International Review of Financial Analysis 95 (2024) 103368

Contents lists available at ScienceDirect

International Review of Financial Analysis

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

Testing out-of-sample portfolio performance using second-order stochastic dominance constrained optimization approach

Peng Xu

Essex Business School, University of Essex, Southend-on-Sea SS1 1LW, England, United Kingdom

ARTICLE INFO

Keywords:
Out-of-sample analysis
Stochastic dominance
Portfolio optimization
Robustness
Empirical likelihood
Incomplete probability information

ABSTRACT

Second-order Stochastic Dominance (SSD) criterion can be used to support portfolio decision making under risk and uncertainty. In this paper, we develop novel robust SSD criteria to capture the strength of dominance and portfolio optimization models utilizing these criteria to identify portfolios whose in-sample SSD dominance over a given benchmark is likely to hold also out-of-sample. The developed models can incorporate incomplete probability information by allowing a set of feasible state probabilities. We also show that these portfolio optimization models can be formulated as linear programming problems. We report results from applying these SSD-based portfolio optimization models with different sets of state probabilities in an empirical application, with a focus on evaluating the out-of-sample portfolio performance of the optimized portfolios.

1. Introduction

Stochastic Dominance (SD) is a popular decision rule to support decision making under risk and uncertainty (Hadar & Russell, 1969; Hanoch & Levy, 1969; Rothschild & Stiglitz, 1970), which does not require explicit specification of the decision maker's (DM's) risk preferences (see, Levy, 2016, for an overview). Within the SD framework, Second-order Stochastic Dominance (SSD) criterion is particularly suitable for financial decision making where investors are normally assumed to be (rationally) risk-averse (Hanoch & Levy, 1969). In particular, portfolio A dominates B by SSD, if the expected utility of portfolio A is greater than or equal to that of portfolio B for all monotonically increasing concave or linear utility functions. Consequently, all risk-averse or -neutral DMs would prefer portfolio A over B. In addition, portfolio A is said to be SSD-efficient, if there exists no any other marketed portfolio that dominates portfolio A by SSD.

The theoretical appeal of the SSD criterion has attracted growing research attention and efforts in developing SSD-based portfolio optimization approaches. In practice, identifying SSD-efficient portfolios corresponds to solving a Linear Programming (LP) problem (see, e.g., Bruni et al., 2017; Consigli et al., 2020; Dentcheva & Ruszczyński, 2003, 2006; Henriksen et al., 2019; Kallio & Hardoroudi, 2018; Kopa et al., 2018; Kopa & Post, 2015; Kuosmanen, 2004; Liesiö et al., 2020; Longarela, 2016; Post, 2003; Roman et al., 2006; Vitali & Moriggia, 2021). Specifically, the SSD-based models by Post (2003) and Kuosmanen (2004) enable to (i) test if a given portfolio is SSD-efficiently diversified in view of all possible portfolios in a particular asset span; (ii) if not, construct a dominating portfolio that is SSD-efficient. The dual formulation of Kopa and Post (2015) maximizes the weighted sum of differences in cumulative portfolio returns and thus, the dominating portfolio obtained achieves SSD efficiency.

Recently, another SSD research strand grows in popularity: Assessing the out-of-sample SSD-related performance of in-sample SSD-efficient portfolios in empirical applications of SSD-based portfolio optimization (see, e.g., Hodder et al., 2015; Liesiö et al., 2020; Post et al., 2018; Roman et al., 2013). In particular, this research strand has been led by some distinct interest in evaluating and analyzing how likely such optimized portfolios obtained by SSD-based models would still dominate a particular benchmark out-of-sample. Hodder et al. (2015) document empirical findings that in-sample SSD-efficient portfolios do not necessarily dominate the benchmark portfolio in out-of-sample evaluation. In order to best preserve dominance by SSD out-of-sample, Liesiö et al. (2020) develop SSD approaches that utilize incomplete probability information to account for several types of parameter uncertainties in estimating state probabilities. Specifically, their approach, like most traditional methods, builds on expected portfolio return maximization, but identifies robust benchmark dominating portfolios with respect to a set of feasible probability vectors, rather than a single vector of state probabilities. In addition, Post et al. (2018) propose another SSD approach that maximizes expected portfolio return and utilizes the Empirical Likelihood approach to elicit the implied state probabilities subject to a set of well-specified moment conditions, capturing stylized facts about empirical finance data.

Although the existing literature has mainly focused on determining the appropriate (set of) state probabilities as an approach to improve

E-mail address: peng.xu@essex.ac.uk.

https://doi.org/10.1016/j.irfa.2024.103368
Received 16 December 2023; Received in revised form 16 May 2024; Accepted 16 May 2024
Available online 20 May 2024

out-of-sample portfolio performance, less attention has been devoted to designing robust SSD criteria to identify in-sample SSD-efficient portfolios that are also likely to dominate the benchmark out-of-sample. Specifically, if there are several portfolios dominating the benchmark in-sample, the most widely used approach is to select the one that yields the highest expected portfolio return. However, when the aim is for the dominance to hold also out-of-sample, it might be more appropriate to select a portfolio whose dominance over the benchmark is in some sense the strongest. Ideally, such novel dominance criteria could be utilized together with incomplete probability information to analyze the joint impacts of the specification of state probabilities along with the SSD criteria have on out-of-sample portfolio performance.

In this paper, we develop new robust approaches for SSD-based portfolio optimization that seek to identify portfolios whose in-sample dominance over the benchmark portfolio holds also out-of-sample. In particular, this development considers both (i) admitting incomplete probability information, and (ii) the design of novel robust SSD criteria that capture the strength of dominance over the benchmark. In terms of probability information, the SSD approaches developed here can use and combine existing estimation methods based on incomplete probability information (Liesiö et al., 2020) and the Empirical Likelihood approach (Post et al., 2018). Standard SSD approaches typically rest on, for instance, maximizing expected portfolio return and assume that the underlying information regarding state probabilities is complete. However, we take a different approach instead by maximizing the minimum distances between the integrated cumulative density functions of portfolio returns and allow incomplete information on state probabilities. We also develop computationally tractable LP models based on such distance measures for identifying robust dominating portfolios whose dominance over a given benchmark is the strongest given a set of feasible state probabilities.

This paper also analyzes and evaluates the out-of-sample performance of the developed SSD-based portfolio optimization models under complete and incomplete probability information. To examine and compare different SSD models, we implement an investment application using empirical data on the returns of industry portfolios and common risk factors. This application setup design follows prior empirical tests deployed in Arvanitis et al. (2021), Hodder et al. (2015), Liesiö et al. (2020), Post et al. (2018) and Roman et al. (2013). Specifically, for each SSD model, we deploy a rolling 12-month estimation window of daily returns for portfolio formation in-sample and rebalance after a 1-month holding period. Then, the obtained optimal portfolio is evaluated on several portfolio performance metrics out-of-sample. Results from this empirical application show that (i) using different SSD models exhibits explicit out-of-sample trade-offs between returns and dominance relations in the sense of SSD; and (ii) under complete probability information, using empirical likelihood (EL) state probabilities in a return maximizing SSD model generally outperforms equal state probabilities. However, we document no consistent performance pattern out-of-sample between using equal and EL state probabilities in the other SSD models, especially when incomplete information on state probabilities is accounted for.

There are also several other relevant streams of research in the literature that cannot be ignored. For instance, Kouaissah (2023) recently develops a robust portfolio optimization approach with weak SSD constraints allowing different distributional assumptions. Specifically, this approach formulates a robust optimization that explicitly specifies a constraint for handling asset returns following heavy-tailed probability distributions in the optimization problem itself. Thus, the resulting reward-risk performance measures obtained by the optimized portfolios are improved and less sensitive to estimation errors in out-of-sample evaluations. Sehgal and Mehra (2020) establish a robust portfolio optimization model with SSD constraints to accommodate uncertainties arising from input asset returns. In particular, these uncertain inputs are allowed to vary in symmetric and bounded intervals so that the optimized portfolios obtained would yield better out-of-sample portfolio performance. More broadly, the analysis, evaluation, and prediction of portfolio performance have always been of key interest to academics and practitioners in the realm of financial analysis research (see, among others, Canepa et al., 2020; Chavez-Bedoya & Rosales, 2021; Cipollini et al., 2021; Guerard et al., 2015; Han & Li, 2022, 2023; Hoang et al., 2015; Kassimatis, 2021; Kazak & Pohlmeier, 2019; Khashanah et al., 2022; Kim & Kang, 2021; Migliavacca et al., 2023; Pho et al., 2021; Post et al., 2019; Roccazzella et al., 2022; Xia et al., 2015). Moreover, stochastic dominance and portfolio optimization (see, Salo et al., 2023, for a recent review) research have a long tradition of supporting decision making under incomplete or imprecise information on risk preferences and probability information (see, among others, Dentcheva & Römisch, 2013; Dentcheva & Ruszczyński, 2010; Dupačová & Kopa, 2012, 2014; Egozcue & Wong, 2010; Keppe & Weber, 1989; La Torre & Mendivil, 2018, 2022; Liesiö et al., 2023; Montes et al., 2014a, 2014b; Wong, 2007).

The rest of the paper is structured as follows. Section 2 introduces the notations and standard definitions related to SSD. Section 3 presents novel robust dominance criteria in the sense of SSD and develops new SSD-based portfolio optimization models. Section 4 addresses the estimation of complete and incomplete state probabilities for the state-space of SSD-based models. Section 5 applies empirically the developed SSD models to industry portfolio optimization under complete and incomplete probability information and analyzes their resulting out-of-sample portfolio performance. Section 6 concludes.

2. Preliminaries

We model the returns of N distinct base assets as real-valued random variables $X_1, \dots, X_N$ on the set of T mutually exclusive and collectively exhaustive states $S = \{s_1, \dots, s_T\}$. The return of asset i in state t is given by $x_{i,t} = X_i(s_t) \in \mathbb{R}^{N\times T}$. A portfolio of these assets is established with a vector of asset weights $\lambda \in \mathbb{R}^N$ capturing the share of capital invested into each asset with no short position allowed $\lambda \geq 0$. Therefore, the set of all possible asset weights can be expressed as

$$\Lambda = \left\{ \lambda \in \mathbb{R}^N_+ \;\middle|\; \sum_{i=1}^{N} \lambda_i = 1 \right\}. \tag{1}$$

The return of a portfolio with weights λ is represented by random variable $X = \sum_{i=1}^N \lambda_i X_i$ whose state-specific returns are denoted by $x_t = X(s_t) = \sum_{i=1}^N \lambda_i x_{i,t}$. The set of all possible portfolios is expressed as return mixtures of the base assets

$$\mathcal{X} = \left\{ \sum_{i=1}^{N} \lambda_i X_i \;\middle|\; \lambda \in \Lambda \right\}. \tag{2}$$

Let Y be any benchmark portfolio whose state-specific returns are represented by $y_t = Y(s_t)$. The state probabilities define a probability vector $p = (p_1, \dots, p_T)$ in the T-dimensional simplex

$$P^0 = \left\{ p \in [0,1]^T \;\middle|\; \sum_{t=1}^{T} p_t = 1 \right\}. \tag{3}$$

We denote the expectation and cumulative density function (CDF) of some portfolio return X by $\mathbb{E}_p[X] = \sum_{t=1}^T p_t x_t$ and

$$F_X(\theta; p) = \mathbb{P}\left(\left\{ s_t \in S \;\middle|\; X(s_t) \leq \theta \right\} \right) = \sum_{\substack{t \\ x_t \leq \theta}} p_t, \tag{4}$$

respectively.

Second-order Stochastic Dominance (SSD) compares the integrals of random variables' cumulative density function (CDF)s. With a finite state-space, these integrals can be evaluated as finite sums and as a result, the integrated CDF of portfolio return X is,

$$F_X^2(\theta; p) = \int_{-\infty}^{\theta} F_X(\tau; p)\, d\tau = \sum_{\substack{t \\ x_t \leq \theta}} p_t(\theta - x_t) = \sum_{t=1}^{T} p_t \max\{\theta - x_t,\ 0\}. \tag{5}$$

With this notation, SSD can be formalized by the following definition.

Definition 1. Portfolio X weakly dominates portfolio Y by Second-order Stochastic Dominance, denoted by $X \succeq Y$, if and only if

$$F_X^2(\theta; p) \leq F_Y^2(\theta; p) \;\forall\; \theta \in \mathbb{R}, \tag{6}$$

where $F_{(\cdot)}^2$ evaluates the integrals of the CDF of some portfolio return.

The dominance relation $\succeq$ is explicitly linked to the Expected Utility Theory (EUT) framework: Portfolio X dominates some benchmark portfolio Y by SSD if and only if any risk-averse or -neutral expected utility maximizing decision maker prefers portfolio X to portfolio Y (Hanoch & Levy, 1969). This is since the expected utility of portfolio X is always higher than or equal to that of Y, i.e.,

$$X \succeq Y \Leftrightarrow \mathbb{E}_p[u(X)] \geq \mathbb{E}_p[u(Y)] \;\forall\; u \in U^*, \tag{7}$$

where set U^* consists of all non-decreasing utility functions that are concave.

SSD-based portfolio optimization models typically consider only complete probability information and generally assume that there exists some pre-specified probability distribution of asset returns characterized by a single state probability vector p^*. In a recent work, this assumption is relaxed by accounting also for incomplete state probability information (Liesiö et al., 2020). Specifically, such incomplete information on state probabilities can be modeled using set-valued state probabilities and this procedure is technically realized by formulating a set of feasible probabilities[1] P such that $P \subseteq P^0$, around a single state probability vector p^* (e.g., centroid of set P).

Liesiö et al. (2020) extend the definition of SSD (Definition 1) to set-valued state probabilities by requiring that the dominance relation holds for all probability vectors in this set. This leads to the dominance relation formalized in the following definition, which we refer to as robust SSD.[2]

Definition 2. Portfolio X weakly dominates portfolio Y by Second-order Stochastic Dominance with respect to the set of feasible probabilities $P \subseteq P^0$, denoted by $X \succeq_P Y$, if and only if

$$F_X^2(\theta; p) \leq F_Y^2(\theta; p) \;\forall\; \theta \in \mathbb{R},\; p \in P. \tag{8}$$

This robust dominance relation $\succeq_P$ also has an EUT-based interpretation: If portfolio X dominates some benchmark portfolio Y by SSD, then a risk-averse or -neutral expected utility maximizing decision maker would prefer portfolio X over Y for all of the state probability vectors in set P (Liesiö et al., 2020). Consequently, the resulting expected utility interpretation of robust SSD can be formally stated as

$$X \succeq_P Y \Leftrightarrow \mathbb{E}_p[u(X)] \geq \mathbb{E}_p[u(Y)] \;\forall\; u \in U^*,\; p \in P. \tag{9}$$

3. Portfolio optimization based on robust second-order stochastic dominance

In this section, we introduce novel robust dominance criteria and develop new portfolio optimization approaches for identifying robust dominating portfolio X whose in-sample strength of dominance over some pre-specified benchmark portfolio Y is the strongest. Specifically, we develop computationally tractable SSD-based portfolio optimization models based on distance measures between the integrated CDFs of portfolios X and Y, with incomplete probability information captured by some set of feasible probabilities P. Without loss of generality, we assume, throughout this section, that the states are indexed in an ascending order of state-specific returns of the benchmark portfolio Y such that $y_1 \leq y_2 \leq \cdots \leq y_{T-1} \leq y_T$.

3.1. Establishing robust second-order stochastic dominance criteria

Perhaps the most obvious approach to utilize robust dominance criteria in portfolio optimization is to require that the optimized portfolio X dominates some pre-specified benchmark portfolio Y for all feasible probability vectors in set P. As there can be several portfolios dominating the benchmark, the selection among these portfolios can be based on maximizing the expected portfolio return under some pre-specified state probabilities p^* in set P. As established by Liesiö et al. (2020), this portfolio optimization problem can be formulated as

$$\max_{X \in \mathcal{X}} \mathbb{E}_{p^*}[X] \;\; s.t. \;\; X \succeq_P Y. \tag{10}$$

If set P consists only of a single vector of state probabilities $P = \{p^*\}$, then model (10) reduces to a standard SSD problem with complete probability information (see, e.g., Kopa & Post, 2015; Kuosmanen, 2004; Post et al., 2018).

However, if the aim is to identify a portfolio whose dominance over the benchmark is robust in the sense that it would hold also out-of-sample, it might not be advisable to select the portfolio with the highest expected return. Thus, we propose selecting the portfolio whose dominance over the benchmark is the strongest. One approach to capture the strength of dominance is to measure the maximal allowed increase in the return of benchmark portfolio such that the established dominance relation would still hold. Formally, this corresponds to finding the maximal value $\varphi \in \mathbb{R}_+$ such that X dominates $Y + \varphi$. This measure can be readily extended to account for incomplete probability information by considering the maximal value of φ across all feasible probability vectors in set P. Thus, identifying a portfolio whose dominance over the benchmark is the strongest corresponds to solving the optimization problem

$$\max_{\varphi \geq 0,\; X \in \mathcal{X}} \varphi \;\; s.t. \;\; X \succeq_P Y + \varphi. \tag{11}$$

Intuitively, φ can be seen as a risk-free rate added to portfolio Y. Alternatively, φ can also be interpreted to measure the minimal horizontal distance between the integrated CDFs of portfolios X and Y.

This interpretation suggests an alternative approach in which the strength of dominance is measured as the minimal vertical distance between the integrated CDFs of portfolios X and Y. For robust SSD to hold between portfolios X and Y such that $X \succeq_P Y$, it is required that the integrated CDF of portfolio X remains below or equal to that of portfolio Y for all feasible probability vectors in set P. This condition implies that $y_1 = \min(Y) \leq \min(X)$ and therefore $F_Y^2(y_1; p) = F_X^2(y_1; p) = 0$ for all $p \in P$. To avoid this situation, in which the minimal vertical distance is always zero, we trim the left tail of return distribution of the benchmark portfolio Y and evaluate the vertical distance between F_X^2 and F_Y^2 only for return levels exceeding the second smallest state-specific return of the benchmark y_2. In particular, this vertical distance is formally defined as

$$\Delta(X, Y) = \min_{\theta \geq y_2,\; p \in P} F_Y^2(\theta; p) - F_X^2(\theta; p). \tag{12}$$

Hence, a robust benchmark dominating portfolio that maximizes this minimal vertical distance is formally obtained as an optimal solution to the optimization problem

$$\max_{X \in \mathcal{X}} \Delta(X, Y) \;\; s.t. \;\; X \succeq_P Y. \tag{13}$$

For a single state probability vector p, Fig. 1 illustrates the robust SSD criteria introduced above. Clearly, portfolio X stochastically dominates portfolio Y by SSD, as the integrated CDF of X remains below or equal to that of Y for all return levels θ. The blue stars and red dots represent the state-specific returns of dominating portfolio X and those of benchmark Y, respectively. Essentially, φ measures the minimal horizontal distance between F_Y^2 and F_X^2, whereas Δ quantifies the minimal vertical distance between F_Y^2 and F_X^2.

[1] Sets of feasible probabilities are also interchangeably known as credal sets in Levi (1980).

[2] A classic definition of robust SSD is presented in Dentcheva and Ruszczyński (2010).

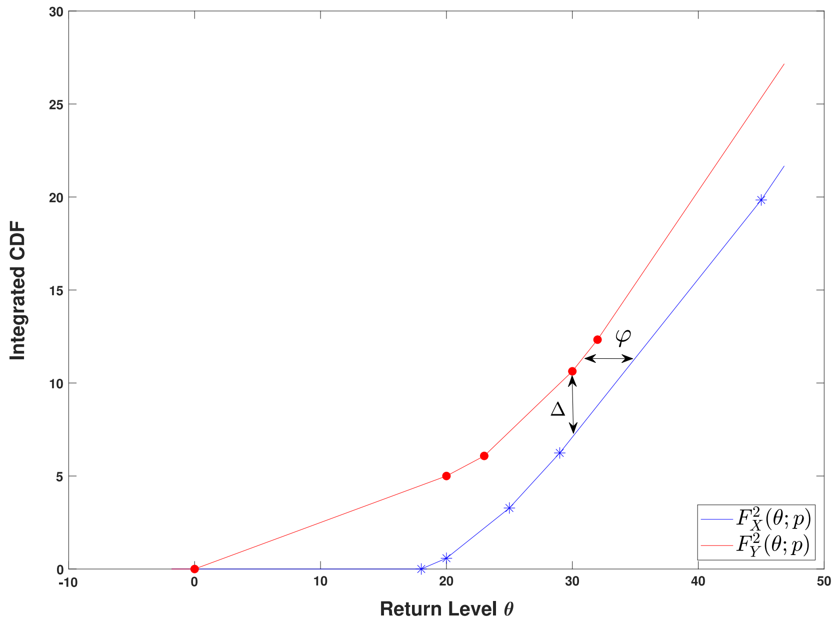

Fig. 1. Illustration of new robust dominance criteria in the sense of SSD.

3.2. Optimization models for identifying robust dominating portfolios

Notably, optimization problem (10) can be formulated as an LP problem if set P has a finite number of extreme points (cf. state probability vectors). This is since it suffices to ensure that portfolio X dominates the benchmark portfolio Y by SSD at these extreme probability vectors, which can be implemented using well-known LP techniques (see, e.g., Dentcheva & Ruszczyński, 2003; Rockafellar & Uryasev, 2000). Technically, this implementation requires introducing an auxiliary decision variable $z \in \mathbb{R}_+^{T\times T}$ and, ultimately, optimization problem (10) can be formulated using the LP setup from Liesiö et al. (2020) (hereinafter referred to as 'SSD–LXK' model)

$$\max_{\lambda\in\Lambda,\ z\in\mathbb{R}_+^{T\times T}} \sum_{t=1}^{T} p_t^* \sum_{i=1}^{N} \lambda_i x_{i,t} \tag{14}$$

$$s.t. \sum_{i=1}^{N} \lambda_i x_{i,s} + z_{t,s} \geq y_t \ \forall\ t, s \in \{1, \dots, T\} \tag{15}$$

$$\sum_{s=1}^{T} p_s z_{t,s} \leq F_Y^2(y_t; p)\ \forall\ t \in \{1, \dots, T\}, \quad \forall\ p \in \text{ext}(\text{conv}(P)), \tag{16}$$

where ext(conv(P)) denotes the set of extreme points of the convex hull of set P.

In order to develop LP formulations for optimization problems (11) and (13), we establish the following auxiliary lemma that will be used in establishing the subsequent theorems. Essentially, the lemma states that the minimal vertical distance between the integrated CDFs of two portfolio returns is always found at a return level equal to one of the state-specific returns of the dominated portfolio.

Lemma 1. *Let V and W be random variables on the discrete state-space $S = \{s_1, \dots, s_T\}$ such that $v_t = V(s_t)$ and $w_t = W(s_t)$. If $F_V^2(v_t; p) \geq F_W^2(v_t; p)$ for all $t \in \{1, \dots, T\}$ and $p \in \text{ext}(\text{conv}(P))$, then $F_V^2(\theta; p) \geq F_W^2(\theta; p)$ for all $\theta \in \mathbb{R}$ and $p \in \text{ext}(\text{conv}(P))$.*

Proof. See Appendix A.

This lemma can be utilized to obtain an LP formulation for optimization problem (11) with a similar structure as LP problem (14)–(16). This result is formally stated by the following theorem.

Theorem 1. *Portfolio $X^* = \sum_{i=1}^{N} \lambda_i X_i$ and $\varphi \in \mathbb{R}_+$ form an optimal solution to optimization problem (11) if and only if there exists $z \in \mathbb{R}_+^{T\times T}$ such that (λ, z, φ) is an optimal solution to the LP problem*

$$\max_{\substack{\lambda\in\Lambda,\ z\in\mathbb{R}_+^{T\times T}\\ \varphi\in\mathbb{R}_+}} \varphi \tag{17}$$

$$s.t. \sum_{i=1}^{N} \lambda_i x_{i,s} + z_{t,s} \geq y_t + \varphi\ \forall\ t, s \in \{1, \dots, T\} \tag{18}$$

$$\sum_{s=1}^{T} p_s z_{t,s} \leq F_Y^2(y_t; p)\ \forall\ t \in \{1, \dots, T\},\ \forall\ p \in \text{ext}(\text{conv}(P)). \tag{19}$$

Proof. See Appendix A.

LP problem (17)–(19), which is hereinafter referred to as 'SSD–φ' model, identifies a robust dominating portfolio X over the benchmark portfolio Y by SSD, for a given set of feasible probabilities P. Essentially, 'SSD–φ' model has $N + T \times T + 1$ decision variables and $T \times T + T \times Q + 1$ linear constraints, where Q denotes the number of extreme state probability vectors of the convex hull of set P.

The LP formulation for optimization problem (13) is formally established by the following theorem.

Theorem 2. *Portfolio $X^* = \sum_{i=1}^{N} \lambda_i X_i$ is an optimal solution to optimization problem (13) if and only if there exist $z \in \mathbb{R}_+^{T\times T}$ and $\delta \in \mathbb{R}_+$ such that (λ, z, δ) is an optimal solution to the LP problem*

$$\max_{\substack{\lambda\in\Lambda,\ z\in\mathbb{R}_+^{T\times T}\\ \delta\in\mathbb{R}_+}} \delta \tag{20}$$

$$s.t. \sum_{i=1}^{N} \lambda_i x_{i,s} + z_{t,s} \geq y_t\ \forall\ t, s \in \{1, \dots, T\} \tag{21}$$

$$\sum_{s=1}^{T} p_s z_{1,s} \leq F_Y^2(y_1; p)\ \forall\ p \in \text{ext}(\text{conv}(P)) \tag{22}$$

$$\sum_{s=1}^{T} p_s z_{t,s} + \delta \leq F_Y^2(y_t; p)\ \forall\ t \in \{2, \dots, T\},\quad \forall\ p \in \text{ext}(\text{conv}(P)). \tag{23}$$

Proof. See Appendix A.

Theorem 2 formulates an LP problem for identifying a robust benchmark dominating portfolio X in the sense of SSD with respect to some set of feasible probabilities P. LP problem (20)–(23) is referred to as 'SSD–Δ' model. From a computational perspective, it seems that SSD–Δ would have the same level of model complexity in comparison to SSD–φ in terms of, for instance, the number of decision variables as well as that of the constraints. Nevertheless, empirical results document quite different solution times from applications of these two SSD models, the details of which are subsequently reported in Section 5.3.

4. Estimating complete and incomplete probability information

In addition to the utilization of robust SSD criteria, another approach to improve out-of-sample portfolio performance is to specify appropriate (set of) state probabilities. In this section, we introduce different methods to estimate state probabilities covering complete and incomplete information. In particular, we first consider the estimation of complete probability information using the traditional Empirical Distribution Function (EDF) approach and also a more recent Empirical Likelihood (EL) approach (Post et al., 2018). Then, we explore the estimation further by allowing incomplete information on such point-estimate state probabilities. Specifically, we utilize the notion of set-valued state probabilities (Liesiö et al., 2020) to capture parameter uncertainties in the estimation, in which multiple probability vectors form a confidence region around a single vector of state probabilities p^*. Moreover, we contribute to the literature by introducing a novel hybrid approach for state probability estimation that technically combines the approaches of Liesiö et al. (2020) and Post et al. (2018).

4.1. Complete probability information

Traditionally, the state-space of SSD-based portfolio optimization models is generated using a sample of T most recent realized asset returns with equal state probabilities $p_t = \frac{1}{T},\ t \in \{1,\dots,T\}$ assigned to each state vector of returns, representing equally probable scenarios of future asset returns (see, e.g., Kopa & Post, 2015; Kuosmanen, 2004). This 'plug-in' procedure equivalently corresponds to consistently estimating the true CDFs of the joint asset return distribution using the Empirical Distribution Function (EDF) approach (see, e.g., Post et al., 2018; also, Liesiö et al., 2020). The EDF is in essence a consistent estimator for state probabilities given that certain statistical assumptions are fulfilled.

However, for most empirical applications of SSD-based portfolio optimization, the sample size T of asset returns in the state-space is unfortunately not 'large' enough in the statistical sense. Thus, favorable asymptotic properties generally fail to guarantee that the estimated CDFs converge in probability to the true CDFs of the joint return distribution. As a consequence, using the EDF estimator to estimate the state probabilities can be vulnerable to estimation errors, as it fails to account for dynamic patterns in empirical finance data. To improve estimation accuracy, Post et al. (2018) propose the Empirical Likelihood (EL) approach to elicit the state probabilities based on a system of pre-specified moment conditions, e.g., capturing empirical stylized facts about returns of various asset classes.

The Blockwise Empirical Likelihood (BEL) approach by Kitamura (1997) can be used to take dynamic patterns into account. Specifically, a sample of T asset returns in the state-space is divided into multiple data blocks with overlapping observations. Let B be the number of return observations in each block and, as a result, the total number of overlapping data blocks is $T^* = T - B + 1$. For instance, consider a sample of time series data with 6 asset return observations $x^1,\dots,x^6$, where $x^t = (x_{1,t},\dots,x_{N,t})'$ and the block size $B = 3$. Hence, by the BEL approach, altogether $6 - 3 + 1 = 4$ data blocks can be formed: Observations x^1, x^2, x^3 are framed into the first block; observations x^2, x^3, x^4 enter the second one; observations x^3, x^4, x^5 fit into the third one. Ultimately, the last (fourth) block contains observations x^4, x^5, x^6. In general, the index of the jth element of the kth block is formally given by $(k-1)+j$, and any return observation x^t is therefore present throughout all data blocks indexed from $t^- = \max(1,\ t-B+1)$ till $t^+ = \min(t,\ T^*)$. Each block is assigned a block-level probability b_k and the state-level EL probabilities can then be obtained from these block-level probabilities through

$$p_t^{EL} = \frac{1}{B}\sum_{k=t^-}^{t^+} b_k \ \forall\ t \in \{1,\dots,T\}. \tag{24}$$

The BEL approach estimates these block-level probabilities such that the expected values they imply for the chosen common risk factors (e.g., market premium, size, value, momentum) belong to some pre-specified intervals. In particular, let there be F common risk factors denoted by random variables $\Phi_1,\dots,\Phi_F$ on the state-space. Moreover, denote the return of common risk factor Φ_f in state t by $\phi_{f,t} = \Phi_f(s_t) \in \mathbb{R}^{F\times T}$. Constraints on these common risk factors can then be formalized as

$$L_f \le \mathbb{E}_{p^{EL}}[\Phi_f] \le U_f, f \in \{1,\dots,F\}, \tag{25}$$

where L_f and U_f denote the lower and upper bounds, respectively, specified for their realized distributions at certain percentile levels. Evaluating the expectation of a particular factor return Φ_f under state probabilities p^{EL} gives

$$\mathbb{E}_{p^{EL}}[\Phi_f] = \sum_{t=1}^{T} p_t^{EL}\phi_{f,t} = \sum_{t=1}^{T}\left(\frac{1}{B}\sum_{k=t^-}^{t^+} b_k\right)\phi_{f,t} = \frac{1}{B}\sum_{t=1}^{T}\sum_{k=t^-}^{t^+} b_k\phi_{f,t}, \tag{26}$$

where $t^- = \max(1,\ t-B+1)$ and $t^+ = \min(t,\ T^*)$. Now, changing the order of summation and utilizing the fact that the kth block consists of the states with indices from k to $(k-1)+B$ yield

$$\mathbb{E}_{p^{EL}}[\Phi_f] = \frac{1}{B}\sum_{k=1}^{T^*}\sum_{t=k}^{(k-1)+B} b_k\phi_{f,t} = \sum_{k=1}^{T^*} b_k \sum_{j=1}^{B}\frac{1}{B}\phi_{f,(k-1)+j}. \tag{27}$$

Among those block-level probabilities $(b_1,\dots,b_{T^*})$ satisfying constraints (25) on expected factor returns, the BEL approach finds the ones that minimize the Kullback–Leibler distance from equal block probabilities $(\frac{1}{T^*},\dots,\frac{1}{T^*})$ by solving the optimization problem

$$\min_{b\in\mathbb{R}_+^{T^*}} \ -\frac{1}{T^*}\sum_{k=1}^{T^*}\ln(b_k) - \ln(T^*) \tag{28}$$

$$L_f \le \sum_{k=1}^{T^*} b_k \sum_{j=1}^{B}\frac{1}{B}\phi_{[f,(k-1)+j]} \le U_f \ \forall\ f \in \{1,\dots,F\} \tag{29}$$

$$\sum_{k=1}^{T^*} b_k = 1, \tag{30}$$

where constraint (29) is obtained by substituting (27) into (25).

4.2. Incomplete probability information

In addition to specifying point-estimate state probabilities, another approach to estimate the state probabilities is to allow incomplete probability information in the estimation. In practice, in most empirical applications of SSD-based portfolio optimization, precise or complete information regarding the state probabilities can be difficult to acquire or justify (see, Liesiö & Salo, 2012; Moskowitz et al., 1993; Vilkkumaa et al., 2018). Liesiö et al. (2020) avoid the exact specification of point-estimate state probabilities by accepting multiple parameter uncertainties in modeling state probabilities, known as *probability ranking*, *confidence region around equal state probabilities*, and *varying sample size*. For instance, the use of set-valued state probabilities enables to build up a confidence region around a vector of equal state probabilities $\bar{p} = (\frac{1}{T},\dots,\frac{1}{T})$ (see, Liesiö et al., 2020). Formally, this type of set-valued state probabilities can be formulated as

$$P_\alpha^{ES} = \left\{ p \in P^0 \ \middle|\ p_t \ge \frac{\alpha}{T},\ \forall\ t \in \{1,\dots,T\} \right\}, \tag{31}$$

where parameter $\alpha \in [0,1]$ controls the minimum probability across all states. Thus, decreasing the value of parameter α increases the size of set P_α^{ES}: The smallest set P_1^{ES} consists of a single vector of equal state probabilities, i.e., $P_1^{ES} = \{\bar{p}\}$, while the largest set P_0^{ES} coincides with the T-dimensional simplex.

Instead of setting the lower bounds proportional to equal state probabilities, we construct another set-valued state probabilities, in which these lower bounds on state probabilities are relaxed in a decrement of β. In particular, with $\bar{p} = (\frac{1}{T}, \dots, \frac{1}{T})$ being again the center point, this type of set can be formally expressed as

$$\tilde{P}_\beta^{ES} = \left\{ p \in P^0 \;\middle|\; p_t \geq \max\left\{ p_t - \beta,\ 0 \right\},\ \forall\, t \in \{1, \dots, T\} \right\}, \tag{32}$$

where parameter β is a positive real number. However, for any $p_t < \beta$, this results in a situation that the lower bound falls negative. As a result, the max operator is introduced to ensure non-negative lower bounds on the state probabilities. Moreover, increasing the value of β leads to a size increase of set $\tilde{P}^{ES}$. Hence, set $\tilde{P}_0^{ES}$ is the smallest in size corresponding to equal state probabilities, a.k.a., $\tilde{P}_0^{ES} = \{\bar{p}\}$.

We suggest also a novel hybrid approach utilizing incomplete specification of the state probabilities that combines the approaches by Post et al. (2018) and Liesiö et al. (2020). The EL approach in Post et al. (2018) solves optimization problem (28)–(30) and selects the one vector of block probabilities $b^* = (b_1, \dots, b_{T^*})$ that minimizes the divergence from equal block probabilities $(\frac{1}{T^*}, \dots, \frac{1}{T^*})$ w.r.t. the Kullback–Leibler distance, subject to a system of linear constraints enforced by moment conditions of common risk factors (cf. constraints (25)). With b^*, we are able to compute EL-estimate state probabilities p^{EL} using Eq. (24), then constructing a confidence region around this center point p^{EL} with the approach from Liesiö et al. (2020) gives a set-valued state probabilities

$$P_\alpha^{EL} = \left\{ p \in P^0 \;\middle|\; p_t \geq \alpha p_t^{EL},\ \forall\, t \in \{1, \dots, T\} \right\}, \tag{33}$$

where parameter $\alpha \in [0,1]$ determines the lower bound on each state probability. Similar to set (31), the smallest set P_α^{EL} with only one vector of EL-estimate state probabilities p^{EL} can be obtained with $\alpha = 1$, i.e., $P_1^{EL} = \{p^{EL}\}$. Subsequently, a decrease in α value leads to a size increase of set P_α^{EL} until the T-dimensional simplex is attained once again at $\alpha = 0$ with the largest set P_0^{EL}. Furthermore, we establish another set-valued state probabilities based on this EL-estimate state probabilities p^{EL}, around which the extreme state probability vectors are symmetric. Formally, replacing the center point $\bar{p}$ in set (32) with p^{EL} yields

$$\tilde{P}_\beta^{EL} = \left\{ p \in P^0 \;\middle|\; p_t \geq \max\left\{ p_t^{EL} - \beta,\ 0 \right\},\ \forall\, t \in \{1, \dots, T\} \right\}. \tag{34}$$

5. Empirical analysis: Industrial portfolio optimization

In this section, we apply the developed SSD-based portfolio optimization approaches to empirical data sets. In particular, we run empirical tests using LP models SSD–LXK (cf. (14)–(16)), SSD–φ (cf. (17)–(19)), and SSD–Δ (cf. (20)–(23)) and evaluate their resulting out-of-sample portfolio performance. We also analyze whether admitting incomplete probability information by establishing a set of feasible probabilities around point-estimate such as equal and empirical likelihood (EL) state probabilities improves the portfolio performance of these SSD methods out-of-sample. Moreover, we compare if the use of EL state probabilities in these SSD approaches results in better out-of-sample performance than equal state probabilities. Finally, we conduct sensitivity analysis examining how the block size affects these SSD models when utilizing EL-based probability sets.

5.1. Data

The empirical data set is comprised of daily returns of three types of assets. First, the base assets $X_1, \dots, X_{49}$ are represented by the Fama–French 49 value-weighted industry portfolios. Second, the benchmark market portfolio Y is an all-share proxy index of the value-weighted average return of all CRSP (The Center for Research in Security Prices) stocks listed on the NYSE, AMEX, or NASDAQ Exchanges. Third, we deploy four common risk factors, namely, market premium (RMRF, Φ_1), size (SMB, Φ_2), value (HML, Φ_3), and momentum (MOM, Φ_4) (Carhart, 1997; Fama & French, 1993). In particular, RMRF captures the excess market return over the risk-free rate. SMB is the market capitalization factor measured as excess returns of small cap stocks over large cap stocks. HML is the factor related to value, i.e., excess returns of high book-to-market (value) stocks to low book-to-market (growth) stocks. MOM stands for the momentum factor and tracks the tendency for asset returns to continue moving along a rising (winners) or falling (losers) path. MOM is also referred to as up-minus-down factor (UMD). Tables 1 and 2 present the descriptive statistics of 49 industry portfolios, the benchmark market portfolio, and the common risk factors. The data sample records all daily return observations over a time period from January 3rd 1927 through December 29th 2017 and thus, includes 23 990 trading days spanning across 91 years.

5.2. Investment strategy, empirical test specification

We use a plain vanilla buy-hold investment strategy allowing no short sales based on a rolling estimation window with a 12-month formation period and portfolio rebalancing occurring after a 1-month holding period. This experimental setup has been applied widely in the existing literature (see, among others, Arvanitis et al., 2021; Hodder et al., 2015; Liesiö et al., 2020; Post et al., 2018; Post & Kopa, 2017) and is also motivated by empirical finance studies on optimal investment strategies (Moskowitz & Grinblatt, 1999). Moreover, Liesiö et al. (2020) report that the likelihood of obtaining full SSD over the benchmark market portfolio declines with longer holding periods.

Specifically, executing this investment strategy on our data sample gives in total 1080 overlapping portfolio formation periods, moving forward by 1-month period at a time. In particular, for each formation period, we use the daily base asset returns in each 12-month formation period as the state-space to solve SSD–LXK model (cf. (14)–(16)), SSD–φ model (cf. (17)–(19)), as well as SSD–Δ model (cf. (20)–(23)). The optimal portfolios are then obtained by these SSD models in-sample for each period by using different types of probability information.

We test particularly two types of set-valued state probabilities P by employing equal and empirical likelihood (EL) state probabilities. For probability sets based on equal state probabilities, we test sets $P = P_\alpha^{ES}$ and $P = \tilde{P}_\beta^{ES}$, for $\alpha \in \{1.00,\ 0.96,\ 0.92,\ 0.90\}$ and $\beta \in \{0,\ 1\times10^{-4},\ 2\times10^{-4},\ 3\times10^{-4}\}$, and evaluate the expectation in (14) with equal state probabilities, i.e., $p^* = \bar{p}$. For probability sets established by using EL state probabilities, another two sets $P = P_\alpha^{EL}$ and $P = \tilde{P}_\beta^{EL}$ are also tested on the above listed parameters α and β, and the expectation evaluation of (14) is therefore based on EL state probabilities, i.e., $p^* = p^{EL}$.

In order to obtain EL-estimate state probabilities p^{EL} through Eq. (24), the EL approach (Post et al., 2018) applies to obtain block-level probabilities $b^* = (b_1, \dots, b_{T^*})$ by solving optimization problem (28)–(30). However, this optimization problem requires a number of model inputs. For instance, (i) block size B needs to be specified beforehand, and the total number of overlapping data blocks T^* is, as a result, determined, (ii) F common risk factor returns, and (iii) the specification of lower and upper bounds, L and U, on the moment conditions of these risk factors (cf. constraint (29)). In particular, with other things being equal, we select one of the best-performing test cases from Post et al. (2018) (see, Table 4) as the inputs for optimization problem (28)–(30). Specifically, we choose to deploy 4 factors ($F =$

Table 1
Descriptive statistics of 49 industry portfolios and benchmark market portfolio.

SIC	Statistics					
	Mean	Std.	Skew.	Kurt.	Min	Max
Agric	0.043	1.50	0.61	18.64	−15.27	23.69
Food	0.044	0.92	−0.06	24.46	−16.04	15.54
Soda	0.055	1.38	−0.28	14.29	−19.22	11.68
Beer	0.054	1.45	0.01	23.04	−24.06	19.91
Smoke	0.052	1.19	0.16	16.20	−13.99	16.22
Toys	0.046	2.14	0.58	29.69	−26.75	39.74
Fun	0.053	1.80	0.18	15.65	−24.11	20.81
Books	0.041	1.55	0.84	28.29	−19.34	33.40
Hshld	0.041	1.16	−0.15	34.96	−21.46	25.87
Clths	0.039	1.14	−0.15	25.94	−18.51	20.49
Hlth	0.043	1.52	−0.19	12.91	−15.45	17.39
MedEq	0.053	1.59	13.06	1076.71	−53.62	111.82
Drugs	0.049	1.14	−0.24	20.16	−18.70	16.70
Chems	0.047	1.27	−0.15	18.88	−18.91	16.86
Rubbr	0.054	1.67	0.58	28.25	−19.79	26.32
Txtls	0.041	1.31	0.10	18.88	−18.40	19.50
BldMt	0.043	1.25	0.05	21.11	−17.96	22.97
Cnstr	0.049	2.00	0.68	18.62	−23.81	29.35
Steel	0.040	1.67	0.58	28.95	−24.04	30.39
FabPr	0.034	1.49	−0.13	8.96	−15.45	11.44
Mach	0.046	1.37	0.30	22.55	−18.06	26.16
ElcEq	0.053	1.56	0.20	16.73	−19.70	24.44
Autos	0.047	1.57	0.35	17.92	−19.72	27.88
Aero	0.063	1.77	0.49	22.41	−19.29	32.00
Ships	0.043	1.51	0.09	10.71	−13.20	16.62

SIC	Statistics					
	Mean	Std.	Skew.	Kurt.	Min	Max
Guns	0.056	1.38	−0.05	11.09	−19.48	14.92
Gold	0.047	2.27	0.42	10.01	−23.38	25.56
Mines	0.045	1.53	0.21	16.89	−17.91	19.85
Coal	0.046	2.12	0.33	15.39	−19.34	27.31
Oil	0.046	1.28	0.08	16.82	−19.50	19.27
Util	0.038	1.09	0.29	26.46	−15.26	17.92
Telcm	0.039	1.03	0.20	20.95	−16.69	15.98
PerSv	0.045	2.01	0.32	28.45	−30.99	30.61
BusSv	0.050	1.97	5.32	246.36	−37.41	61.56
Hardw	0.054	1.53	−0.02	18.87	−23.52	21.65
Softw	0.044	2.36	0.65	14.01	−20.76	24.19
Chips	0.052	1.75	0.16	26.42	−30.57	37.90
LabEq	0.050	1.42	−0.05	12.24	−18.78	15.93
Paper	0.075	3.27	8.62	309.60	−45.65	150.00
Boxes	0.049	1.25	−0.18	14.20	−21.43	12.59
Trans	0.040	1.35	0.12	15.27	−17.56	18.49
Whlsl	0.040	1.61	3.13	189.22	−44.44	66.92
Rtail	0.045	1.13	0.00	17.00	−18.01	17.81
Meals	0.048	1.34	−0.03	13.00	−15.48	19.40
Banks	0.052	1.47	0.31	25.73	−20.43	23.05
Insur	0.045	1.37	0.35	22.01	−17.15	18.93
RlEst	0.037	2.14	1.10	24.76	−21.23	36.78
Fin	0.048	1.57	0.05	28.22	−28.65	23.28
Other	0.032	1.48	−0.05	15.25	−20.26	16.84
Bench	0.042	1.07	−0.12	19.73	−17.41	15.76

Table 2
Descriptive statistics of 4 common risk factors.

Risk factors	Statistics					
	Mean	Std.	Skew.	Kurt.	Min	Max
RMRF	0.029	1.07	−0.11	19.75	−17.44	15.76
SMB	0.005	0.59	−0.77	25.85	−11.62	8.21
HML	0.016	0.59	0.74	19.00	−5.98	8.43
MOM	0.027	0.75	−1.61	30.33	−18.33	7.01

4, Carhart, 1997, 4-factor model), instead of 3 ($F = 3$, Fama & French, 1993, 3-factor model), in our test setup because addition of the MOM factor yields indeed informative out-of-sample results. In addition, for each formation period, we use the daily factor returns in the identical 12-month formation period (mentioned above) to compute L_f and U_f, which correspond to 10th and 90th percentiles of each risk factor's historical distribution, respectively, for $f \in \{1, \ldots, 4\}$. Moreover, in order to evaluate the impact of varying block size B, these empirical tests are carried out using a set of different block sizes $B \in \{2, 3, 5, 10\}$.

In the end, for each optimal portfolio obtained from SSD–LXK, SSD–φ, and SSD–Δ models, we evaluate its resulting out-of-sample performance on several metrics using daily returns from the subsequent 1-month holding period. Hence, a total of 1080 holding periods are evaluated out-of-sample for each test case of P.

5.3. Computation

LP models SSD–LXK, SSD–φ, and SSD–Δ are implemented in MATLAB and solved with Gurobi 9.1.1. on Aalto University's high performance computing cluster *Triton* with an Intel Xeon Processor @ 2.50 GHz and 16 GB of RAM per executor node. In addition, optimization problem (28)–(30) is solved by the nonlinear *fmincon* solver from the MATLAB optimization toolbox on a standard laptop with an Intel Core i7 Processor @ 2.60 GHz and 32 GB of RAM.

Solving SSD–LXK model takes some 3 to 420 s. The solution time of SSD–φ model varies between 14 and 671 s, while solving SSD–Δ model requires 6 to 237 s. In fact, solving SSD–φ model is more time consuming. Naturally, for larger probability sets, it takes more time on average to solve the underlying LP problem but there is no significant difference observed in the solution time between using probability sets based on equal and EL state probabilities. Moreover, solving optimization problem (28)–(30) to obtain EL probability estimates takes approximately 10 s.

5.4. Results

Tables 3 and 4 report the out-of-sample performance of in-sample optimized portfolios from LP models SSD–LXK, SSD–φ, and SSD–Δ with different probability information. In particular, Table 3 presents the results obtained by using two distinct set-valued state probabilities based on equal state probabilities P_α^{ES} (see, Eq. (31)) and $\tilde{P}_\beta^{ES}$ (see, Eq. (32)), for $\alpha \in \{1.00, 0.96, 0.92, 0.90\}$ and $\beta \in \{0, 1 \times 10^{-4}, 2 \times 10^{-4}, 3 \times 10^{-4}\}$. Table 4 shows the corresponding outcomes from the use of probability sets P_α^{EL} (see, Eq. (33)) and $\tilde{P}_\beta^{EL}$ (see, Eq. (34)), both of which are established with empirical likelihood (EL) state probabilities obtained with the block size $B = 2$. Some preliminary empirical tests showed that it becomes impossible to identify a portfolio dominating the benchmark in-sample, when α falls below the value of 0.9 in sets P_α^{ES} and P_α^{EL} or when β increases beyond the value of 0.0003 for sets $\tilde{P}_\beta^{ES}$ and $\tilde{P}_\beta^{EL}$.

In general, using SSD–LXK model yields always the highest spreads in comparison to SSD–φ and SSD–Δ models. The use of SSD–φ model gives the second best spreads following SSD–LXK, whereas the least spreads are earned by SSD–Δ model, for all probability sets based on both equal and EL state probabilities. With a size increase in sets P_α^{ES} and P_α^{EL}, for $\alpha \in \{1.00, 0.96, 0.92, 0.90\}$, as well as in $\tilde{P}_\beta^{ES}$ and $\tilde{P}_\beta^{EL}$, for $\beta \in \{0, 1 \times 10^{-4}, 2 \times 10^{-4}, 3 \times 10^{-4}\}$, applying SSD–LXK and SSD–φ models exhibits diminishing returns in mean spreads. However, allowing incomplete information on state probabilities with SSD–Δ model always outperforms complete probability information specified by a single vector of either equal (i.e., $P_1^{ES} = \tilde{P}_0^{ES}$) or EL state probabilities (i.e., $P_1^{EL} = \tilde{P}_0^{EL}$). Moreover, under such complete probability information, using SSD–Δ model offers the lowest active risk (i.e., the standard deviation of spreads), as well as the lowest downside risk in terms of Conditional Value-at-Risk (CVaR$_{5\%}$). Although increasing the size of sets P_α^{ES} and P_α^{EL}, for $\alpha \in \{1.00, 0.96, 0.92, 0.90\}$, as well as that of sets $\tilde{P}_\beta^{ES}$ and $\tilde{P}_\beta^{EL}$, for $\beta \in \{0, 1 \times 10^{-4}, 2 \times 10^{-4}, 3 \times 10^{-4}\}$

Table 3

Out-of-sample portfolio performance with probability sets based on equal state probabilities. The first four columns show the mean, standard deviation, skewness, and $\mathrm{CVaR}_{5\%}$ of annualized excess returns (spreads) to the benchmark portfolio. The next three columns under risk-adjusted metrics present annualized Sharpe, Sortino, and Information (Info.) ratios. The second last two columns under asset diversification report portfolio turnover and average number of industries (out of 49) included in the optimal portfolios. The last two columns present two dominance measures, almost-SSD (ASSD) and full SSDs over the benchmark.

Set P	Model	Spread over benchmark				Risk–adjusted metrics			Asset diversification		Dominance	
		Mean %	Std. %	Skew ness	$\mathrm{CVaR}_{5\%}$ %	Sharpe ratio	Sortino ratio	Info. ratio	Turnover	Industries # in 49	ASSD ε	SSDs %
P_1^{ES}	SSD–LXK	5.934	15.33	0.90	−17.75	0.599	0.965	0.387	0.758	5.145	0.2836	14.54
=	SSD–φ	5.529	13.05	0.75	−17.34	0.624	1.029	0.424	0.681	6.768	0.1822	24.35
$\tilde{P}_0^{ES}$	SSD–Δ	1.400	9.47	0.18	−16.31	0.550	0.243	0.148	0.451	39.184	0.0259	45.65
	SSD–LXK	5.387	11.62	1.00	−12.92	0.636	1.284	0.464	0.726	7.469	0.1733	25.28
$P_{0.96}^{ES}$	SSD–φ	4.117	7.83	1.08	−8.01	0.604	1.717	0.526	0.667	12.837	0.1300	28.43
	SSD–Δ	2.440	8.46	0.47	−11.93	0.597	0.575	0.288	0.531	36.893	0.0442	43.06
	SSD–LXK	4.487	8.79	1.10	−7.68	0.621	1.702	0.511	0.687	11.576	0.1187	29.81
$P_{0.92}^{ES}$	SSD–φ	2.706	5.04	1.05	−4.58	0.550	1.715	0.536	0.545	19.872	0.1146	29.44
	SSD–Δ	2.629	6.40	0.88	−8.03	0.580	1.020	0.411	0.526	35.411	0.0382	44.44
	SSD–LXK	3.896	7.04	1.01	−6.38	0.602	1.887	0.553	0.640	14.398	0.1129	31.14
$P_{0.90}^{ES}$	SSD–φ	2.077	3.89	1.07	−3.59	0.510	1.783	0.534	0.479	23.389	0.1042	30.12
	SSD–Δ	2.520	5.36	1.08	−5.89	0.548	1.360	0.470	0.492	34.955	0.0390	43.56
	SSD–LXK	5.657	12.23	0.88	−14.38	0.632	1.251	0.463	0.737	6.631	0.2006	23.15
$\tilde{P}_{1\times10^{-4}}^{ES}$	SSD–φ	4.489	9.90	1.29	−10.51	0.592	1.372	0.454	0.702	10.580	0.1414	26.94
	SSD–Δ	2.219	8.95	0.28	−13.61	0.595	0.462	0.248	0.510	37.483	0.0352	44.91
	SSD–LXK	4.806	10.91	1.16	−11.80	0.638	1.216	0.441	0.718	8.581	0.1535	26.67
$\tilde{P}_{2\times10^{-4}}^{ES}$	SSD–φ	3.545	6.64	0.92	−7.08	0.590	1.611	0.534	0.634	15.104	0.1262	30.00
	SSD–Δ	2.581	7.73	0.65	−10.39	0.589	0.725	0.334	0.538	36.052	0.0431	42.87
	SSD–LXK	4.464	8.81	1.15	−7.94	0.624	1.640	0.507	0.678	11.606	0.1202	28.70
$\tilde{P}_{3\times10^{-4}}^{ES}$	SSD–φ	2.673	5.08	1.13	−5.05	0.552	1.636	0.526	0.547	19.812	0.1141	28.98
	SSD–Δ	2.626	6.45	0.90	−8.37	0.576	0.995	0.407	0.524	35.573	0.0372	44.63

Table 4

Out-of-sample portfolio performance with probability sets based on empirical likelihood (EL) state probabilities, obtained by block size $B = 2$. The first four columns show the mean, standard deviation, skewness, and $\mathrm{CVaR}_{5\%}$ of annualized excess returns (spreads) to the benchmark portfolio. The next three columns under risk-adjusted metrics present annualized Sharpe, Sortino, and Information (Info.) ratios. The second last two columns under asset diversification report portfolio turnover and average number of industries (out of 49) included in the optimal portfolios. The last two columns present two dominance measures, almost-SSD (ASSD) and full SSDs over the benchmark.

Set P	Model	Spread over benchmark				Risk–adjusted metrics			Asset diversification		Dominance	
		Mean %	Std. %	Skew ness	$\mathrm{CVaR}_{5\%}$ %	Sharpe ratio	Sortino ratio	Info. ratio	Turnover	Industries # in 49	ASSD ε	SSDs %
P_1^{EL}	SSD–LXK	6.067	15.12	1.05	−16.80	0.596	1.060	0.401	0.747	5.117	0.2771	14.35
=	SSD–φ	5.569	13.28	1.04	−16.60	0.613	1.069	0.419	0.677	6.739	0.1783	24.26
$\tilde{P}_0^{EL}$	SSD–Δ	1.241	9.59	0.23	−16.33	0.543	0.212	0.129	0.415	40.033	0.0278	45.09
	SSD–LXK	5.390	11.81	0.97	−13.12	0.628	1.237	0.456	0.718	7.465	0.1748	24.54
$P_{0.96}^{EL}$	SSD–φ	4.039	7.85	1.09	−8.07	0.601	1.618	0.514	0.667	12.841	0.1334	28.06
	SSD–Δ	2.429	8.41	0.43	−12.18	0.598	0.572	0.289	0.528	36.907	0.0438	42.96
	SSD–LXK	4.479	8.80	1.17	−7.76	0.621	1.716	0.509	0.684	11.596	0.1199	30.00
$P_{0.92}^{EL}$	SSD–φ	2.693	4.93	1.01	−4.53	0.550	1.732	0.546	0.544	19.894	0.1082	29.72
	SSD–Δ	2.678	6.34	0.90	−7.76	0.583	1.070	0.422	0.524	35.513	0.0380	44.72
	SSD–LXK	3.966	6.95	0.95	−6.02	0.615	1.953	0.570	0.639	14.390	0.1155	31.02
$P_{0.90}^{EL}$	SSD–φ	2.129	3.79	1.10	−3.28	0.525	1.971	0.563	0.476	23.381	0.1011	30.28
	SSD–Δ	2.539	5.27	1.01	−5.68	0.562	1.372	0.482	0.490	34.930	0.0376	43.43
	SSD–LXK	5.773	12.41	1.00	−13.98	0.631	1.300	0.465	0.726	6.639	0.2045	23.06
$\tilde{P}_{1\times10^{-4}}^{EL}$	SSD–φ	4.380	9.94	1.22	−10.74	0.588	1.276	0.441	0.698	10.569	0.1430	27.31
	SSD–Δ	2.290	8.99	0.27	−13.73	0.599	0.476	0.255	0.506	37.685	0.0351	45.19
	SSD–LXK	4.769	11.01	1.11	−12.27	0.633	1.177	0.433	0.715	8.564	0.1520	27.59
$\tilde{P}_{2\times10^{-4}}^{EL}$	SSD–φ	3.529	6.60	0.95	−7.28	0.587	1.608	0.535	0.628	15.124	0.1216	30.28
	SSD–Δ	2.549	7.74	0.61	−10.48	0.586	0.705	0.329	0.533	36.119	0.0413	43.43
	SSD–LXK	4.404	8.87	1.18	−8.41	0.620	1.571	0.497	0.677	11.580	0.1273	28.89
$\tilde{P}_{3\times10^{-4}}^{EL}$	SSD–φ	2.707	4.95	1.05	−4.95	0.552	1.714	0.547	0.545	19.725	0.1122	29.81
	SSD–Δ	2.645	6.36	0.91	−8.03	0.578	1.038	0.416	0.521	35.548	0.0384	44.07

generally bring both active and downside risks persistently lower for all SSD models, SSD–φ excels in particular with respect to risk control in contrast to SSD–LXK and SSD–Δ models, especially in mitigating large negative returns in the downside as evaluated by $\mathrm{CVaR}_{5\%}$.

On risk-adjusted return measures, using SSD–LXK and SSD–φ models results in better performance in terms of Sortino and Information ratios compared to SSD–Δ model, for all probability sets constructed with equal and EL state probabilities. Moreover, using SSD–LXK model outperforms SSD–φ and SSD–Δ models on Sharpe ratios, when a large probability set as such is used (see, Table B.8 for test of significance[3] in Appendix B). More specifically, a higher Sharpe (Sortino) ratio implies that each unit of risk (in the downside) earns a larger spread, while a better Information ratio indicates that a higher spread is earned for each unit of active risk. In general, increasing the size of sets P_α^{ES}, P_α^{EL},

[3] In this setting, the bootstrap algorithm developed by Ledoit and Wolf (2008) allows for testing the statistical significance of Sharpe and Information ratios only.

$\tilde{P}^{ES}_{\beta}$, and $\tilde{P}^{EL}_{\beta}$, for all α and β values, leads to improved Sortino and Information ratios for SSD–Δ model but declined Sharpe ratios for SSD–φ model. For almost all SSD models, an increasing pattern in Sortino and Information ratios can be observed from applying a larger EL-based probability set P^{EL}_{α} or $\tilde{P}^{EL}_{\beta}$. However, one exception is noted for SSD–LXK model with probability set $\tilde{P}^{EL}_{2\times10^{-4}}$.

Generally, the use of SSD–Δ model reports the lowest portfolio turnovers for all probability sets utilizing equal and EL state probabilities throughout all α and β values, except for the two largest sets $P^{ES}_{0.90}$ and $P^{EL}_{0.90}$; however, it also requires the largest diversification across the underlying asset span (49 industries) in comparison with SSD–LXK and SSD–φ models. When the size of a particular probability set increases, applying SSD–LXK and SSD–φ models typically leads to lower portfolio turnovers, but increases the average number of industries in the optimal benchmark dominating portfolios. Conversely, a diminishing pattern in the average number of industries is observed from using SSD–Δ model with respect to a set size increase.

Moreover, utilizing SSD–Δ model achieves the strongest out-of-sample second-order stochastic dominance performance measured on two metrics, i.e., almost SSD (see, Tzeng et al., 2013; also Tsetlin et al., 2015) and full SSDs over the benchmark portfolio, respectively. Applying SSD–φ model provides the second best average ε values of almost SSD and likelihood of obtaining full SSDs out-of-sample, as compared with the modest performer SSD–LXK. On one hand, for SSD–LXK and SSD–φ models, in general, increasing the size of all probability sets established by equal and EL state probabilities, for all α and β values, results in a greater likelihood of obtaining full dominance over the benchmark by SSD, with one exception found though with sets $\tilde{P}^{ES}_{3\times10^{-4}}$ and $\tilde{P}^{EL}_{3\times10^{-4}}$. On the other hand, however, the greatest likelihood of full SSDs is obtained by SSD–Δ model with the smallest probability set $P^{ES}_{1} = \tilde{P}^{ES}_{0}$. Interestingly, using SSD–Δ model exhibits no consistent dominance pattern based on the almost SSD or full SSDs metrics with regard to a size increase of the probability set.

In the top panel, Fig. 2 illustrates some of the resulting out-of-sample portfolio performance obtained by applying these SSD approaches with sets P^{ES}_{α} and $\tilde{P}^{ES}_{\beta}$ (see, Table 3), while that obtained with sets P^{EL}_{α} and $\tilde{P}^{EL}_{\beta}$ (see, Table 4) is shown in the panel on the bottom. Specifically, in each subplot, we use blue stars, red circles, and black crosses to denote the out-of-sample mean returns of spread and full dominance by SSD obtained by using SSD–LXK, SSD–φ, and SSD–Δ models, respectively. Moreover, the blue, red, and black arrows point to indicate a size increase of the underlying probability set for SSD–LXK, SSD–φ, and SSD–Δ models, respectively.

Finally, we observe a clear trade-off between out-of-sample portfolio returns and dominance over the benchmark for all SSD models with all probability sets. Hence, this result suggests that it is impossible to systematically improve on out-of-sample SSDs without making trade-offs among different levels of out-of-sample returns. In addition, we notice that the red arrows are remarkably steeper than the blue ones. This finding implies that the out-of-sample spreads earned by utilizing SSD–φ model decline at a faster rate than SSD–LXK model due to improved out-of-sample SSDs.

5.5. Sensitivity of results to block size

In order to utilize probability sets with empirical likelihood (EL) state probabilities, block size B needs to be specified as a parameter in its estimation procedure. To evaluate the impact of varying block sizes on out-of-sample portfolio performance, we replicate the empirical tests in Section 5.2 carried out with the block size $B = 2$ by using another three block sizes $B \in \{3, 5, 10\}$. Specifically, these tests are carried out by solving first optimization problem (28)–(30) to obtain the EL state probabilities p^{EL} for each block size B, then constructing different probability sets P based on these EL-estimate state probabilities. In what follows, we identify robust dominating portfolios by SSD by solving LP models SSD–LXK, SSD–φ, and SSD–Δ with these established probability sets P^{EL}_{α} and $\tilde{P}^{EL}_{\beta}$, for $\alpha \in \{1.00, 0.96, 0.92, 0.90\}$ and $\beta \in \{0, 1\times10^{-4}, 2\times10^{-4}, 3\times10^{-4}\}$. The results obtained for each block size B are presented in Tables 5–7.

In general, when block size B increases, for $B \in \{3, 5, 10\}$, using SSD–LXK model earns still the highest spreads, followed by SSD–φ and SSD–Δ models with sets P^{EL}_{α} and $\tilde{P}^{EL}_{\beta}$ across all α and β values. With the smallest probability set $P^{EL}_{1} = \tilde{P}^{EL}_{0}$, the use of SSD–$\Delta$ model remains to give the lowest active and downside risks, as well as the best likelihood of obtaining full SSDs. For all α and β values with sets P^{EL}_{α} and $\tilde{P}^{EL}_{\beta}$, applying SSD–LXK and SSD–φ models keeps outperforming SSD–Δ model on risk-adjusted metrics such as Sortino and Information ratios, whereas the best Sharpe ratios are still earned by SSD–LXK model with large EL-based probability sets (see, Table B.8 for test of significance in Appendix B). The outperformance on dominance related metrics such as almost SSD and full SSDs comes from utilizing SSD–Δ model. Moreover, increasing the size of an EL-based probability set improves on risk-adjusted metrics consistently, for instance, Sortino and Information ratios for almost all SSD models, as well as on dominance metrics, e.g., almost SSD for all SSD–LXK and SSD–φ models. Notably, this finding suggests that the out-of-sample portfolio performance of all SSD models is robust to variations in block size B for EL-based probability sets.

Additionally, admitting incomplete information on equal and empirical likelihood (EL) state probabilities for all models SSD–LXK, SSD–φ, and SSD–Δ reflects no consistent portfolio performance pattern out-of-sample. However, under complete probability information, using set $P^{EL}_{1} = \tilde{P}^{EL}_{0}$ outperforms set $P^{ES}_{1} = \tilde{P}^{ES}_{0}$ for SSD–LXK model. Specifically, using EL state probabilities can earn a larger spread up to 6.569% over the benchmark portfolio in contrast to 5.934% with equal state probabilities. Consequently, this brings us to conclude that EL state probabilities can be used to better support portfolio return maximization models based on SSD, when complete information on state probabilities is available.

6. Conclusions and discussion

In this paper, we have developed portfolio optimization approaches based on novel robust dominance criteria. In particular, these robust criteria are designed to maximize the minimum distances between pairs of integrated CDFs of portfolio returns. Therefore, portfolio optimization models utilizing these new criteria can identify portfolios whose in-sample dominance over the benchmark is the strongest. We also suggested a novel hybrid method to admit incomplete probability information by combining the existing approaches of Post et al. (2018) and Liesiö et al. (2020). Specifically, this hybrid method first estimates a single vector of state probabilities using the Empirical Likelihood (EL) approach (Post et al., 2018). Then, it builds up a confidence region of feasible probability vectors around this point-estimate state probabilities.

The developed SSD optimization approaches were tested through an empirical application by evaluating out-of-sample performance of optimized industrial portfolios. In particular, we compared the out-of-sample performance of these in-sample optimized portfolios obtained with several different probability sets based on deploying equal and EL state probabilities.

Our results from the empirical application document an explicit trade-off pattern between returns and stochastic dominance relations out-of-sample. For SSD-based portfolio optimization models, in general, utilizing robust dominance measures based on distances between the integrated CDFs of portfolios returns accomplishes higher likelihood of obtaining out-of-sample dominance over the benchmark portfolio in the sense of SSD. However, such optimal portfolios also earn modestly lower returns out-of-sample. On the other hand, SSD approaches that identify return maximizing portfolios in-sample naturally obtain the best out-of-sample returns, however, at the cost of losing future dominance over the benchmark. Moreover, admitting incomplete probability

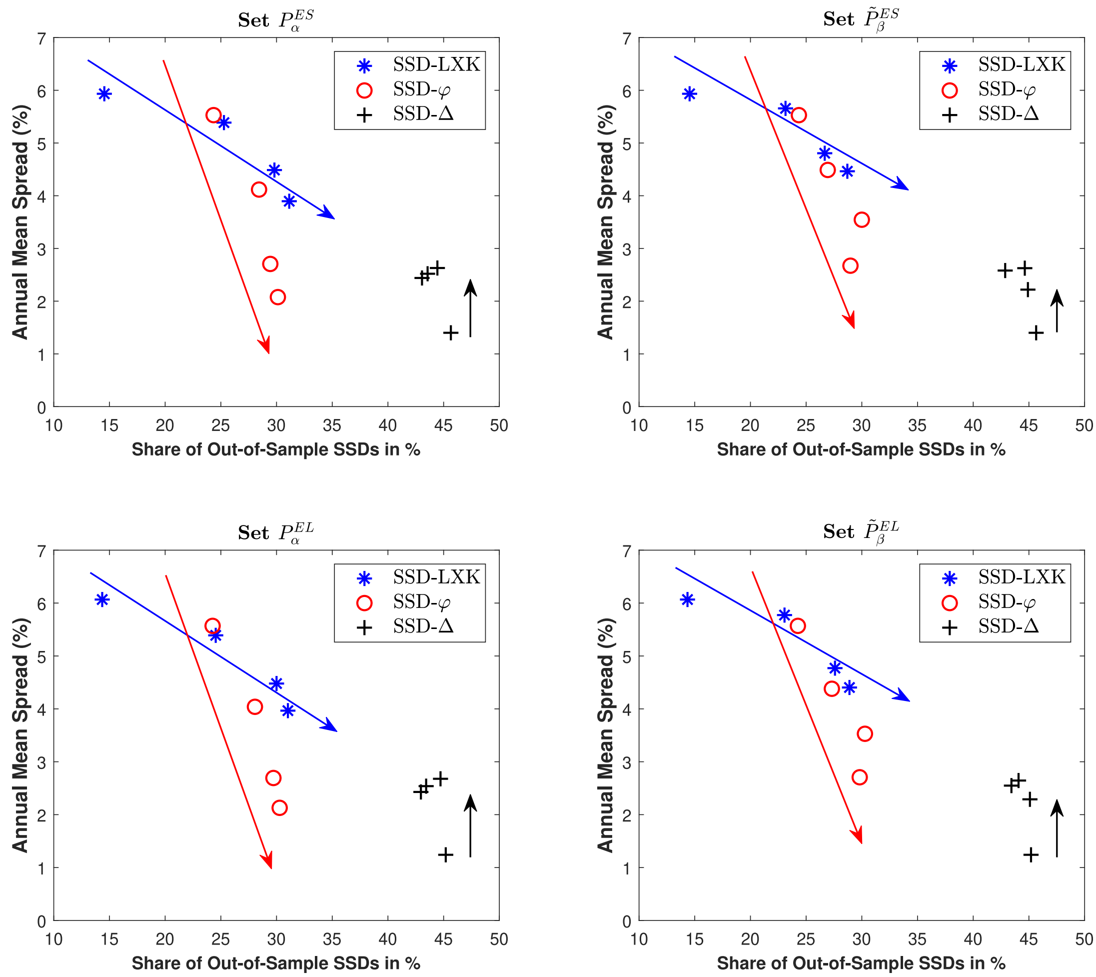

Fig. 2. Illustration of out-of-sample mean returns of spread and obtained shares of full dominance by SSD from the use of different SSD models with different probability sets based on equal and empirical likelihood state probabilities. Arrows point to indicate a larger size of probability sets.

Table 5
Out-of-sample portfolio performance with probability sets based on empirical likelihood (EL) state probabilities, obtained by block size $B = 3$. The first four columns show the mean, standard deviation, skewness, and $\text{CVaR}_{5\%}$ of annualized excess returns (spreads) to the benchmark portfolio. The next three columns under risk-adjusted metrics present annualized Sharpe, Sortino, and Information (Info.) ratios. The second last two columns under asset diversification report portfolio turnover and average number of industries (out of 49) included in the optimal portfolios. The last two columns present two dominance measures, almost-SSD (ASSD) and full SSDs over the benchmark.

Set P	Model	Spread over benchmark				Risk–adjusted metrics			Asset diversification		Dominance	
		Mean %	Std. %	Skew ness	$\text{CVaR}_{5\%}$ %	Sharpe ratio	Sortino ratio	Info. ratio	Turnover	Industries # in 49	ASSD ε	SSDs %
P^{EL}_{1}	SSD–LXK	6.262	15.38	1.08	−17.50	0.600	1.072	0.407	0.743	5.103	0.2742	14.91
=	SSD–φ	5.401	13.06	0.83	−17.03	0.613	1.002	0.414	0.675	6.760	0.1788	24.17
$\tilde{P}^{EL}_{0}$	SSD–Δ	1.387	9.60	0.18	−16.73	0.546	0.237	0.144	0.452	39.119	0.0271	45.65
	SSD–LXK	5.430	12.07	1.04	−13.54	0.628	1.228	0.450	0.714	7.462	0.1730	24.63
$P^{EL}_{0.96}$	SSD–φ	4.031	7.94	1.10	−8.13	0.599	1.592	0.508	0.661	12.844	0.1330	27.69
	SSD–Δ	2.472	8.50	0.45	−12.30	0.600	0.580	0.291	0.523	36.928	0.0453	43.24
	SSD–LXK	4.466	8.75	1.11	−7.73	0.620	1.718	0.510	0.681	11.565	0.1216	30.56
$P^{EL}_{0.92}$	SSD–φ	2.669	4.93	1.08	−4.44	0.547	1.730	0.541	0.543	19.894	0.1120	30.09
	SSD–Δ	2.655	6.33	0.86	−7.58	0.580	1.064	0.420	0.520	35.471	0.0385	44.81
	SSD–LXK	3.889	6.92	0.94	−6.30	0.612	1.854	0.562	0.637	14.387	0.1149	30.93
$P^{EL}_{0.90}$	SSD–φ	2.140	3.77	1.05	−3.38	0.525	1.936	0.568	0.475	23.401	0.1069	30.74
	SSD–Δ	2.573	5.25	1.02	−5.70	0.563	1.406	0.490	0.487	34.891	0.0377	44.07
	SSD–LXK	5.699	12.62	0.95	−14.92	0.626	1.217	0.452	0.720	6.647	0.2038	23.24
$\tilde{P}^{EL}_{1\times10^{-4}}$	SSD–φ	4.418	9.99	1.30	−10.57	0.589	1.308	0.442	0.695	10.580	0.1490	26.57
	SSD–Δ	2.261	9.02	0.31	−13.91	0.596	0.471	0.251	0.504	37.643	0.0356	44.72
	SSD–LXK	4.882	11.15	1.16	−12.22	0.635	1.214	0.438	0.707	8.570	0.1562	28.06
$\tilde{P}^{EL}_{2\times10^{-4}}$	SSD–φ	3.602	6.62	0.94	−7.30	0.588	1.664	0.544	0.625	15.090	0.1220	31.20
	SSD–Δ	2.583	7.81	0.66	−10.52	0.589	0.718	0.331	0.533	36.068	0.0431	42.78
	SSD–LXK	4.454	8.85	1.13	−8.14	0.625	1.627	0.503	0.674	11.526	0.1298	29.26
$\tilde{P}^{EL}_{3\times10^{-4}}$	SSD–φ	2.680	4.92	1.12	−4.80	0.551	1.711	0.545	0.545	19.738	0.1093	30.19
	SSD–Δ	2.668	6.34	0.87	−8.01	0.577	1.046	0.421	0.520	35.490	0.0380	43.52

Table 6

Out-of-sample portfolio performance with probability sets based on empirical likelihood (EL) state probabilities, obtained by block size $B = 5$. The first four columns show the mean, standard deviation, skewness, and $\text{CVaR}_{5\%}$ of annualized excess returns (spreads) to the benchmark portfolio. The next three columns under risk-adjusted metrics present annualized Sharpe, Sortino, and Information (Info.) ratios. The second last two columns under asset diversification report portfolio turnover and average number of industries (out of 49) included in the optimal portfolios. The last two columns present two dominance measures, almost-SSD (ASSD) and full SSDs over the benchmark.

Set P	Model	Spread over benchmark				Risk–adjusted metrics			Asset diversification		Dominance	
		Mean %	Std. %	Skewness	$\text{CVaR}_{5\%}$ %	Sharpe ratio	Sortino ratio	Info. ratio	Turnover	Industries # in 49	ASSD ε	SSDs %
P_1^{EL}	SSD–LXK	6.201	15.62	1.27	−16.75	0.589	1.085	0.397	0.736	5.072	0.2840	14.54
=	SSD–φ	5.515	13.02	0.88	−16.90	0.615	1.053	0.423	0.663	6.744	0.1766	23.98
$\tilde{P}_0^{EL}$	SSD–Δ	1.408	9.49	0.22	−16.26	0.544	0.246	0.148	0.426	40.107	0.0280	44.35
	SSD–LXK	5.519	12.37	1.01	−14.30	0.629	1.198	0.446	0.704	7.470	0.1778	24.81
$P_{0.96}^{EL}$	SSD–φ	4.056	8.00	1.15	−8.09	0.596	1.611	0.507	0.655	12.822	0.1325	27.96
	SSD–Δ	2.513	8.60	0.46	−12.19	0.601	0.586	0.292	0.516	36.808	0.0437	43.06
	SSD–LXK	4.425	8.82	1.10	−7.66	0.615	1.663	0.502	0.673	11.563	0.1234	29.63
$P_{0.92}^{EL}$	SSD–φ	2.662	4.91	1.06	−4.43	0.546	1.725	0.542	0.539	19.894	0.1101	31.20
	SSD–Δ	2.682	6.32	0.85	−7.48	0.581	1.079	0.424	0.516	35.325	0.0394	44.17
	SSD–LXK	3.754	6.93	0.93	−6.53	0.606	1.723	0.542	0.630	14.354	0.1131	30.46
$P_{0.90}^{EL}$	SSD–φ	2.095	3.73	1.05	−3.38	0.522	1.875	0.561	0.473	23.388	0.1092	31.20
	SSD–Δ	2.552	5.21	0.99	−5.68	0.563	1.396	0.490	0.484	34.875	0.0365	43.80
	SSD–LXK	5.805	12.79	1.05	−14.63	0.629	1.267	0.454	0.715	6.630	0.2022	23.15
$\tilde{P}_{1\times10^{-4}}^{EL}$	SSD–φ	4.330	10.08	1.32	−10.60	0.587	1.247	0.430	0.689	10.546	0.1464	26.30
	SSD–Δ	2.305	9.04	0.34	−13.90	0.599	0.484	0.255	0.497	37.696	0.0369	43.89
	SSD–LXK	4.905	11.29	1.10	−12.48	0.633	1.183	0.435	0.700	8.571	0.1656	26.48
$\tilde{P}_{2\times10^{-4}}^{EL}$	SSD–φ	3.582	6.59	0.91	−7.28	0.589	1.653	0.543	0.622	15.041	0.1236	30.46
	SSD–Δ	2.520	7.81	0.67	−10.67	0.586	0.699	0.323	0.529	36.122	0.0404	43.43
	SSD–LXK	4.390	8.87	1.10	−8.12	0.618	1.583	0.495	0.668	11.536	0.1344	29.26
$\tilde{P}_{3\times10^{-4}}^{EL}$	SSD–φ	2.624	4.85	1.12	−4.77	0.549	1.682	0.541	0.539	19.728	0.1074	30.00
	SSD–Δ	2.676	6.35	0.86	−7.73	0.577	1.053	0.421	0.514	35.544	0.0390	44.63

Table 7

Out-of-sample portfolio performance with probability sets based on empirical likelihood (EL) state probabilities, obtained by block size $B = 10$. The first four columns show the mean, standard deviation, skewness, and $\text{CVaR}_{5\%}$ of annualized excess returns (spreads) to the benchmark portfolio. The next three columns under risk-adjusted metrics present annualized Sharpe, Sortino, and Information (Info.) ratios. The second last two columns under asset diversification report portfolio turnover and average number of industries (out of 49) included in the optimal portfolios. The last two columns present two dominance measures, almost-SSD (ASSD) and full SSDs over the benchmark.

Set P	Model	Spread over benchmark				Risk–adjusted metrics			Asset diversification		Dominance	
		Mean %	Std. %	Skewness	$\text{CVaR}_{5\%}$ %	Sharpe ratio	Sortino ratio	Info. ratio	Turnover	Industries # in 49	ASSD ε	SSDs %
P_1^{EL}	SSD–LXK	6.569	15.81	1.40	−16.07	0.592	1.219	0.415	0.721	5.073	0.2829	15.28
=	SSD–φ	5.641	13.24	1.00	−16.98	0.613	1.115	0.426	0.650	6.719	0.1774	24.26
$\tilde{P}_0^{EL}$	SSD–Δ	1.320	9.58	0.19	−16.53	0.550	0.226	0.138	0.421	39.760	0.0342	44.72
	SSD–LXK	5.326	12.32	1.06	−14.33	0.622	1.157	0.432	0.694	7.476	0.1799	23.52
$P_{0.96}^{EL}$	SSD–φ	3.992	7.78	1.15	−7.57	0.596	1.671	0.513	0.648	12.824	0.1386	28.24
	SSD–Δ	2.568	8.49	0.52	−12.12	0.604	0.622	0.302	0.506	37.081	0.0426	42.78
	SSD–LXK	4.211	8.94	1.13	−8.28	0.607	1.506	0.471	0.656	11.506	0.1268	29.63
$P_{0.92}^{EL}$	SSD–φ	2.629	4.74	0.91	−4.55	0.544	1.713	0.554	0.530	19.879	0.1051	30.83
	SSD–Δ	2.630	6.23	0.77	−7.63	0.576	1.059	0.422	0.504	35.370	0.0397	44.26
	SSD–LXK	3.652	6.85	0.91	−6.74	0.596	1.656	0.533	0.618	14.358	0.1171	30.83
$P_{0.90}^{EL}$	SSD–φ	1.994	3.62	0.88	−3.51	0.519	1.749	0.551	0.462	23.370	0.1060	31.11
	SSD–Δ	2.499	5.19	0.92	−5.91	0.558	1.350	0.482	0.476	34.967	0.0399	42.78
	SSD–LXK	5.765	13.19	1.19	−14.35	0.618	1.236	0.437	0.700	6.665	0.2075	22.96
$\tilde{P}_{1\times10^{-4}}^{EL}$	SSD–φ	4.524	10.05	1.36	−10.73	0.591	1.362	0.450	0.674	10.586	0.1525	26.11
	SSD–Δ	2.385	9.03	0.34	−13.88	0.602	0.507	0.264	0.483	37.769	0.0382	44.44
	SSD–LXK	4.762	11.21	1.21	−11.41	0.618	1.214	0.425	0.689	8.591	0.1672	25.65
$\tilde{P}_{2\times10^{-4}}^{EL}$	SSD–φ	3.337	6.53	0.85	−7.11	0.578	1.494	0.511	0.618	15.083	0.1227	29.72
	SSD–Δ	2.539	7.75	0.63	−10.75	0.589	0.711	0.328	0.516	36.239	0.0423	43.52
	SSD–LXK	4.106	8.78	1.09	−8.44	0.607	1.443	0.468	0.654	11.485	0.1294	29.17
$\tilde{P}_{3\times10^{-4}}^{EL}$	SSD–φ	2.586	4.76	1.00	−4.82	0.547	1.642	0.543	0.535	19.684	0.1107	30.28
	SSD–Δ	2.642	6.24	0.83	−7.55	0.574	1.072	0.423	0.505	35.368	0.0409	44.07

information and applying set-valued state probabilities in some SSD models lead to improved performance in achieving full SSDs out-of-sample; however, this finding does not necessarily hold true for all models. In addition, our results indicate that SSD-based portfolio optimization models can benefit from the use of EL state probabilities, which take into account empirical stylized facts about returns of various asset classes.

The developed approaches for identifying robust dominating portfolios are also of practical significance in view of real-world applications. They can be readily applied to support financial decision making under uncertainty. Specifically, using these models can be beneficial in decision situations where portfolio managers wish to construct robust portfolios that are not sensitive to uncertain, volatile variations in future asset returns, but where well-established preference and/or probability information is unavailable or difficult to obtain. In general, these methods can be effective risk management tools under market turmoil for fund managers facing an uncertain market outlook.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No.

Acknowledgments

The author would like to gratefully thank the editor, anonymous reviewers, Professors Juuso Liesiö, Jyrki Wallenius, and Tri M. Tran for insightful and constructive comments and suggestions. The computational resources on Triton computing cluster provided by the Aalto Science-IT project are also acknowledged.

Funding

This research was supported by the Foundation for Economic Education (Liikesivistysrahasto) [Grant no. 170429, 200316], the HSE Support Foundation [Grant no. 7-1415], and Marcus Wallenberg Foundation.

Appendix A. Proofs of Lemma and Theorems

Proof of Lemma 1. Consider function $g(\theta) = F_V^2(\theta;p) - F_W^2(\theta;p)$ for an arbitrary $p \in \text{ext}(\text{conv}(P))$. This is a piece-wise linear function such that $g(-\infty) = 0$ and $g(\infty) = C$, where C is some positive constant. Then, $g(\theta)$ attains its minimum at some $\theta \in \arg\min_\theta g(\theta) = \{v_1,\dots,v_T,w_1,\dots,w_T\}$. Suppose that $\min_\theta g(\theta) = g(w_t)$ for some w_t such that $w_t \neq v_s$, for any $s \in \{1,\dots,T\}$ and consider a small neighborhood of returns around w_t such that $[w_t-\varepsilon,\ w_t+\varepsilon] \cap \{v_1,\dots,v_T\} = \emptyset$. Then, evaluate the derivative of $g(\theta)$ at $w_t-\varepsilon$ and $w_t+\varepsilon$, respectively, which yields

$$\frac{\partial g(w_t-\varepsilon)}{\partial\theta} = F_V(w_t-\varepsilon;p) - F_W(w_t-\varepsilon;p), \tag{A.1}$$

$$\frac{\partial g(w_t+\varepsilon)}{\partial\theta} = F_V(w_t+\varepsilon;p) - F_W(w_t+\varepsilon;p). \tag{A.2}$$

Since $F_V(w_t-\varepsilon;p) = F_V(w_t+\varepsilon;p)$ and $F_W(\theta;p)$ is increasing in θ, we obtain $\frac{\partial g(w_t-\varepsilon)}{\partial\theta} \geq \frac{\partial g(w_t+\varepsilon)}{\partial\theta}$, and therefore $g(\theta)$ is concave on the interval $[w_t-\varepsilon,\ w_t+\varepsilon]$ and hence $g(\theta)$ cannot obtain its minimum at w_t. Thus, the minimum of $g(\theta)$ is obtained at some $\theta = v_t$. Hence, we have

$$\min_\theta \left(F_V^2(\theta;p) - F_W^2(\theta;p)\right) = \min_\theta g(\theta) = \min_t g(v_t) = \min_t \left(F_V^2(v_t;p) - F_W^2(v_t;p)\right). \tag{A.3}$$

Now, assume that $F_V^2(v_t;p) \geq F_W^2(v_t;p)$ for all $t \in \{1,\dots,T\}$ and $p \in \text{ext}(\text{conv}(P))$. Then, Eq. (A.3) implies that $F_V^2(\theta;p) \geq F_W^2(\theta;p)$ for all $\theta \in \mathbb{R}$ and $p \in \text{ext}(\text{conv}(P))$. □

Proof of Theorem 1. (i) Take any feasible solution $X = \sum_{i=1}^N \lambda_i X_i$ and φ such that (X,φ) is a feasible solution to optimization problem (11), i.e., $X \geq_P Y+\varphi$. Constructing z such that $z_{t,s} = \max\{y_t + \varphi - \sum_{i=1}^N \lambda_i x_{i,s},\ 0\}$ for each $t,s \in \{1,\dots,T\}$, results in $z \in \mathbb{R}_+^{T\times T}$ that satisfies constraint (18). To show that z satisfies also constraint (19), we evaluate the LHS of (19) for an arbitrary $t \in \{1,\dots,T\}$ and $p \in \text{ext}(\text{conv}(P))$, which gives $\sum_{s=1}^T p_s z_{t,s} = \sum_{s=1}^T p_s \max\{y_t+\varphi - \sum_{i=1}^N \lambda_i x_{i,s},\ 0\} = \sum_{s=1}^T p_s \max\{y_t+\varphi-x_s,\ 0\} \overset{(5)}{=} F_X^2(y_t+\varphi;p)$. Since (X,φ) is a feasible solution to optimization problem (11), $X \geq_P Y+\varphi$ holds and Theorem 1 of Liesiö et al. (2020) implies that $F_X^2(y_t+\varphi;p) \leq F_{Y+\varphi}^2(y_t+\varphi;p)$ for all $t \in \{1,\dots,T\}$ and $p \in \text{ext}(\text{conv}(P))$. By Eq. (5), evaluating the RHS of this inequality gives

$$\begin{aligned} F_{Y+\varphi}^2(y_t+\varphi;p) &= \int_{-\infty}^{y_t+\varphi} F_{Y+\varphi}(\tau;p)\,d\tau = \sum_{\substack{s \\ y_s+\varphi\leq y_t+\varphi}} p_s\left[y_t+\varphi-(y_s+\varphi)\right] \\ &= \sum_{\substack{s \\ y_s\leq y_t}} p_s(y_t-y_s) \\ &= \sum_{t=1}^T p_s \max\{y_t-y_s,\ 0\} = F_Y^2(y_t;p), \end{aligned} \tag{A.4}$$

which is the RHS of constraint (19). Hence, (λ,z,φ) is a feasible solution to LP problem (17)–(19).

(ii) Take any feasible solution (λ,z,φ) to LP problem (17)–(19). Constraint (18) then implies that $z_{t,s} \geq \max\{y_t+\varphi-\sum_{i=1}^N \lambda_i x_{i,s},\ 0\}$ holds. Thus, evaluating constraint (19) for an arbitrary $t \in \{1,\dots,T\}$ and $p \in \text{ext}(\text{conv}(P))$ yields

$$\begin{aligned} \underbrace{F_Y^2(y_t;p)}_{\overset{(A.4)}{=} F_{Y+\varphi}^2(y_t+\varphi;p)} &\geq \sum_{s=1}^T p_s z_{t,s} \geq \sum_{s=1}^T p_s \max\left\{y_t+\varphi-\sum_{i=1}^N \lambda_i x_{i,s},\ 0\right\} \\ &= \sum_{s=1}^T p_s \max\{y_t+\varphi-x_s,\ 0\} \overset{(5)}{=} F_X^2(y_t+\varphi;p), \end{aligned}$$

where $X = \sum_{i=1}^N \lambda_i X_i$. Thus, we have $F_{Y+\varphi}^2(y_t+\varphi;p) \geq F_X^2(y_t+\varphi;p)$ for all $t \in \{1,\dots,T\}$ and $p \in \text{ext}(\text{conv}(P))$, which by Lemma 1 implies that $F_{Y+\varphi}^2(\theta;p) \geq F_X^2(\theta;p)$ holds for all $\theta \in \mathbb{R}$ and $p \in \text{ext}(\text{conv}(P))$. Then, Theorem 1 of Liesiö et al. (2020) implies that $X \geq_P Y+\varphi$ holds. Hence, (X,φ) is a feasible solution to optimization problem (11).

Now, let $\left(\sum_{i=1}^N \lambda_i X_i, \varphi\right)$ be an optimal solution to (11). Then, by (i), there exists a feasible solution (λ,z,φ) to LP problem (17)–(19). Suppose that this solution is not optimal to (17)–(19), in which case then there exists another feasible solution (λ',z',φ') with a higher objective function value, i.e., $\varphi' > \varphi$. Then, by (ii), $\left(\sum_{i=1}^N \lambda_i' X_i, \varphi'\right)$ is also a feasible solution to (11), which contradicts the assumption that $\left(\sum_{i=1}^N \lambda_i X_i, \varphi\right)$ is the optimal solution to (11). Hence, (λ,z,φ) is an optimal solution to (17)–(19).

In turn, let (λ,z,φ) be an optimal solution to LP problem (17)–(19). Then, by (ii), $\left(\sum_{i=1}^N \lambda_i X_i, \varphi\right)$ is a feasible solution to optimization problem (11). Assume that $\left(\sum_{i=1}^N \lambda_i X_i, \varphi\right)$ is not optimal to (11). Then, there must exist another feasible solution $\left(\sum_{i=1}^N \lambda_i' X_i, \varphi'\right)$ yielding a higher objective function value $\varphi' > \varphi$. Then, by (i), there exists z' such that (λ',z',φ') is also feasible to (17)–(19), which contradicts the assumption that (λ,z,φ) is the optimal solution to (17)–(19). Therefore, $\left(\sum_{i=1}^N \lambda_i X_i, \varphi\right)$ is an optimal solution to (11). □

Proof of Theorem 2. (i) Take any feasible solution $X = \sum_{i=1}^N \lambda_i X_i$ to optimization problem (13), i.e., $X \geq_P Y$. Then $\Delta(X,Y) \overset{(12)}{=} \min_{\theta\geq y_2,\ p\in P} F_Y^2(\theta;p) - F_X^2(\theta;p) \geq 0$, where the inequality follows from the fact that $X \geq_P Y$ implies $F_Y^2(\theta;p) - F_X^2(\theta;p) \geq 0$ for all $\theta \in \mathbb{R}$ and $p \in P$ by Definition 2. Then, setting $\delta = \Delta(X,Y)$ and $z_{t,s} = \max\{y_t - \sum_{i=1}^N \lambda_i x_{i,s},\ 0\}$ for each $t,s \in \{1,\dots,T\}$ gives $(z,\delta) \in \mathbb{R}_+^{T\times T} \times \mathbb{R}$ that clearly satisfies constraint (21) and $\delta \geq 0$. To show that z satisfies first constraint (22), we evaluate its LHS for an arbitrary $p \in \text{ext}(\text{conv}(P))$, which yields $\sum_{s=1}^T p_s z_{1,s} = \sum_{s=1}^T p_s \max\{y_1 - \sum_{i=1}^N \lambda_i x_{i,s},\ 0\} = \sum_{s=1}^T p_s \max\{y_1 - x_s,\ 0\} \overset{(5)}{=} F_X^2(y_1;p)$. Since $X = \sum_{i=1}^N \lambda_i X_i$ is a feasible solution to optimization problem (13), $X \geq_P Y$ holds, then it must hold also $y_1 = \min(Y) \leq \min(X)$. This condition gives $F_Y^2(y_1;p) = F_X^2(y_1;p) = 0$, which then satisfies constraint (22). To show that (z,δ) satisfies constraint (23), we evaluate the LHS of (23) for an arbitrary $t \in \{2,\dots,T\}$ and $p \in \text{ext}(\text{conv}(P))$, which gives $\sum_{s=1}^T p_s z_{t,s} + \delta = \sum_{s=1}^T p_s \max\{y_t - \sum_{i=1}^N \lambda_i x_{i,s},\ 0\} + \delta = \sum_{s=1}^T p_s \max\{y_t - x_s,\ 0\} + \delta \overset{(5)}{=} F_X^2(y_t;p) + \delta$. Substituting $\delta = \Delta(X,Y)$ gives

$$F_X^2(y_t;p) + \delta = F_X^2(y_t;p) + \Delta(X,Y) = F_X^2(y_t;p)$$

Table B.8
Statistical significance of Sharpe and Information ratios in Tables 3–7.

Prob.		Equal state (ES) $\bar{p}$		Empirical likelihood (EL) state p^{EL}							
				$B=2$		$B=3$		$B=5$		$B=10$	
Set P	Model	Sharpe ratio	Info. ratio	Sharpe ratio	Info. ratio	Sharpe ratio	Info. ratio	Sharpe ratio	Info. ratio	Sharpe ratio	Info. ratio
$P_1^{(\cdot)}$	SSD–LXK	0.599	0.387**	0.596	0.401***	0.600	0.407***	0.589	0.397***	0.592	0.415***
=	SSD–φ	0.624	0.424***	0.613	0.419***	0.613	0.414***	0.615	0.423***	0.613	0.426***
$\tilde{P}_0^{(\cdot)}$	SSD–Δ	0.550	0.148	0.543	0.129	0.546	0.144	0.544	0.148	0.550	0.138
	SSD–LXK	0.636	0.464***	0.628	0.456**	0.628	0.450**	0.629	0.446**	0.622	0.432**
$P_{0.96}^{(\cdot)}$	SSD–φ	0.604	0.526***	0.601	0.514***	0.599	0.508***	0.596	0.507***	0.596	0.513***
	SSD–Δ	0.597	0.288	0.598	0.289	0.600	0.291	0.601	0.292	0.604	0.302
	SSD–LXK	0.621*	0.511**	0.621*	0.509*	0.620*	0.510**	0.615*	0.502*	0.607	0.471
$P_{0.92}^{(\cdot)}$	SSD–φ	0.550	0.536***	0.550	0.546***	0.547	0.541**	0.546	0.542**	0.544	0.554***
	SSD–Δ	0.580	0.411	0.583	0.422	0.580	0.420	0.581	0.424	0.576	0.422
	SSD–LXK	0.602***	0.553**	0.615***	0.570**	0.612***	0.562**	0.606***	0.542*	0.596***	0.533
$P_{0.90}^{(\cdot)}$	SSD–φ	0.510	0.534*	0.525	0.563**	0.525	0.568*	0.522	0.561*	0.519	0.551*
	SSD–Δ	0.548	0.470	0.562	0.482	0.563	0.490	0.563	0.490	0.558	0.482
	SSD–LXK	0.632	0.463***	0.631	0.465***	0.626	0.452***	0.629	0.454**	0.618	0.437**
$\tilde{P}_{1\times10^{-4}}^{(\cdot)}$	SSD–φ	0.592	0.454**	0.588	0.441**	0.589	0.442**	0.587	0.430**	0.591	0.450**
	SSD–Δ	0.595	0.248	0.599	0.255	0.596	0.251	0.599	0.255	0.602	0.264
	SSD–LXK	0.638*	0.441**	0.633	0.433*	0.635*	0.438**	0.633	0.435**	0.618	0.425*
$\tilde{P}_{2\times10^{-4}}^{(\cdot)}$	SSD–φ	0.590	0.534***	0.587	0.535***	0.588	0.544***	0.589	0.543***	0.578	0.511***
	SSD–Δ	0.589	0.334	0.586	0.329	0.589	0.331	0.586	0.323	0.589	0.328
	SSD–LXK	0.624**	0.507*	0.620*	0.497*	0.625**	0.503*	0.618*	0.495*	0.607	0.468
$\tilde{P}_{3\times10^{-4}}^{(\cdot)}$	SSD–φ	0.552	0.526***	0.552	0.547***	0.551	0.545**	0.549	0.541**	0.547	0.543**
	SSD–Δ	0.576	0.407	0.578	0.416	0.577	0.421	0.577	0.421	0.574	0.423

***,**,* indicate that the Sharpe ratio reported from the corresponding SSD–LXK model with set P (or $\tilde{P}$) is statistically greater than the resulting Sharpe ratios earned by SSD–φ and SSD–Δ models at the 1%, 5%, and 10% significance levels, respectively.
***,**,* indicate that the Information ratio reported from the corresponding SSD–LXK or SSD–φ model with set P (or $\tilde{P}$) is statistically greater than the resulting Information ratio earned by SSD–Δ model at the 1%, 5%, and 10% significance levels, respectively.

$$
\begin{aligned}
&+ \min_{\theta\geq y_2,\ p'\in P} \left(F_Y^2(\theta;p') - F_X^2(\theta;p') \right) \\
&\leq F_X^2(y_t;p) + F_Y^2(y_t;p) - F_X^2(y_t;p) \\
&= F_Y^2(y_t;p),
\end{aligned}
$$

which is the RHS of constraint (23). Hence, (λ, z, δ), where $\delta = \Delta\left(\sum_{i=1}^N \lambda_i X_i, Y\right)$, is a feasible solution to LP problem (20)–(23).

(ii) Take any feasible solution (λ, z, δ) to LP problem (20)–(23). Constraint (21) implies then $z_{t,s} \geq \max\left\{y_t - \sum_{i=1}^N \lambda_i x_{i,s},\ 0\right\}$ holds. Then, for an arbitrary $p \in \mathrm{ext}(\mathrm{conv}(P))$, evaluate constraint (22), which yields then

$$
\begin{aligned}
F_Y^2(y_1;p) &\geq \sum_{s=1}^T p_s z_{1,s} \geq \sum_{s=1}^T p_s \max\left\{ y_1 - \sum_{i=1}^N \lambda_i x_{i,s},\ 0 \right\} \\
&= \sum_{s=1}^T p_s \max\left\{y_1 - x_s,\ 0\right\} \overset{(5)}{=} F_X^2(y_1;p),
\end{aligned}
$$

where $X = \sum_{i=1}^N \lambda_i X_i$. Then, evaluating constraint (23) for an arbitrary $t \in \{2,\dots,T\}$ and $p \in \mathrm{ext}(\mathrm{conv}(P))$ gives

$$
\begin{aligned}
F_Y^2(y_t;p) &\geq \sum_{s=1}^T p_s z_{t,s} + \delta \geq \sum_{s=1}^T p_s \max\left\{ y_t - \sum_{i=1}^N \lambda_i x_{i,s},\ 0 \right\} + \delta \\
&= \sum_{s=1}^T p_s \max\left\{y_t - x_s,\ 0\right\} + \delta \overset{(5)}{=} F_X^2(y_t;p) + \delta, \qquad \text{(A.5)}
\end{aligned}
$$

where $X = \sum_{i=1}^N \lambda_i X_i$. Hence, we obtain $F_Y^2(y_t;p) \geq F_X^2(y_t;p)$ for all $t \in \{1,\dots,T\}$ and $p \in \mathrm{ext}(\mathrm{conv}(P))$. Lemma 1 implies then $F_Y^2(\theta;p) \geq F_X^2(\theta;p)$ holds for all $\theta \in \mathbb{R}$ and $p \in \mathrm{ext}(\mathrm{conv}(P))$, which by Theorem 1 of Liesiö et al. (2020) implies $X \succeq_P Y$ holds. Moreover, the objective function value $\Delta(X,Y)$ (cf. (12)) satisfies

$$
\begin{aligned}
\Delta(X,Y) &= \min_{\substack{\theta\geq y_2 \\ p\in P}} \left(F_Y^2(\theta;p) - F_X^2(\theta;p) \right) \\
&= \min_{p\in\mathrm{ext}(\mathrm{conv}(P))} \min_{\theta\geq y_2} \left(F_Y^2(\theta;p) - F_X^2(\theta;p) \right), \qquad \text{(A.6)}
\end{aligned}
$$

since $F_Y^2(\theta;p) - F_X^2(\theta;p)$ is linear in p (see, Eq. (5)) and therefore the minimum is attained at some $p \in \mathrm{ext}(\mathrm{conv}(P))$. Furthermore, for any $p \in \mathrm{ext}(\mathrm{conv}(P))$, $F_Y^2(\theta;p) - F_X^2(\theta;p)$ is a piece-wise linear function in θ with non-linearities at points $\theta \in \{x_1,\dots,x_T, y_2,\dots,y_T\}$. Thus, by using similar arguments as in the proof of Lemma 1, it can be shown that

$$
\min_{\theta\geq y_2} \left(F_Y^2(\theta;p) - F_X^2(\theta;p) \right) = \min_{\substack{t \\ t\geq 2}} \left(F_Y^2(y_t;p) - F_X^2(y_t;p) \right) \geq \delta, \qquad \text{(A.7)}
$$

where the last inequality follows from (A.5). Together, (A.6) and (A.7) imply that $\Delta(X,Y) \geq \delta$. Thus, $X = \sum_{i=1}^N \lambda_i X_i$ is a feasible solution to optimization problem (13) such that $\Delta(X,Y) \geq \delta$.

Now, let $\sum_{i=1}^N \lambda_i X_i$ be an optimal solution to (13). Then, by (i), there exists a feasible solution (λ, z, δ) to LP problem (20)–(23) such that $\delta = \Delta\left(\sum_{i=1}^N \lambda_i X_i, Y\right)$. Assume now (λ, z, δ) is not optimal to (20)–(23), then there must exist another feasible solution (λ', z', δ') giving a better objective function value, i.e., $\delta' > \delta$. Then, by (ii), $\sum_{i=1}^N \lambda'_i X_i$ is also a feasible solution to (13) with $\Delta\left(\sum_{i=1}^N \lambda'_i X_i, Y\right) \geq \delta' > \delta = \Delta\left(\sum_{i=1}^N \lambda_i X_i, Y\right)$, which contradicts the assumption that $\sum_{i=1}^N \lambda_i X_i$ is the optimal solution to (13). Hence, (λ, z, δ) is an optimal solution to (20)–(23).

In turn, let (λ, z, δ) be an optimal solution to LP problem (20)–(23). Then, by (ii), $\sum_{i=1}^N \lambda_i X_i$ is a feasible solution to optimization problem (13) and $\Delta\left(\sum_{i=1}^N \lambda_i X_i, Y\right) \geq \delta$. Suppose that $\sum_{i=1}^N \lambda_i X_i$ is not optimal to (13). Then, there exists another feasible solution $\sum_{i=1}^N \lambda'_i X_i$ with a better objective function value, i.e., $\Delta\left(\sum_{i=1}^N \lambda'_i X_i, Y\right) > \Delta\left(\sum_{i=1}^N \lambda_i X_i, Y\right)$. Then, by (i), there exist z' and δ' such that (λ', z', δ') is also feasible to (20)–(23) and $\delta' = \Delta\left(\sum_{i=1}^N \lambda'_i X_i, Y\right) > \Delta\left(\sum_{i=1}^N \lambda_i X_i, Y\right) \geq \delta$, which contradicts the assumption that (λ, z, δ) is the optimal solution to (20)–(23). Therefore, $\sum_{i=1}^N \lambda_i X_i$ is an optimal solution to (13). □

Appendix B. Additional table of Section 5

See Table B.8.

References

Arvanitis, S., Post, T., & Topaloglou, N. (2021). Stochastic bounds for reference sets in portfolio analysis. *Management Science*, *67*(12), 7737–7754.

Bruni, R., Cesarone, F., Scozzari, A., & Tardella, F. (2017). On exact and approximate stochastic dominance strategies for portfolio selection. *European Journal of Operational Research*, *259*(1), 322–329.

Canepa, A., de la O. González, M., & Skinner, F. S. (2020). Hedge fund strategies: A non-parametric analysis. *International Review of Financial Analysis*, *67*, Article 101436.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57–82.

Chavez-Bedoya, L., & Rosales, F. (2021). Reduction of estimation risk in optimal portfolio choice using redundant constraints. *International Review of Financial Analysis*, *78*, Article 101930.

Cipollini, F., Gallo, G. M., & Palandri, A. (2021). A dynamic conditional approach to forecasting portfolio weights. *International Journal of Forecasting*, *37*(3), 1111–1126.

Consigli, G., Moriggia, V., & Vitali, S. (2020). Long-term individual financial planning under stochastic dominance constraints. *Annals of Operations Research*, *292*, 973–1000.

Dentcheva, D., & Römisch, W. (2013). Stability and sensitivity of stochastic dominance constrained optimization models. *SIAM Journal on Optimization*, *23*(3), 1672–1688.

Dentcheva, D., & Ruszczyński, A. (2003). Optimization with stochastic dominance constraints. *SIAM Journal on Optimization*, *14*(2), 548–566.

Dentcheva, D., & Ruszczyński, A. (2006). Portfolio optimization with stochastic dominance constraints. *Journal of Banking & Finance*, *30*(2), 433–451.

Dentcheva, D., & Ruszczyński, A. (2010). Robust stochastic dominance and its application to risk-averse optimization. *Mathematical Programming*, *123*, 85–100.

Dupačová, J., & Kopa, M. (2012). Robustness in stochastic programs with risk constraints. *Annals of Operations Research*, *200*, 55–74.

Dupačová, J., & Kopa, M. (2014). Robustness of optimal portfolios under risk and stochastic dominance constraints. *European Journal of Operational Research*, *234*, 434–441.

Egozcue, M., & Wong, W. (2010). Gains from diversification on convex combinations: A majorization and stochastic dominance approach. *European Journal of Operational Research*, *200*(3), 893–900.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Guerard, J. B., Markowitz, H., & Xu, G. (2015). Earnings forecasting in a global stock selection model and efficient portfolio construction and management. *International Journal of Forecasting*, *31*(2), 550–560.

Hadar, J., & Russell, W. (1969). Rules for ordering uncertain prospects. *American Economic Review*, *59*(1), 25–34.

Han, Y., & Li, J. (2022). Should investors include green bonds in their portfolios? Evidence for the USA and europe. *International Review of Financial Analysis*, *80*, Article 101998.

Han, Y., & Li, J. (2023). The impact of global economic policy uncertainty on portfolio optimization: A black-litterman approach. *International Review of Financial Analysis*, *86*, Article 102476.

Hanoch, G., & Levy, H. (1969). The efficiency analysis of choices involving risk. *Review of Economic Studies*, *36*(3), 335–346.

Henriksen, T. E., Pichler, A., Westgaard, S., & Frydenberg, S. (2019). Can commodities dominate stock and bond portfolios? *Annals of Operations Research*, *282*, 155–177.

Hoang, T.-H.-V., Lean, H. H., & Wong, W.-K. (2015). Is gold good for portfolio diversification? A stochastic dominance analysis of the Paris stock exchange. *International Review of Financial Analysis*, *42*, 98–108.

Hodder, J., Jackwerth, J., & Kolokolova, O. (2015). Improved portfolio choice using second-order stochastic dominance. *Review of Finance*, *19*(4), 1623–1647.

Kallio, M., & Hardoroudi, N. D. (2018). Second-order stochastic dominance constrained portfolio optimization: Theory and computational tests. *European Journal of Operational Research*, *264*, 675–685.

Kassimatis, K. (2021). Mean-variance versus utility maximization revisited: The case of constant relative risk aversion. *International Review of Financial Analysis*, *78*, Article 101932.

Kazak, E., & Pohlmeier, W. (2019). Testing out-of-sample portfolio performance. *International Journal of Forecasting*, *35*(2), 540–554.

Keppe, H.-J., & Weber, M. (1989). Stochastic dominance with incomplete information on probabilities. *European Journal of Operational Research*, *43*(3), 350–355.

Khashanah, K., Simaan, M., & Simaan, Y. (2022). Do we need higher-order comoments to enhance mean-variance portfolios? evidence from a simplified jump process. *International Review of Financial Analysis*, *81*, Article 102068.

Kim, D., & Kang, K. H. (2021). Conditional value-at-risk forecasts of an optimal foreign currency portfolio. *International Journal of Forecasting*, *37*(2), 838–861.

Kitamura, Y. (1997). Empirical likelihood methods with weakly dependent processes. *The Annals of Statistics*, *25*(5), 2084–2102.

Kopa, M., Moriggia, V., & Vitali, S. (2018). Individual optimal pension allocation under stochastic dominance constraints. *Annals of Operations Research*, *260*, 255–291.

Kopa, M., & Post, T. (2015). A general test for SSD portfolio efficiency. *OR Spectrum*, *37*, 703–734.

Kouaissah, N. (2023). Robust reward-risk performance measures with weakly second-order stochastic dominance constraints. *The Quarterly Review of Economics and Finance*, *88*, 53–62.

Kuosmanen, T. (2004). Efficient diversification according to stochastic dominance criteria. *Management Science*, *50*(10), 1390–1406.

La Torre, D., & Mendivil, F. (2018). Portfolio optimization under partial uncertainty and incomplete information: A probability multimeasure-based approach. *Annals of Operations Research*, *267*, 267–279.

La Torre, D., & Mendivil, F. (2022). Stochastic efficiency and inefficiency in portfolio optimization with incomplete information: A set-valued probability approach. *Annals of Operations Research*, *311*, 1085–1098.

Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, *15*(5), 850–859.

Levi, I. (1980). *The enterprise of knowledge: An essay on knowledge, credal probability, and chance*. MIT Press.

Levy, H. (2016). *Stochastic dominance: Investment decision making under uncertainty* (3rd ed.). Springer.

Liesiö, J., Kallio, M., & Argyris, N. (2023). Incomplete risk-preference information in portfolio decision analysis. *European Journal of Operational Research*, *304*(3), 1084–1098.

Liesiö, J., & Salo, A. (2012). Scenario-based portfolio selection of investment projects with incomplete probability and utility information. *European Journal of Operational Research*, *217*, 162–172.

Liesiö, J., Xu, P., & Kuosmanen, T. (2020). Portfolio diversification based on stochastic dominance under incomplete probability information. *European Journal of Operational Research*, *286*(2), 755–768.

Longarela, I. R. (2016). A characterization of the SSD-efficient frontier of portfolio weights by means of a set of mixed-integer linear constraints. *Management Science*, *62*(12), 3549–3554.

Migliavacca, M., Goodell, J. W., & Paltrinieri, A. (2023). A bibliometric review of portfolio diversification literature. *International Review of Financial Analysis*, *90*, Article 102836.

Montes, I., Miranda, E., & Montes, S. (2014a). Decision making with imprecise probabilities and utilities by means of statistical preference and stochastic dominance. *European Journal of Operational Research*, *234*(1), 209–220.

Montes, I., Miranda, E., & Montes, S. (2014b). Stochastic dominance with imprecise information. *Computational Statistics & Data Analysis*, *71*, 868–886.

Moskowitz, T., & Grinblatt, M. (1999). Do industries explain momentum? *The Journal of Finance*, *54*(4), 1249–1290.

Moskowitz, H., Preckel, P. V., & Yang, A. (1993). Decision analysis with incomplete utility and probability information. *Operations Research*, *41*(5), 864–879.

Pho, K. H., Ly, S., Lu, R., Hoang, T. H. V., & Wong, W.-K. (2021). Is bitcoin a better portfolio diversifier than gold? A copula and sectoral analysis for China. *International Review of Financial Analysis*, *74*, Article 101674.

Post, T. (2003). Empirical tests for stochastic dominance efficiency. *The Journal of Finance*, *58*(5), 1905–1931.

Post, T., Karabati, S., & Arvanitis, S. (2018). Portfolio optimization based on stochastic dominance and empirical likelihood. *Journal of Econometrics*, *206*, 167–186.

Post, T., Karabati, S., & Arvanitis, S. (2019). Robust optimization of forecast combinations. *International Journal of Forecasting*, *35*(3), 910–926.

Post, T., & Kopa, M. (2017). Portfolio choice based on third-degree stochastic dominance. *Management Science*, *63*(10), 3381–3392.

Roccazzella, F., Gambetti, P., & Vrins, F. (2022). Optimal and robust combination of forecasts via constrained optimization and shrinkage. *International Journal of Forecasting*, *38*(1), 97–116.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, *2*(3), 21–41.

Roman, D., Darby-Dowman, K., & Mitra, G. (2006). Portfolio construction based on stochastic dominance and target return distributions. *Mathematical Programming*, *108*, 541–569.

Roman, D., Mitra, G., & Zverovich, V. (2013). Enhanced indexation based on second-order stochastic dominance. *European Journal of Operational Research*, *228*(1), 273–281.

Rothschild, M., & Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, *2*, 225–243.

Salo, A., Doumpos, M., Liesiö, J., & Zopounidis, C. (2023). Fifty years of portfolio optimization. *European Journal of Operational Research*.

Sehgal, R., & Mehra, A. (2020). Robust portfolio optimization with second order stochastic dominance constraints. *Computers & Industrial Engineering*, *144*, Article 106396.

Tsetlin, I., Winkler, R. L., Huang, R. J., & Tzeng, L. Y. (2015). Generalized almost stochastic dominance. *Operations Research*, *63*(2), 363–377.

Tzeng, L. Y., Huang, R. J., & Shih, P.-T. (2013). Revisiting almost second-degree stochastic dominance. *Management Science*, *59*(5), 1250–1254.

Vilkkumaa, E., Liesiö, J., Salo, A., & Ilmola-Sheppard, L. (2018). Scenario-based portfolio model for building robust and proactive strategies. *European Journal of Operational Research*, *266*(1), 205–220.

Vitali, S., & Moriggia, V. (2021). Pension fund management with investment certificates and stochastic dominance. *Annals of Operations Research*, *299*, 273–292.
Wong, W. (2007). Stochastic dominance and mean-variance measures of profit and loss for business planning and investment. *European Journal of Operational Research*, *182*, 829–843.
Xia, H., Min, X., & Deng, S. (2015). Effectiveness of earnings forecasts in efficient global portfolio construction. *International Journal of Forecasting*, *31*(2), 568–574.