample performance summary is presented in Table 8.1 In terms of both returns and volatility, the cointegrated tracking portfolio outperforms the TEV model, being only slightly more volatile than the equally weighted portfolio of all funds. Again, this performance is remarkable, because it is achieved with only 10 to 15 percent of the funds in the benchmark universe. As expected, the correlation with the equally weighted benchmark is lower than in the case of the TEV model, but remains very high (0.88). One respect in which the cointegrated portfolio is less attractive is turnover, which at 10.7 is much higher than in the TEV model. Thus, the additional feature of cointegration appears to be costly particularly in terms of transaction costs. This can be due to three reasons:

1. The cointegration constraint is very strong, and it requires significant changes to the portfolio structure from one period to another.
2. The ADF fund selection criterion does not result in a stable group of funds to be further optimized under the cointegration constraint.
3. The ADF fund selection criterion is not consistent with the cointegration constraint, which is not well supported by the data on the selected funds.

The last alternative is the most unlikely.

In order to identify the cause of the portfolio instability, we have replaced the ADF fund selection criterion with the correlation criterion used in the TEV model. Even though this criterion is less consistent with the
cointegration constraint than the ADF criteria, it produced a relatively stable portfolio structure for the TEV model. When implementing the correlation-based selection criterion in conjunction with the cointegration model, we find that the turnover is significantly reduced (8.9, as compared to 10.7 for the ADF selection). As expected, the correlation with the benchmark is very high (0.94). However, the drawback is the increase in volatility, which results in a lower information ratio than previously (0.77, as compared to 1.37 for the ADF selection and 1.26 for the TEV model).

Clearly, the increased instability of the ADF based cointegration model has to do with the features of the hedge fund returns. The average of the lowest 33 ADF statistics computed on the residuals of univariate regressions of the benchmark cumulative returns on the individual fund cumulative returns is approximately 2.7, which is well above the critical value for the Engle-Granger test for cointegration (usually less than −4, depending on the number of funds in the portfolio and the length of the data sample). Despite the fact that at each rebalancing moment we manage to find a portfolio of hedge funds that is cointegrated with the equally weighted index, the stochastic common factors driving fund returns are weaker than in the case of equities, and cointegration relationships are more difficult to find.

Of all the portfolios analyzed, the ADF cointegration portfolio was the most accurate tracker of the index for the largest part of the sample period, (until September 2001). Over the next six months it underperformed the benchmark, and from mid-2002 to the end of the sample it outperformed it. Overall, it matched the return over five years of the equally weighted index of all funds. Note that from mid-2002 to the end of the sample, the performance of the ADF and the TEV portfolios was not synchronized. Hence one may consider combining them to produce a more accurate replica of the benchmark and to reduce the portfolio volatility.

Common Factors Replication Portfolios

As expected, the portfolio replicating the first principal component is much more volatile than the equally weighted portfolio of all funds. However, its returns are also significantly enhanced to 12 percent per annum, as compared to 10 percent for the benchmark. Overall, the information ratio of this portfolio is less than the one of the benchmark, but their correlation remains very high (0.94). The turnover of this portfolio is between those of the TEV model and the ADF models. In terms of higher moments, the Principal Component (PC1) portfolio is the most similar to the benchmark of all three portfolios.

The common factors portfolio clearly outperformed the equally weighted index of all funds until February 2000. Given that during the five
years prior to this moment technology funds were very popular front runners and were highly correlated as a group, the portfolio replicating the first principal component was overweighting them relative to the benchmark, and hence it made significant gains. When the technology bubble burst, this portfolio became much more volatile than the benchmark, despite the fact that in terms of returns, the difference between them did not changed dramatically.

**CONCLUSION**

Despite the modeling complexity caused by biases present in data, noisy correlation structure, alphas that are difficult to estimate, and institutional limitations to trading, alternative investments represent attractive opportunities. The diversification potential of investments in hedge funds has long been advocated, and the next natural step is to look for models to replicate the performance of hedge funds indices. Such models can be a valuable tool for managing funds of hedge funds portfolios. Aiming to develop such a fund selection and optimal allocation process for funds of hedge funds, we have analyzed the out-of-sample performance of a number of index-tracking models that were originally designed for equity portfolios and have adjusted them to fit the special features of alternative investments.

We have shown that that it is possible to obtain fair replicas of hedge fund benchmarks that preserve most of their features and comprise only a small number of funds. Each of the models investigated appears to suit a different investment profile: The TEV portfolio generates a reasonable performance associated with low turnover; the ADF cointegration portfolio represents a more accurate replica of the benchmark, but during turbulent periods of regime changes its turnover can be high; finally, the PC1 portfolio represents the model of choice for investors aiming at enhancing index returns while keeping a high correlation with the benchmark and a reasonable turnover.