Equity indexing: optimize your passive investments

Carol Alexander and *Anca Dimitriu* discuss two strategies for enhanced index tracking designed to best suit a passive investment framework.

Introduction

The phenomenon of equity indexing has recently attracted considerable interest from both academics and practitioners. Equity indexing is the most popular form of passive investment, aiming to replicate the risk and return characteristics of a benchmark, usually a wide stock market index. The passive investment industry as a whole has witnessed a remarkable growth during the last ten years, with a huge number of funds pegging their holdings to broad market indexes. The very reason for adopting a passive strategy rests in a belief in market efficiency, which provides the theoretical foundation of indexing. Traditional capital market theory states that the market portfolio offers the highest level of return per unit of risk and the only way investors can beat the market is by taking greater risks. Additionally, active management has been shown to under-perform compared to its passive alternative due to transaction costs and administration fees, mostly in bull, but also in bear markets. For example, the S&P active/passive scorecard for the last quarter of 2003 shows that the majority of active funds have failed to beat their relevant index even in the bear market of the last few years.

In these circumstances, passive investment and, in particular, indexing, are natural choices. However, despite their increased popularity, the number of models available for equity indexing and their level of sophistication are rather limited.

Traditionally, equity indexing has targeted price- and capitalization-weighted indexes. These indexes can be easily replicated with portfolios comprising the entire set of stocks and mirroring the benchmark weights, as long as there are no changes in the index composition or in the number of shares in each issue. Such portfolios are self-adjusting to changes in the stock prices and do not require any rebalancing, except for special events like mergers, additions or deletions, splits or dividends. Despite the self-replication advantage, holding all the stocks in the benchmark may not always be desirable or possible because of difficulties in purchasing odd lots to exactly match the market weights and increased transaction costs/market impact related to trading less liquid stocks. Also, the higher the number of stocks in a particular index, the more significant will be the set-up, maintenance and disinvestment costs generated by a portfolio replicating that index.

When fewer stocks are held in the replicating portfolio, their weights are no longer self-adjusting and require periodic rebalancing. More involved strategies are also required for tracking equally-weighted indexes, since frequent rebalancing is required in order to maintain equal dollar amounts in each stock.

The search for methods to optimally replicate a benchmark has focused on one of the features of an indexing strategy, i.e. the tracking error as a measure of the tracking accuracy. The tracking error is usually defined as the root-mean-square deviation of the tracking portfolio returns from the benchmark returns. When the optimization problem for an indexing strategy is formulated in terms of tracking error, the usual objective is its minimization. Additional constraints/alternative formulations of the objective function may concern the correlation of the portfolio returns with the benchmark returns or the transaction costs involved in rebalancing the portfolio. Occasionally, the optimization is supplemented or even replaced by basic capitalization and stratification considerations. Reviews of the indexing models available can be found in Rudd (1980), Meade and Salkin (1989) and, most recently, in Larsen and Resnick (1998, 2001).

However, optimization models based on tracking error or on correlation measures have significant drawbacks, which limit their applicability to a passive investment framework. First, the attempt to minimize the in-sample tracking error with respect to an index which, 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 over-fit the data in-sample, but this is done at the expense of an additional out-of-sample tracking error. Moreover, the in-sample over-fitting will result in a very unstable portfolio structure, which implies frequent rebalancing and significant transaction costs.

> Traditionally optimization models for passive investments are constructed on tracking error and correlation measures.

Apart from the problems discussed above, an optimization based on correlation has additional weaknesses which are generated by the very nature of correlation as a measure of dependence. It is a short-term statistic, which lacks stability; it is only applicable to stationary variables, such as stock returns, which requires prior de-trending of level variables (i.e. stock prices) and has the disadvantage of losing valuable information (i.e. the common trends in prices); and its estimation is very sensitive to the presence of outliers, non-stationarity or volatility clustering, which limit the use of a long data history. All of these exacerbate the general problems created by optimization and small sample over-fitting.

These limitations are well known and are usually dealt with, in an active management setting, through fine-tuning of model parameters such as the length and quality of the data used to calibrate the portfolio, the choice of optimization target, implementation of filtered rebalancing, etc. However, stability in the portfolio structure and transaction costs are central issues for passive investment and can only be dealt with by rethinking the optimization model to accommodate the objectives and limitations of passive investment.

To this end, we proposed two models designed to best suit a passive investment framework: cointegration-based index tracking (Alexander 1999) and common-trend replication (Alexander and Dimitriu 2003b). Both models produce stable portfolios that have strong relationships with either the benchmark itself, or with only one of its components, i.e. the common trend of the stocks included in the benchmark. Their enhanced stability results in a low level of rebalancing and, consequently, reduced transaction costs. Moreover, we find that, while being highly correlated with their benchmarks, the replicating portfolios succeed in over-performing them in given market circumstances. By constructing well-diversified portfolios that have a stable relationship with the market index and exploit certain market phenomena, these strategies qualify eventually as enhanced index-tracking strategies. In the following, we will provide a brief overview of these two models.

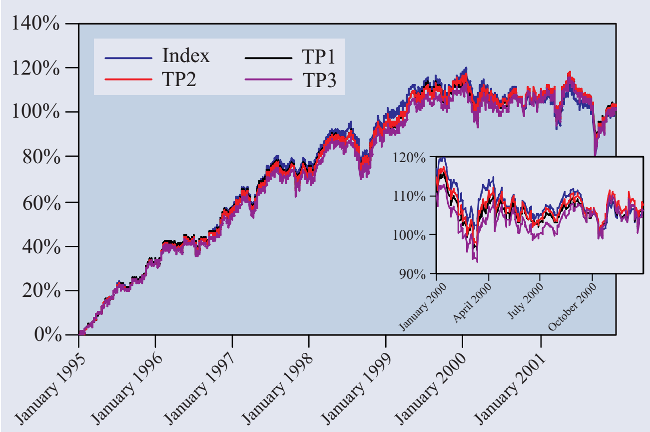

Figure 1. Cumulative returns on DJIA and tracking portfolios based on 25 stocks.

Cointegration-based index tracking

In the last decade, cointegration has been widely applied in financial econometrics in connection with time-series analysis and macroeconomics. It has evolved as an extremely powerful statistical technique because it allows the application of simple estimation methods (such as least-squares regression and maximum likelihood) to non-stationary variables. However, its relevance to investment analysis has been rather limited so far, due to the fact that the standard in portfolio management and risk measurement is the correlation analysis of asset returns, despite its well-known limitations. The fundamental remark justifying the application of cointegration to the analysis of stock prices is that a system of stock prices can share common stochastic trends (Stock and Watson 1988). According to Beveridge and Nelson (1981), a variable has a stochastic trend and is integrated of order one, if its first difference has a stationary invertible ARMA(p, q) representation plus a deterministic component. Since such models seem to characterize many financial variables, in particular stock prices and stock market indexes, it follows that these variables can be described by stochastic trends and, in this case, cointegration exists when there is at least one stationary linear combination of their prices.

A stationary linear combination of several stock prices and a market index indicates that the spread, or price differential between the index and a portfolio of those stocks, is mean reverting. This finding does not provide any information for forecasting the individual prices in the system, or the position of the system at some point in the future, but it does provide the valuable information that, irrespective of its position, the prices in the system will remain linked on a long-run basis.

The most general form of the cointegration-based index-tracking model allows the replication of all types of indexes, with different numbers of stocks. The rationale for constructing portfolios based on a cointegration relationship with the market index rests on the following features of cointegration: first, the price 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; and second, the stock weights, being based on the history of prices rather than returns, have an enhanced stability. These enhanced characteristics are a result of making full use of the information contained in level variables such as stock prices. Moreover, cointegration relationships between the market index and portfolios comprising all or only part of their stocks should be easy to find since market indexes, either equally weighted or capitalization weighted, are just linear combinations of stock prices.

Alexander and Dimitriu (2002) present an exhaustive analysis of the performance characteristics of cointegration optimal tracking portfolios, together with various other statistical arbitrage strategies derived from them. When applied to constructing trading strategies in the Dow Jones Industrial Average (DJIA), the cointegration technique produced encouraging results, out of sample and after transaction costs. As shown by figure 1, the tracking portfolios have a strong relationship with their benchmark. Their enhanced stability results in a low amount of rebalancing and, consequently, reduced transaction costs. An interesting finding is that the tracking portfolio comprising the same stocks as the market index produces positive abnormal returns in certain market conditions, which are quite substantial even after accounting for transaction costs. Moreover, the periods during which most of the abnormal return is accumulated coincide with the main market crises during the sample period: the Asian crisis, the Russian crisis, the technology market crash and September 11th.

Since the only information used to construct the tracking strategy is the history of the stock prices, the cause of

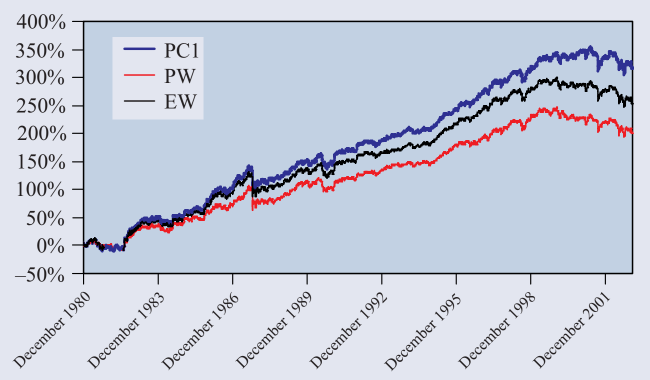

Figure 2. Cumulative performance of the common-trend replication portfolio in the DJIA universe versus a price-weighted benchmark and an equally-weighted benchmark comprising the same stocks.

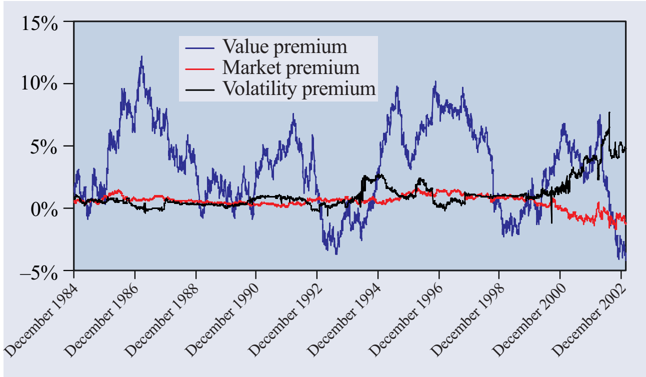

Figure 3. Time distribution of the over-performance sources.

the abnormal return should be linked to the time-variability characteristics of the stock prices in the system. Alexander and Dimitriu (2003a) introduce a new measure of the cohesion of stock prices within the market index, called 'index dispersion', and find that this is a leading indicator of the abnormal return, where their relationship is based on a switching process of two market regimes. The entire abnormal return is shown to be associated with only one of the regimes, which has been the prevalent regime during the last few years.

The cointegration relationship specified between the portfolio and a price-weighted benchmark can be interpreted as a relative pricing model. For as long as stock prices are oscillating around the past equilibrium levels, the strategy generates accurate replicas of the benchmark. But in volatile markets, where returns are low and prices are moving towards new levels, the strategy produces consistent excess returns. We argue that the reason why the cointegration strategy has periods when it significantly over-performs its benchmark but no periods of significant under-performance, is the asymmetric behaviour of stock prices, the fact that prices tend to fall faster than they rise. The cointegration portfolio, being based on a historical price equilibrium, exploits general stock-market declines and recovery periods (as was the case, for example, during the recent 'boom and bust' in technology stocks) even though it is not specifically designed for this purpose.

Since the over-performance occurs during especially volatile periods, one cannot exclude the presence of a risk factor—related to the presence of speculative bubbles—for which the abnormal return of the cointegration strategy represents a risk premium. Even if such a risk factor does not exist, the anomalies identified by cointegration are temporary and occur only in special market circumstances.

Additionally, the cointegration index-tracking strategy can be easily extended into active management by setting up portfolios to track artificial indexes, such as index plus and index minus, and trading on their spread. Alexander and Dimitriu (2002) show that the statistical arbitrage portfolios yield returns according to the spread between the benchmarks tracked and have lower volatility than the market, low correlation with it, and near to normal-return distributions. Their performance depends on the number of stocks in the portfolio, the stock selection method, the spread between the benchmarks tracked and the calibration period. By using 'fund of funds' and 'portable alpha' techniques, we also show how the characteristics of the cointegration index-tracking and statistical-arbitrage portfolios could be significantly improved by combining them to create strongly market-neutral or enhanced index-tracking portfolios.

Common trend replication

The second model, proposed by Alexander and Dimitriu (2003b), is a general portfolio-construction model based on principal component analysis. From all possible portfolios containing the stocks 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 stock returns. Such a property makes it the optimal portfolio for capturing only the common trend in a system of stocks, thus filtering out a significant amount of variation that can be ascribed to 'noise'.

The standard approach in constructing factor mimicking portfolios uses factor loadings in the stock selection process, as in Fama and French (1993) and Fung and Hsieh (1997). In these methods, principal-component analysis is used as a stock-selection technique and the portfolio construction is a separate stage, based either on a standard optimization, or on an arbitrary method such as equal weighting. We propose a different approach in which principal-component analysis is used for asset allocation rather than stock selection. A portfolio replicating the first principal component is constructed directly from the normalized eigenvectors of the variance–covariance matrix of stock returns. By construction, this portfolio captures the largest proportion of the variation in the stock returns and it is naturally suited for a passive investment framework: it requires a fully invested portfolio of all stocks, but involves a very small amount of rebalancing because it captures the major common trend in stock returns. This procedure involves a single optimization, the one producing the principal components. Moreover, there is no arbitrary choice of the allocation model, such as equal weighting of stocks.

The common-trend replication model is different from the traditional approaches to portfolio optimization in more than one respect. Firstly, it is maximizing and not minimizing portfolio variance, and this might appear counterintuitive at first glance. However, when combined with the unit norm constraint on the factor loadings, the result is a balanced portfolio with a stable structure which also explains most of the joint variance in the system of stocks. Secondly, it is not aiming at stock selection, but rather at diversifying over the entire universe of stocks. All stocks will be represented in the portfolio replicating the common trend and the portfolio will be fairly evenly balanced if there is a high level of correlation in the stock returns. Finally, despite being a passive investment model, the benchmark does not enter into the methodology anywhere. This eliminates the problems associated with using an inappropriate benchmark in the portfolio construction, but limits the relevance of traditional indexing-performance measures such as tracking error, so caution is needed when interpreting such results.

When applying the strategy in different stock markets, we find that the first principal component captures the market factor, being highly correlated with the benchmark returns. Moreover, the factor weights prove to be very stable in time, so transaction costs are minimal. However, what does come as some surprise is that, out of sample, the portfolio replicating the first principal component in the DJIA universe, while being highly correlated with its equally-weighted and price-weighted benchmarks, significantly over-performs both of them (figure 2). One cause of the over-performance is shown to be the mean reversion in returns for the group of stocks which are over-weighted by the portfolio, i.e. the stocks that have had higher volatility and have also been highly correlated as a group during the portfolio calibration period. There are two behavioural phenomena that could be driving the mean reversion for these stocks: the attention-capturing effect and investors' over-reaction, both of them resulting in different forms of herding behaviour. Indeed, there is a close relationship between the abnormal return and a measure of investors herding towards the market factor.

The strategy's over-performance can be decomposed into a market premium, a value premium and a volatility premium. The straddle pattern created by the volatility premium is particularly appealing for investors, reducing the exposure to large negative benchmark returns and increasing the exposure to positive benchmark returns. The distribution in time of the over-performance sources has been shown to evolve, from a value-dominated over-performance towards an increased volatility premium associated with the volatile years towards the end of our data sample (see figure 3). This supports the behavioural mechanisms thought to be driving the mean reversion in stock returns, to the extent that during volatile periods investors tend to herd less and this prevents mean reversion from taking place.

Moreover, these findings are not restricted to the Dow Jones index. A common pattern in strategy performance has been identified in three major stock markets, S&P100, FTSE100 and CAC40: there is high correlation between the strategy results for the two European indexes, and high correlation between the results for the two US indexes, but lower correlation between the results on the European and US indexes. The differences in the patterns of the US and European results, however small, present a potential for diversification. Extending the analysis to other stock markets, less correlated with the US and European ones, could uncover even better diversification opportunities.

To summarize and conclude, the findings related to the two indexing models have wide implications for the passive investment industry. They show that, without any stock selection or explicit timing attempts, which are attributes of active management, solely through smart optimization, the benchmark performance can be significantly enhanced, even after accounting for transaction costs. Moreover, these strategies can be applied to replicate and/or enhance any type of value- or capitalization-weighted benchmark, not only wide market indexes.

References

Alexander C 1999 Optimal hedging using cointegration *Phil. Trans. R. Soc.* A **357** 2039–58

Alexander C and Dimitriu A 2002 The cointegration alpha: enhanced index tracking and long–short market neutral strategies *ISMA Centre Discussion Paper Series in Finance* 2002-08 (available from http://www.ismacentre.rdg.ac.uk/dp)

Alexander C and Dimitriu A 2003a Abnormal returns in equity markets: evidence from a dynamic indexing strategy *ISMA Centre Discussion Paper Series in Finance* 2003-02 (available from http://www.ismacentre.rdg.ac.uk/dp)

Alexander C and Dimitriu A 2003b Sources of over-performance in equity markets: mean reversion, common trends and herding *J. Portfolio Management* at press (*ISMA Centre Discussion Paper Series in Finance* 2003-08 (available from http://www.ismacentre.rdg.ac.uk/dp))

Beveridge S and Nelson C R 1981 A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle *J. Monetary Economics* **7** 151–74

Fama E F and French K R 1993 Common risk factors in the returns on stocks and bonds *J. Financial Economics* **33** 3–56

Fung W and Hsieh D A 1997 Empirical characteristics of dynamic trading strategies: the case of hedge funds *Rev. Financial Stud.* **10** 275–302

Larsen G A and Resnick B 1998 Empirical insights on indexing *J. Portfolio Management* **25** 51–60

Larsen G A and Resnick B 2001 Parameter estimation techniques, optimization frequency, and equity portfolio return enhancement *J. Portfolio Management* **27** 27–34

Meade N and Salkin G R 1989 Index funds—construction and performance measurement *J. Oper. Res. Soc.* **40** 871–9

Rudd A 1980 Optimal selection of passive portfolios *Financial Management* **9** 57–66

Stock J H and Watson M W 1988 Testing for common trends *J. Am. Stat. Assoc.* **83** 1097–107

Carol Alexander (c.alexander@ismacentre.rdg.ac.uk) is the Professor of Risk Management and Director of Research and Anca Dimitriu (a.dimitriu@ismacentre.rdg.ac.uk) is a PhD Student at the ISMA Centre, The University of Reading, Whiteknights Park, Reading RG6 6BA, UK.

Copyright of Quantitative Finance is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.