ARTICLE IN PRESS

JID: EOR [m5G;October 27, 2016;15:27]

European Journal of Operational Research 000 (2016) 1–8

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

On exact and approximate stochastic dominance strategies for portfolio selection

Renato Bruni[a,*], Francesco Cesarone[b], Andrea Scozzari[c], Fabio Tardella[d]

[a] Dipartimento di Ingegneria Informatica, Automatica e Gestionale, Sapienza Università di Roma, Rome, Italy
[b] Dipartimento di Studi Aziendali, Università di Roma Tre, Rome, Italy
[c] Facoltà di Economia, Università degli Studi Niccolò Cusano - Telematica, Rome, Italy
[d] Dipartimento Metodi e Modelli per l'Economia, il Territorio e la Finanza, Sapienza Università di Roma, Rome, Italy

ARTICLE INFO

Article history:
Received 22 April 2015
Accepted 3 October 2016
Available online xxx

Keywords:
Applied probability
Stochastic dominance
Portfolio optimization
Expected shortfall
Index tracking

ABSTRACT

One recent and promising strategy for Enhanced Indexation is the selection of portfolios that stochastically dominate the benchmark. We propose here a new type of approximate stochastic dominance rule which implies other existing approximate stochastic dominance rules. We then use it to find the portfolio that approximately stochastically dominates a given benchmark with the best possible approximation. Our model is initially formulated as a Linear Program with exponentially many constraints, and then reformulated in a more compact manner so that it can be very efficiently solved in practice. This reformulation also reveals an interesting financial interpretation. We compare our approach with several exact and approximate stochastic dominance models for portfolio selection. An extensive empirical analysis on real and publicly available datasets shows very good out-of-sample performances of our model.

1. Introduction

In this work we develop portfolio optimization methods for Enhanced Indexation (EI) based on various types of Stochastic Dominance (SD) criteria, and we compare their empirical performances. References on EI can be found in, e.g., Canakgoz and Beasley (2008), Guastaroba and Speranza (2012), Bruni, Cesarone, Scozzari, and Tardella (2015). SD approaches to EI exhibit particular advantages and have an intuitive meaning in terms of Expected Utility Theory (see e.g., Levy, 1992, 2006). Furthermore, several relations between SD approaches and mean-risk optimization have been identified in the literature (see e.g., Gotoh and Konno, 2000 and references therein).

In most cases the optimization models for EI based on stochastic dominance have a large number of constraints, since a large number of conditions are needed to ensure SD. However, they can often be solved in reasonable time by taking advantage of polyhedral techniques developed in the field of Combinatorial Optimization. Ruszczyński and Vanderbei (2003) propose mean-risk models that are solvable by linear programing and generate portfolios whose returns are nondominated according to Second-order Stochastic Dominance (SSD). One of the first Enhanced Indexation models based on SD is also in Kuosmanen (2004). He derives and implements the first programs dealing with the exact First-order Stochastic Dominance (FSD) and SSD rules. Later, Luedtke (2008) describes compact linear programing formulations where the objective is to maximize the portfolio expected return with SSD constraints over the benchmark. An efficient practical approach to EI for large markets has been proposed by Fábián, Mitra, Roman, and Zverovich (2011) and Roman, Mitra, and Zviarovich (2013), who directly apply a SSD strategy to construct a portfolio whose return distribution dominates the one of a benchmark. More recently, Hodder, Jackwerth, and Kolokolova (2015) successfully apply the exact SSD methods of Kuosmanen (2004) and Kopa and Post (2015), while Longarela (2015) provides a description of the set of all SSD-efficient portfolios by means of a family of mixed-integer linear constraints. Third-order Stochastic Dominance has also been recently applied to EI by Post and Kopa (2016).

As shown by Leshno and Levy (2002), relaxations of SD may provide advantages over exact SD in several economical contexts. Hence, they propose an approximate SD rule, called *Almost Stochastic Dominance*, and they identify the corresponding classes of utility functions for the case of first and second order stochastic dominance. An oversight in their work has been corrected in Tzeng, Huang, and Shih (2013), and further generalizations and characterizations have been provided in Levy, Leshno, and Leibovitch (2010), Tzeng et al. (2013), Post and Kopa (2013), Guo, Post, Wong, and Zhu (2014), Denuit, Huang, Tzeng, and Wang (2014), and Tsetlin, Winkler, Huang, and Tzeng (2015). However, no

* Corresponding author. Fax: +39 06 77274074.
E-mail address: bruni@dis.uniroma1.it (R. Bruni).

http://dx.doi.org/10.1016/j.ejor.2016.10.006
0377-2217/© 2016 Elsevier B.V. All rights reserved.

applications of Almost Stochastic Dominance to portfolio selection seem to be available. This might be due to the difficulty of implementing Almost Stochastic Dominance rules in this setting, but also to the abundance of portfolios that typically dominate the benchmark already with standard SD rules.

Lizyayev and Ruszczynski (2012) have introduced a different relaxation of SD, which we call here *Lizyayev–Ruszczynski Almost Stochastic Dominance* (LR-ASD). In this case, the authors focus on computationally tractable conditions, and describe the optimization models for the practical implementation of first and second-order rules. They also describe potential applications of the LR-ASD rules to portfolio selection. However, they do not provide empirical results on real datasets, but only on some illustrative examples. Furthermore, also in this case one could question the advantage of a relaxed SD rule over the standard one which already guarantees an abundance of portfolios dominating the benchmark.

In contrast to the previous cases, under classical no-arbitrage assumptions, the existence of a portfolio dominating the benchmark is ruled out when using the standard Zero order (also called statewise) stochastic dominance. Thus, some kind of relaxed Zero order stochastic dominance is needed to find a portfolio dominating the benchmark. A preliminary study in this direction has been presented in Bruni, Cesarone, Scozzari, and Tardella (2012), obtaining promising empirical and computational results on some real-world datasets.

We compare here several new and known variants of exact and approximate SD models for portfolio selection, and we analyze in detail their practical performances by means of an extensive comparative evaluation. Specifically, in Section 2 we briefly describe the main exact and approximate SD rules, and we define the Zero-order ε-Stochastic Dominance (ZεSD) rule, which implies both the Almost Stochastic Dominance rule introduced by Leshno and Levy (2002) and the one introduced by Lizyayev and Ruszczynski (2012). In Section 3 we present a cumulative version (CZεSD) of ZεSD and we apply it to the EI problem. The EI model based on CZεSD requires that the cumulative performance of the selected portfolio on all subsets of past observations outperforms that of the index up to an ε tolerance. This gives rise to a very large LP model which can however be reformulated in a compact manner and solved efficiently. Such reformulation also provides an interesting financial interpretation of the CZεSD approach to EI in terms of expected shortfall. In Section 4 we present empirical results on some major real world markets showing the practical effectiveness of several SD based approaches for portfolio selection and in particular of the one based on CZεSD.

To sum up, the main contributions of this work are the definition of new types of approximate stochastic dominance rules, their relations with the existing ones, and their application and interpretation in portfolio selection problems.

2. Exact and approximate stochastic dominance relations

According to Expected Utility Theory (see e.g., von Neumann & Morgenstern, 1944), a random variable is preferred to another if it presents a larger value of the expected utility. However, this approach depends on the specification of a utility function, which is a fairly subjective matter. On the other hand, Stochastic Dominance (SD), which is strictly related to Expected Utility Theory, is able to provide a (partial) order in the space of random variables avoiding the specification of a particular utility function, and for this reason it is particularly attractive to approach portfolio selection problems.

We now briefly recall the most common Stochastic Dominance order relations. Let A and B be two random variables, with distribution functions $F_A(\alpha) = \mathbf{Pr}(A \le \alpha)$ and $F_B(\alpha) = \mathbf{Pr}(B \le \alpha)$ for $\alpha \in \mathbb{R}$.

Definition 1 (Zero-order Stochastic Dominance (ZSD))**.** A is preferred to B w.r.t. ZSD if

$$F_{A-B}(0) = \mathbf{Pr}(A - B \le 0) = 0. \tag{1}$$

In terms of the realizations a_t and b_t of A and B at time t, this means that $a_t \ge b_t$ almost everywhere.

Definition 2 (First-order Stochastic Dominance (FSD))**.** A is preferred to B w.r.t. FSD if

$$F_A(\alpha) \le F_B(\alpha) \qquad \forall \alpha \in \mathbb{R}. \tag{2}$$

Definition 3 (Second-order Stochastic Dominance (SSD))**.** A is preferred to B w.r.t. SSD if

$$\int_{-\infty}^{\alpha} F_A(\tau)d\tau \le \int_{-\infty}^{\alpha} F_B(\tau)d\tau \qquad \forall \alpha \in \mathbb{R}. \tag{3}$$

Note that, for the sake of simplicity, in the above definitions we omit the frequently added requirement for the strict inequality in at least one case. SD relations of any order ν can be defined. When increasing the order, the corresponding condition becomes less restrictive: the νth order SD implies the $(\nu+1)$th order SD, while the opposite is not necessarily true (see e.g., Levy, 2006).

The ZSD relation represents the behavior of a decision maker who prefers a random variable over another only when the first gives better outcomes than the second in (almost) all states of the world. On the other hand, higher order SD relations are less demanding and can be linked to Expected Utility Theory in terms of different classes of utility functions. Indeed, A is preferred to B w.r.t. FSD if and only if $E[u(A)] \ge E[u(B)]$ for all non-decreasing utility functions u; A is preferred to B w.r.t. SSD if and only if the same holds for all non-decreasing and concave utility functions (see e.g., Levy, 1992).

As showed, e.g., in Leshno and Levy (2002), there are cases where the above SD relations are not able to order the returns of two investments, even though most decision makers would prefer one investment over the other. Therefore, some relaxations of the above exact SD relations have been proposed in the literature with the aim of increasing their ability to establish preferences among investments. We first describe the one proposed by Leshno and Levy (2002) with the name of *Almost Stochastic Dominance*. This relationship can be specified for any order $\nu \ge 1$. With our notation, the one corresponding to the first order is:

Definition 4 (Leshno–Levy Almost First-order Stochastic Dominance (LL-AFSD))**.** Given a tolerance $\eta > 0$, A is preferred to B w.r.t. LL-AFSD if

$$\int_{S_1} (F_A(\tau) - F_B(\tau))d\tau \le \eta \int_{\alpha'}^{\alpha''} |F_A(\tau) - F_B(\tau)|d\tau, \tag{4}$$

where $[\alpha', \alpha'']$ is the combined range of outcomes of A and B, and $S_1 = \{\tau \in [\alpha', \alpha''] : F_A(\tau) > F_B(\tau)\}$.

The underlying idea is to allow an area of possible violation of the classical SD, the so-called *actual violation area*, containing preferences of investors that can be considered *economically irrelevant*, as explained in detail in Leshno and Levy (2002). This corresponds to the exclusion of "extreme" utility functions and allows to fit in the theory situations where most of the investors would prefer investment A over investment B, but neither investment dominates the other with the usual FSD or SSD rules.

Another recent relaxation of Stochastic Dominance, still defined for any order $\nu \ge 1$, is proposed by Lizyayev and Ruszczynski (2012), who also provide the optimization models corresponding to First- and Second-order SD relations. However, the First-order relation requires, in this case, a large number of binary variables, so we focus on the more applicable Second-order condition.