European Journal of Operational Research 228 (2013) 273–281

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

Enhanced indexation based on second-order stochastic dominance

Diana Roman [a,b,*], Gautam Mitra [a,c,1], Victor Zverovich [a,d,2]

[a] CARISMA (The Centre for the Analysis of Risk and Optimisation Modelling Applications), School of Information Systems, Computing and Mathematics, Brunel University, London UB8 3PH, UK
[b] Consultant at Optirisk Systems, Uxbridge, UK
[c] Director of Optirisk Systems, Uxbridge, UK
[d] Software engineer at Optirisk Systems, Uxbridge, UK

ARTICLE INFO

Article history:
Received 6 July 2011
Accepted 21 January 2013
Available online 1 February 2013

Keywords:
Finance
Portfolio choice
Index tracking
Mean-risk
Stochastic dominance

ABSTRACT

Second order Stochastic Dominance (SSD) has a well recognised importance in portfolio selection, since it provides a natural interpretation of the theory of risk-averse investor behaviour. Recently, SSD-based models of portfolio choice have been proposed; these assume that a reference distribution is available and a portfolio is constructed, whose return distribution dominates the reference distribution with respect to SSD. We present an empirical study which analyses the effectiveness of such strategies in the context of enhanced indexation. Several datasets, drawn from FTSE 100, SP 500 and Nikkei 225 are investigated through portfolio rebalancing and backtesting. Three main conclusions are drawn. First, the portfolios chosen by the SSD based models consistently outperformed the indices and the traditional index trackers. Secondly, the SSD based models do not require imposition of cardinality constraints since naturally a small number of stocks are selected. Thus, they do not present the computational difficulty normally associated with index tracking models. Finally, the SSD based models are robust with respect to small changes in the scenario set and little or no rebalancing is necessary.

In this paper we present a unified framework which incorporates (a) SSD, (b) downside risk (Conditional Value-at-Risk) minimisation and (c) enhanced indexation.

1. Introduction

Second order Stochastic Dominance (SSD) has a well recognised importance in financial portfolio selection, due to its connection to the theory of risk-averse investor behaviour and tail risk minimisation. However, until recently, stochastic dominance was considered only as a theoretical tool and not as an active portfolio strategy, because the models applying this concept were regarded as intractable or at least very demanding from a computational point of view. Computationally tractable and scalable portfolio optimisation models which apply the concept of SSD were proposed recently (Dentcheva and Ruszczyński, 2006; Roman et al., 2006; Fabian et al., 2011). These portfolio optimisation models assume that a benchmark, that is, a desirable "reference" distribution is available; a portfolio is then "actively" constructed, whose return distribution dominates the reference distribution with respect to SSD.

Index tracking models are commonly referred to as "passive" asset allocation strategies. They also assume that a reference distribution (that of a financial index) is available. A portfolio is constructed with the aim of replicating, or tracking, the financial index. Traditionally, this is done by minimising the tracking error: the standard deviation of the differences between the portfolio and index returns (Roll, 1992). Other methods have been proposed (for a review of these methods, see for example Beasley et al., 2003; Canakgoz and Beasley, 2008). The "passive" portfolio strategy of index tracking is based on the well established "Efficient Market Hypothesis" (Fama, 1970) which implies that financial indices achieve the best returns over time.

A common problem with index tracking models is raised by their computational difficulty; this is due to implementing regulatory or trading constraints, e.g. cardinality constraints that limit the number of stocks in the chosen portfolio. It is known that most index tracking models naturally select a very large number of stocks in the composition of the portfolio. Cardinality constraints overcome this problem, but they require introduction of binary variables and thus the resulting model becomes more difficult to solve. Most of the literature in the field is concerned with overcoming this computational difficulty; see for example Beasley, Beasley et al. (2003), Canakgoz and Beasley (2008).

Enhanced indexation models are related to index tracking, in the sense that they also consider the return distribution of an index as a reference. They however aim to outperform the index by generating "excess" return (diBartolomeo, 2000; Scowcroft and Sefton,

* Corresponding author at: CARISMA (The Centre for the Analysis of Risk and Optimisation Modelling Applications), School of Information Systems, Computing and Mathematics, Brunel University, UB8 3PH, UK. Tel.: +44 1895 267502.

E-mail addresses: diana.roman@brunel.ac.uk (D. Roman), gautam.mitra@brunel.ac.uk (G. Mitra), viktar.zviarovich@brunel.ac.uk (V. Zverovich).

[1] Tel.: +44 1895 265186.

[2] Tel.: +44 1895 265186.

0377-2217/$ - see front matter
http://dx.doi.org/10.1016/j.ejor.2013.01.035

2003). Enhanced indexation is a very new area of research and there is no generally accepted portfolio construction method in this field (Canakgoz and Beasley, 2008). The same computational issues as in index tracking are encountered.

Although the idea of enhanced indexation was formulated as early as 2000, the (few) enhanced indexation methods were proposed later in the research community (a review in Canakgoz and Beasley, 2008). Moreover, these methods are mostly concentrated on overcoming the computational difficulty raised by restricting the cardinality of the portfolios – not on answering the question if they do attain their stated purpose, i.e. obtain return in excess of the index.

From a theoretical perspective, enhanced indexation calls for further justification. The Efficient Market Hypothesis (EMH) is based on the key assumption that security prices fully reflect all available information – see Elton and Gruber (1995) also Lo (2005) for an insightful review of this topic. However, this hypothesis has been continuously challenged; the mere fact that academicians and practitioners commonly use "active" (i.e. non-index tracking) strategies is an indication for this. An attempt to reconcile the advocates and opponents of the EMH is the "adaptive market hypothesis" (Lo, 2005). Here, the idea is that the market "adapts" to the information received and is generally efficient but there are periods of time when it is not – and thus, these periods can be used by investors to make profit in excess of the market index. This would justify, from a theoretical point of view, the quest for techniques that seek to obtain excess return as compared to financial indices. Enhanced indexation aims to discover and exploit market inefficiencies.

There are very few empirical studies comparing performance of enhanced index funds with that of their proxy indices; a review is given in Krause (2009), see also Ahmed and Nanda (2005). However, no indication is given on the *methods* used for constructing the index funds. These studies mostly come to the conclusion that, although overall the universe of enhanced funds does not seem to outperform the market, there are situations when outperformance does occur and persists for some periods of time; this seems to be in line with the adaptive market hypothesis. Another conclusion is that there may be *specific* types of funds that do add value; however, as stated before, there is no indication on how to construct these types of funds.

In this paper, we analyse the *effectiveness* of our previously proposed SSD-based portfolio models (Roman et al., 2006; Fabian et al., 2011) as enhanced indexation strategies applied to three markets: FTSE 100, Nikkei 225 and SP 500. We also investigate aspects related to the practical application of these portfolio models: cardinality control and rebalancing.

The motivation and contribution of this work are as follows. We aim to show that strategies stemming from risk-averse models of economic behaviour (SSD) can be used as a criterion to dominate (and thereby enhance) a financial index. The resulting active portfolios improve upon the passive strategy of index tracking.

In our previous papers (Roman et al., 2006; Fabian et al., 2011), we proposed asset allocation strategies and showed that the resulting return distributions could dominate a financial index from a theoretical point of view only (i.e. with respect to SSD); however, providing excess return on the index is a separate issue. The results of our study showing the realised historical performance of the chosen portfolios, measured over time and compared with the historical performance of the index, provide empirical evidence that these models achieve the stated purpose of enhanced indexation, that is, to generate excess return.

This evidence – that a (completely described) asset allocation strategy can outperform financial indices in a rather consistent manner – is somewhat different from the findings of other authors, see Ahmed and Nanda (2005). Our study fills a gap in the literature; a specified enhanced indexation strategy is applied to several markets and compared against the indices performance over an extended period of time.

It has been recently shown that very large SSD-based models can be solved in seconds, using solution methods which apply the cutting plane approach, as proposed by Fabian et al. (2011). However, imposing additional constraints that add trading realism (for example cardinality constraints, which require additional binary variables) could increase dramatically the computational time. We empirically show that SSD-based models naturally select a small number of stocks in the composition of the portfolio, thus the cardinality constraints may be eliminated.

Another aspect of interest in real-life portfolio trading is the amount of rebalancing needed; how and when does the current portfolio change when new information comes into place. Within the stochastic optimisation paradigm, this aspect is in connection with the stability of the portfolio optimisation model to changes in the input data. In this paper, we investigate the changes in the solution portfolios over time triggered by new data on the asset returns.

The rest of the paper is organised as follows. In Section 2 we introduce index tracking and enhanced indexation. In Section 3 we discuss how Second order Stochastic Dominance (SSD) is used as a choice criterion in portfolio selection. In Section 4 we formulate the proposed models for enhanced indexation based on SSD. The numerical experiments are presented in Section 5. Three datasets, drawn from FTSE 100, Nikkei 225 and SP 500 are used for backtesting the proposed models in a rebalancing frame. Conclusions are presented in Section 6.

2. Index tracking and enhanced indexation

Let n denote the number of the assets into which we may invest at the beginning of a fixed time period. A portfolio $x = (x_1,\ldots,x_n)^T$ represents the proportions of initial capital invested in the different assets. Let the random vector $R = (R_1,\ldots,R_n)^T$ denote the returns of the assets at the end of the investment period. The return of the portfolio x is denoted by $R_x = R^T x$, a random variable.

Let $X \subset \mathbb{R}^n$ denote the set of the feasible portfolios. We assume that X is a convex polyhedron; for example, in the simplest case,

$$X = \{(x_1,\ldots,x_n)/\sum_{j=1}^{n} x_j = 1,\ x_j \geqslant 0,\ \forall j \in \{1,\ldots,n\}\}$$

It is usual to assume that the future returns of the assets are discrete random variables with a finite number of outcomes, obtained by scenario generation or finite sampling of historical data (this is also the assumption used throughout this paper). Consider S scenarios and p_i the probability of scenario i, $i \in \{1,\ldots,S\}$; $\sum_{i=1}^{S} p_i = 1$. Let r_{ij} be the return of asset j under scenario i, $i \in \{1,\ldots,S\}, j \in \{1,\ldots,n\}$. Thus, the random variable representing the return of asset j is finitely distributed over $\{r_{1j},\ldots,r_{Sj}\}$ with probabilities $p_1,\ldots,p_S$. The random variable R_x representing the return of portfolio $x = (x_1,\ldots,x_n)$ is finitely distributed over $\{r_{x1},\ldots,r_{xS}\}$, where $r_{xi} = x_1 r_{i1} + \ldots + x_n r_{in}$, $\forall i \in \{1,\ldots,S\}$.

The primary problem in "active" portfolio selection is how to find a portfolio x such that its return R_x is "maximised". (Since R_x is a random variable, this requires further clarification. There are various models of choice under risk that specify a preference relation among random returns. A portfolio x is then chosen such that its return R_x is non-dominated with respect to the preference relation considered. We resume this discussion in Section 4).

Index tracking models are a somewhat special category; they are a "passive" portfolio selection strategy. Their aim is to track a financial index's return as close as possible, thus, to "minimise" the difference between R_x and the (known) return distribution R_I