Liesiö, Juuso; Xu, Peng; Kuosmanen, Timo

Portfolio Diversification based on Stochastic Dominance under Incomplete Probability Information

Published in:
European Journal of Operational Research

DOI:
10.1016/j.ejor.2020.03.042

Published: 16/10/2020

Portfolio Diversification based on Stochastic Dominance under Incomplete Probability Information

Juuso Liesiö[a,*], Peng Xu[a], Timo Kuosmanen[a]

[a]*Management Science Group, Department of Information and Service Management, Aalto University School of Business, P.O. Box 21220, FI-00076 Aalto, Finland*

Abstract

Identifying efficient portfolio diversification strategies subject to stochastic dominance (SD) criteria usually assumes that the state-space of future asset returns can be captured by a fixed sample of equally probable historical returns. This paper relaxes this assumption by developing SD criteria under incomplete information on state probabilities. Specifically, we identify portfolios that dominate a given benchmark for any state probabilities in a given set. The proposed approach is applied to analyze if industrial diversification can be utilized to outperform the market portfolio. The results from this application demonstrate that the use of set-valued state probabilities can help to improve out-of-sample performance of SD-based portfolio optimization.

Keywords: decision analysis, finance, stochastic dominance, portfolio diversification, incomplete probability information

1. Introduction

Stochastic dominance (SD) is a widely used analytical tool for comparing decision alternatives with uncertain outcomes that does not require exact specification of the decision maker's (DM's) utility function (for a comprehensive overview see, e.g., Levy 2016). In finance, SD has been used to analyze the efficiency of asset portfolio diversification strategies. This application area was long dominated by the classical mean-variance models, which – unlike the SD criteria – consider only two moments of the return distribution (Markowitz 1952). The strength of mean-variance models lies in their ability to test if the diversification of a benchmark portfolio is efficient in view of observed asset returns, and if not, identify a diversification strategy that yields an equal mean return with minimal variance. In contrast, early SD approaches were designed to establish if one empirical return distribution dominates another (Levy & Hanoch 1970, Aboidi & Thon 1994), and hence using them to confirm whether a particular portfolio is efficient would require explicit enumeration of all alternative diversification strategies. The first steps to bridge the gap between mean-variance and SD approaches were taken by Shalit & Yitzhaki (1994). Their marginal conditional stochastic dominance approach identifies pairs of assets in the benchmark portfolio such that increasing the weight of one asset while decreasing the weight of the other will result in a dominating portfolio. Ultimately, the models developed by Post (2003) and Kuosmanen (2004) made it possible to establish whether a given benchmark portfolio is efficiently diversified in the context of all marketed portfolios. If the

*Corresponding author

Email addresses: juuso.liesio@aalto.fi (Juuso Liesiö), peng.xu@aalto.fi (Peng Xu), timo.kuosmanen@aalto.fi (Timo Kuosmanen)

benchmark is inefficient, these models are also able to identify a dominating portfolio that is itself SD-efficient, i.e., is not dominated by any marketed portfolio. As a result of these seminal contributions, there is a growing literature on SD-based portfolio optimization (e.g., Wong 2007, Lozano & Gutiérrez 2008, Egozcue & Wong 2010, Post & Kopa 2013, Longarela 2016, Fang & Post 2017, Bruni et al. 2017, Post & Kopa 2017, Kallio & Deghan Hardoroudi 2018, Kallio & Hardoroudi 2019, Levy 2019, Huang et al. 2020).

This SD-literature has mainly focused on decision making under risk in which there exists a well-specified probability distribution captured by a single vector of state probabilities. In particular, many models are based on using a sample of historical asset returns to represent the state-space of equally likely states. Arguably, this empirical distribution function (EDF) approach is conceptually sound as observed vectors of asset returns represent plausible scenarios for future returns (Post et al. 2018). Moreover, if the returns are serially independent and identically distributed, this approach yields statistically consistent estimates for the portfolios' return distributions. Equal state probabilities are also sufficient in applications where stochastic dominance is not used for investment decisions concerning the future, but to analyze portfolio efficiency in historical return data (e.g., market portfolio efficiency; Kuosmanen 2004).

When stochastic dominance is used to support selecting a portfolio in view of an uncertain future, using historical return observations as equally likely states can be difficult to justify. For instance, practical applications often have to rely on small data samples, at least relative to the number of base assets, and hence favorable asymptotic properties do not necessarily imply that the obtained state-space offers a reasonable representation of future returns. Alternatively, the state space can be generated by sampling some statistical model that forecasts assets' future returns in which case the state-space is not restricted to cover only historical return realizations. Using such models necessitates making assumptions about the model inputs (e.g., economic growth, inflation) and thus it is advisable to simulate future asset returns using multiple scenarios for the input parameter values. However, obtaining point-estimate probabilities for the generated states requires specifying point-estimate probabilities for the chosen input-scenarios, which can be difficult as these scenarios represent different possible futures.

This paper seeks to extend the SD-based portfolio models beyond decision making under risk with well-specified probability distributions by introducing models that allow the use of a continuum of probability distributions to capture Knightian uncertainty. Specifically, this paper relaxes the requirement of point-estimate state probabilities and develops models that identify a portfolio of marketed assets whose dominance over a given benchmark portfolio is robust under incomplete probability information. In particular, we introduce optimization models that identify the return maximizing portfolio that dominates a given benchmark portfolio for all state probabilities in a possibly non-convex and non-connected set of feasible probability vectors. We show that such a set can be used to model, for instance, (i) ordinal statements on the state probabilities (e.g., recent observations are more probable than older ones), (ii) a range of possible sample sizes, and (iii) confidence intervals constructed around point-estimate state probabilities.

Although our main focus is to develop second-order stochastic dominance (SSD) models, we also consider first-order stochastic dominance (FSD) under incomplete probability information. While SSD-based models are more suitable to support rational investment decision as they assume risk aversion or neutrality, models based on FSD may be a more suitable criterion for descriptive research, in which decision makers are not assumed to follow the axioms of Expected Utility Theory (see, e.g., Starmer 2000). Both of these SD-criteria result in linear programming models for identifying an asset portfolio that dominates a given benchmark under set-valued state probabilities, but FSD requires the introduction of binary decision variables. The linearity of these models is beneficial in view of applications, as it allows to deploy powerful off-the-shelf optimization software to solve the resulting optimization problems.

We apply these models to empirical data on industry portfolio returns and test the out-of-sample performance of the resulting portfolios. Our results show that portfolios that dominate the market portfolio under incomplete probability information are more likely to dominate the market portfolio also when comparing the out-of-sample realized returns of the two portfolios. While these models are developed in view of financial applications, they are readily applicable for identifying robust solutions in project portfolio selection and project scheduling problems, where the decision alternatives correspond to the feasible solutions of a mixed integer linear programming model and uncertainties are captured with a finite state-space (see, e.g., Gustafsson & Salo 2005, Gutjahr 2015, Baptista et al. 2019).

The developed models also offer a computationally efficient approach for conducting sensitivity analysis for standard SSD and FSD models based on point estimate probabilities. In particular, analyzing how sensitive the results obtained from standard models are to variations in the state probabilities requires optimizing the models repeatedly for multiple state probability vectors. Yet, there are no guarantees that any of these identified optimal portfolios is robust in the sense that its dominance over the benchmark portfolio holds for a larger set of state probability vectors. In turn, the models developed here require only a single optimization run to identify a portfolio whose stochastic dominance over the benchmark is not sensitive to variations in the state probabilities.

Beyond the SD literature described above, this paper intersects with several other strands of research worth acknowledging. For instance, Dupačová & Kopa (2014) use a contamination analysis to determine if stochastic dominance between two portfolios holds when the state probabilities are allowed to take any values obtained as a linear combination of the empirical probability distribution and some fixed contamination distribution. Post & Potí (2017) measure the inefficiency of a portfolio by analyzing the required divergence from empirical probabilities that would make the portfolio optimal for some utility function exhibiting decreasing absolute risk aversion. In stochastic optimization, Dentcheva & Ruszczyński (2010) develop optimality conditions for problems containing SD constraints that are required to hold for a set of different probability distributions. Decision analysis (DA) research has a long tradition of using partial orderings of decision alternatives in a setting where only incomplete information about probabilities and risk preferences is available (Pearman & Kmietowicz 1986, Keppe & Weber 1989, Moskowitz et al. 1993, Liesiö & Salo 2012, Jiang et al. 2018, Vilkkumaa et al. 2018). However, the methodological focus of this research has been on establishing dominance between fixed pairs of alternatives (cf. portfolios), whereas the optimization models developed here aim to construct a dominating portfolio from the available base assets.

This paper is structured as follows. Section 2 introduces the notation and standard definitions related to SD-based portfolio efficiency analysis, and Section 3 then extends these concepts to account for set-valued state probabilities. Section 4 develops LP models to identify benchmark dominating portfolios under incomplete probability information, and Section 5 applies the developed models to identify industrial diversification strategies dominating the market portfolio and analyzes their out-of-sample performance. Section 6 concludes. Models for identifying portfolios with robust dominance over the benchmark in the sense of FSD are presented in the appendix.

2. Preliminaries

The returns of the m base assets are modeled as non-negative real-valued random variables $X_1, ..., X_m$ on the set of mutually exclusive and collectively exhaustive states $S = \{s_1, ..., s_n\}$. The states can correspond to a sample of returns obtained from historical observations or to a random sample drawn from a suitable joint distribution representing future asset returns (multivariate log-normal or GARCH models, for instance). Specification of the state-space will be discussed in detail in Section 3. The return of the jth asset in the ith state is denoted by $x_{ji} = X_j(s_i) \in \mathbb{R}^+$. A portfolio of these assets is characterized by a vector of asset weights $\lambda = (\lambda_1, ..., \lambda_m)^T$ capturing the share of initial capital allocated for each asset. The asset weights belong to the set

$$\Lambda = \left\{ \lambda \in \mathbb{R}^m \;\middle|\; \sum_{j=1}^{m} \lambda_j = 1 \right\}, \tag{1}$$

although additional restrictions may also apply. For instance, short positions may be prohibited by requiring that the asset weights can take only non-negative values.

The return of a portfolio with weights λ is captured by the random variable $X = \sum_{j=1}^{m} \lambda_j X_j$ whose state-specific returns are denoted by $x_i = X(s_i) = \sum_{j=1}^{m} \lambda_j x_{ji}$. The set of all portfolio return distributions thus corresponds to the set of all random variables obtained as mixtures of the base assets' returns, i.e.,

$$\mathcal{X} = \{\sum_{j=1}^{m} \lambda_j X_j \mid \lambda \in \Lambda\}. \tag{2}$$

We use Y and $y_i = Y(s_i)$ to denote the fixed benchmark portfolio and its state-specific returns, respectively. This benchmark can be some mixture of the base assets (i.e., $Y \in \mathcal{X}$) or it can represent some desired target return distribution that cannot be replicated using the marketed assets (i.e., $Y \notin \mathcal{X}$) .

The state probabilities form a vector $p = (p_1, ...p_n)$ in the n-dimensional simplex

$$P^0 = \left\{ p \in [0,1]^n \;\middle|\; \sum_{i=1}^{n} p_i = 1 \right\}. \tag{3}$$

The return distribution of any portfolio X is clearly dependent on values of the state probabilities p. To make this link explicit we denote the expectation and cumulative distribution function (CDF) of portfolio X by $\mathbb{E}_p[X] = \sum_{i=1}^{n} p_i x_i$ and

$$F_X(t; p) = \mathbb{P}(\{s_i \in S | X(s_i) \le t\}) = \sum_{\substack{i \in \{1,...,n\} \\ x_i \le t}} p_i, \tag{4}$$

respectively. Furthermore, we denote the integral of the CDF by

$$F_X^2(t; p) = \int_{-\infty}^{t} F_X(r; p) dr = \sum_{\substack{i \in \{1,...,n\} \\ x_i \le t}} p_i (t - x_i). \tag{5}$$

Second-order stochastic dominance (SSD) compares these integrals to provide a partial ranking of probability distribution.

Definition 1. Portfolio $X \in \mathcal{X}$ weakly dominates portfolio $Y \in \mathcal{X}$ in the sense of second-order stochastic dominance, denoted by $X \succeq^2 Y$, if

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

where $F_{(\cdot)}^2$ is given by (5).

If a portfolio dominates another in the sense of SSD, then any expected utility maximizing risk-averse or -neutral decision maker would prefer the former portfolio over the latter (Hanoch & Levy 1969). This result can be formally stated as

$$X \succeq^2 Y \Leftrightarrow \mathbb{E}_p[u(X)] \geq \mathbb{E}_p[u(Y)] \text{ for all } u \in U, \tag{7}$$

where U is the set of all non-decreasing concave utility functions.

SSD can be used to support portfolio selection by specifying some benchmark portfolio Y and identifying a portfolio $X \in \mathcal{X} = \{\sum_{j=1}^{m} \lambda_j X_j \mid \lambda \in \Lambda\}$ that dominates Y in the sense of SSD. There can be multiple dominating portfolios and selecting the one that maximizes the expected return leads to the optimization problem

$$\max_{X \in \mathcal{X}} \{\mathbb{E}_p[X] \mid X \succeq^2 Y\}. \tag{8}$$

Clearly it would be technically possible to use some other objective function in optimization problem (8) instead of expected portfolio return. However, our choice to use the expected portfolio return stems from the reasoning that any portfolio dominating the chosen benchmark is acceptable to the DM in terms of risk and thus the choice from a set of such dominating portfolios can be based on risk-neutral preferences. For instance, the optimal portfolio X^* obtained from (8) can be relevant to fund managers whose clients usually have different risk preferences: Any risk-averse or risk-neutral client maximizing the expected utility would prefer portfolio X^* to the benchmark portfolio Y.

3. Stochastic Dominance under Incomplete Probability Information

The state-space for problem (8) is usually constructed by taking a sample of n most recent observations of assets' and benchmark portfolio's state-specific returns (x_{ji}, y_i, $i \in \{1, ..n\}$) and assigning an equal probability ($p_i = 1/n$) to each state. This approach is appealing from a practical stand-point, and it can also be supported by statistical arguments. In particular, each vector of observed asset returns arguably represents a plausible scenario of future returns. Moreover, if one assumes the joint distribution of the asset returns does not change over time, then these scenarios should be assigned equal probabilities. Essentially, this approach corresponds to using the Empirical Distribution Functions (EDFs) as estimates of the true CDFs of the assets' returns, and under the above assumption the EDF converges to the 'true' CDF in probability as the sample size n increases.

However, such asymptotic properties do not guarantee that an EDF is close to the underlying CDF with small data sets often deployed in applications. Moreover, asset returns can exhibit dynamic patterns such as price reversals and volatility jumps, which violate the assumption that the joint distribution of asset returns remains unchanged over time. In response to these challenges, Post et al. (2018) suggest the use of empirical likelihood (EL) approach to estimate the state probabilities. In this approach the state probabilities are required to satisfy some pre-specified constraints (moment conditions) capturing conditioning information on, for instance, empirical stylized facts about common risk factors in the financial market (Fama & French 1993). Generally, there exist several state probability vectors that satisfy these constraints, and the EL approach selects the vector that is closest to equal state probabilities with respect to the Kullback-Leibler distance.

The models developed in this paper take a different approach, which can be seen as a generalization of the approach suggested by Post et al. (2018): Rather than seeking to specify 'correct' values for the state probabilities in problem (8), we take the alternative approach of accepting the uncertainty in these values and identify a portfolio X which stochastically dominates the benchmark portfolio Y even if these values vary. Technically, this is accomplished by relaxing the assumption of a single probability vector $\tilde{p} \in P^0$ and considering instead a set of feasible probability vectors $P \subseteq P^0$, where the set of all possible probability vectors P^0 is given by equation (3).

This general definition of probability set P enables to capture several types of parameter uncertainties. For instance, it can be used to build confidence regions around point-estimate probabilities $\tilde{p}$ obtained either through the standard EDF approach (i.e., $\tilde{p}_i = \frac{1}{n}$) or the more advanced EL-based approach by Post et al. (2018). In particular, the set of feasible probabilities

$$P = \left\{ p \in P^0 \,\middle|\, (1-\alpha)\tilde{p}_i \leq p_i \leq (1+\alpha)\tilde{p}_i, i \in \{1, \ldots, n\} \right\} \tag{9}$$

allows a $\alpha\%$ variation in each state probability $\tilde{p}_i$. Alternatively, only the lower bounds can be enforced to obtain the large set of feasible probabilities

$$P = \left\{ p \in P^0 \,\middle|\, (1-\alpha)\tilde{p}_i \leq p_i, i \in \{1, \ldots, n\} \right\}. \tag{10}$$

The resulting upper bound for the probability of the ith state is $p_i \leq (1-\alpha)\tilde{p}_i + \alpha$. Figure 1 illustrates the relationship between sets (9) and (10) in a three-state setting.

Alternatively, the EL approach can be utilized to characterize the entire set of feasible probabilities – not only the center point of the confidence region as in (9) and (10). For instance, consider an EL model that uses a single factor f defined as the linear combination $f = \sum_{j=1}^{n} w_j X_j$ of the base asset returns, and suppose the model is specified to find a probability vector such that the implied expected factor value falls between the 10th and 90th percentiles of the factor's historical distribution. Then, the EL estimate $\tilde{p}$ is obtained as the optimal solution to the optimization problem

$$\begin{aligned} \min_{\tilde{p} \in P} \quad & D(\tilde{p}), \\ & P = \left\{ p \in P^0 \,\middle|\, L \leq \sum_{i=1}^{n} p_i f(s_i) \leq U \right\}, \end{aligned} \tag{11}$$

where $D(\cdot)$ denotes the Kullback-Leibler divergence form equal state probabilities, and L and U denote the 10th and 90th percentiles of the historical distribution, respectively. The model developed here does not require the use of any single probability vector but can directly utilize probability set (11). Similarly, it is possible to utilize a set of feasible probabilities obtained by bounding the expected values of multiple factors.

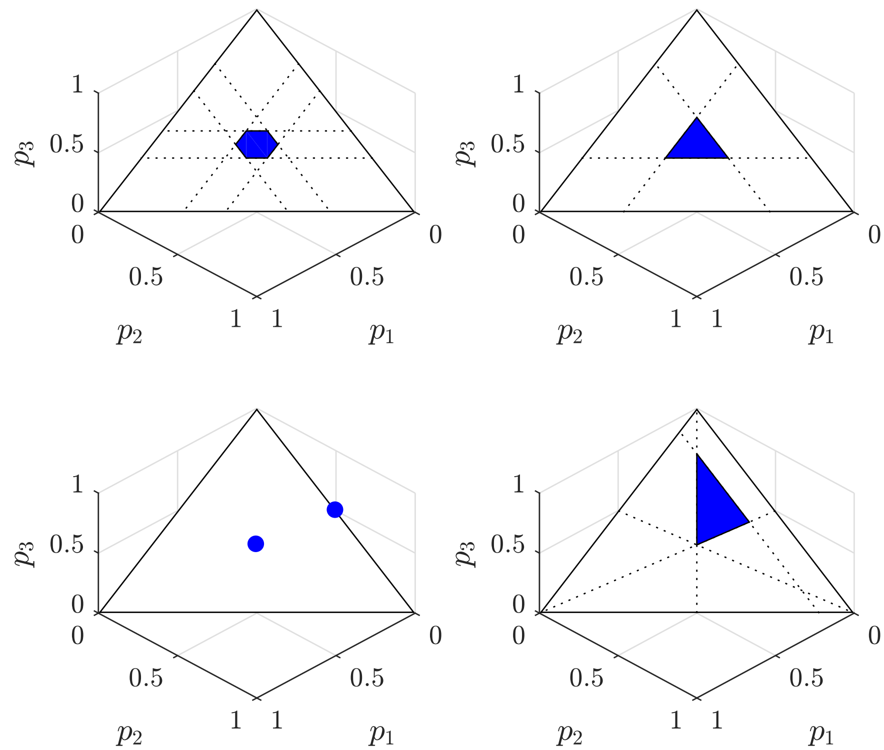


Figure 1: Top figures illustrate the sets of feasible probabilities (9) (left) and (10) (right), when $\tilde{p} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and $\alpha = 0.2$. Left and right bottom figures illustrate the sets given by (12) ($\underline{n} = 2$) and (13) ($\alpha = 1/3$), respectively.

The set of feasible probabilities P can also be used to capture incomplete information on the sample size, which is effectively determined by the number of states with strictly positive probabilities. Hence, allowing this number to vary across the vectors included in set P enables to take into account multiple sample sizes simultaneously. For instance, the standard EDF approach can be extended to consider multiple sample sizes by deploying a finite set of probability vectors. Suppose the n states represent the observed asset returns in a temporal order with $i = n$ corresponding to the most recent observations. Then, allowing sample size to vary between $\underline{n}$ and n results in the set of feasible probabilities

$$P = \left\{ p \in P^0 \;\middle|\; p = (0, \dots, 0, \underbrace{\frac{1}{k}, \dots, \frac{1}{k}}_{k \text{ elements}}), k \in \{\underline{n}, ..., n\} \right\}. \tag{12}$$

For long windows of annual data it can be reasonable to assume that recent observations of asset returns are more informative than remote observations. This assumption can be modeled by deploying the set of probability vectors that assigns a higher probability to more recent observations

$$P_\alpha = \left\{ p \in P^0 \;\middle|\; \frac{\alpha}{n} \leq p_{i-1} \leq p_i, i \in \{2, \dots, n\} \right\}, \tag{13}$$

where parameter $\alpha \in [0, 1]$ can be used to control the minimum probability across observations.

Although our theoretical results hold for any non-empty closed set $P \subseteq P^0$, practical implementation of the optimization problems developed later in Section 4 requires that the convex hull of P has a finite number of extreme points. It is worth highlighting that this requirement is not as restrictive as convexity or connectedness of P: For instance, the set defined by (12) is not convex or continuous, but its convex hull has $n-(\underline{n}-1)$ extreme points (cf. Figure 1).

The SD criteria can be extended to settings with set-valued state probabilities by requiring that dominance holds for all probability vectors in P. Importantly, the resulting robust SD relation inherits the transitivity of standard SD.

Definition 2. Portfolio $X \in \mathcal{X}$ dominates portfolio $Y \in \mathcal{X}$ in the sense of SSD with regard to the set of feasible probabilities $P \subseteq P^0$, denoted by $X \succeq_P^2 Y$, if

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

Strict dominance $X \succ_P^2 Y$ holds if $X \succeq_P^2 Y$ and $\neg(Y \succeq_P^2 X)$, and the notation $X \sim_P^2 Y$ is used when both $X \succeq_P^2 Y$ and $Y \succeq_P^2 X$ hold. Moreover, Definition 1 is obtained as the special cases of Definition 2 when the set of feasible probabilities consists of a single probability vector, i.e., $P = \{p\}$.

This robust SSD can be used to identify portfolios that are not sensitive to parameter uncertainty: If $X \succ_P^2 Y$, then X will stochastically dominate the benchmark portfolio Y for all probability vectors in set P in the sense of SSD. An alternative interpretation for Definition 2 can be established by utilizing expected utility interpretation of SSD (see equation (7)): $X \succeq_P^2 Y$ if and only if any risk-averse or -neutral decision maker prefers portfolio X over Y for the state probabilities contained in set P. Formally, this result can be written as

$$X \succeq_P^2 Y \Leftrightarrow \mathbb{E}_p[u(X)] \geq \mathbb{E}_p[u(Y)] \text{ for all } u \in U, p \in P. \tag{14}$$

In general, the dominance condition in Definition 2 is more restrictive than the standard SSD (Definition 1). In particular, even if one portfolio dominates another in the sense of standard SSD, it is possible that this dominance does not hold for all probability vectors in set P. Hence, the restrictiveness of the condition can be controlled through the size of set P. At one extreme the set composes of a single probability vector $P = \{p\}$ and thus the dominance condition coincides with standard SSD. At the other extreme P contains all vectors whose elements are non-negative probabilities that sum up to one. The sets of feasible probabilities (9)-(13) are between these two extremes.

Reducing the set of feasible probabilities will not affect any existing dominance relations: That is, if $X \succeq_P^2 Y$, then $X \succeq_{P'}^2 Y$ for any $P' \subseteq P$. This property has an intuitive interpretation that more precise information on the state probabilities will generally lead to more conclusive ranking of the portfolios. Another property of the relation $\succeq_P^2$ is that extending the set of feasible probabilities P by adding new probability vectors to it does not affect the dominance between two portfolios, if these new probability vectors fall inside the convex hull of P. In other words, two different sets whose convex hulls are equal will imply the same dominance relations between any portfolios. Hence, a sufficient condition for dominance to hold for all probability vectors in P is that it holds everywhere on the 'border' of P's convex hull. This property is formally stated by the following theorem, which uses $\text{conv}(\cdot)$ to denote the convex hull and $\text{ext}(\cdot)$ to denote the extreme points of a set.

Theorem 1. *Let $P \subseteq P^0$. Then*

$$X \succeq_P^2 Y \quad \Leftrightarrow \quad F_X^2(t;p) \leq F_Y^2(t;p) \;\forall\; t \in \mathbb{R}, p \in \text{ext}(\text{conv}(P)).$$

A detailed proof of Theorem 1 which is based on directly using Definition 2 is presented in the appendix. However, an intuitive justification for Theorem 1 can also be derived using the equivalent expected utility definition of SSD under incomplete probability information (14). In particular, because the expected utility difference between two portfolios is a linear with regard to the state probabilities, this difference attains its minimum values at the extreme points of the convex hull of the set of feasible probabilities. At these points the minimum expected utility difference across the set of utility functions U is non-negative if and only if dominance holds (see equation (7)).

One consequence of Theorem 1 is that with the largest possible set of feasible probabilities $P = P^0$, a portfolio dominates another if it has a greater or equal return in each state. This is because the n extreme points of P^0 are unit vectors in which all but one of the states have zero probability. Hence, comparing CDFs' integrals in each of these extreme points reduces to the comparisons of which of the two portfolios has a higher state-specific outcome. This result is formalized by the following corollary.

Corollary 1. *Let $P = P^0$. Then*

$$X \succeq_P^2 Y \Leftrightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \geq \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

A practical implication of Corollary 1 is that in applications some probability information is likely needed in order to identify benchmark dominating portfolios. In particular, it seems unlikely that in a practical application there would exist portfolios that yield a greater or equal return than the benchmark portfolio in all states.

4. Identifying Dominating Portfolios under Incomplete Probability Information

Using Definition 2 to identify a portfolio $X \in \mathcal{X}$ that stochastically dominates a given benchmark portfolio Y with regard to the set of feasible probabilities P is not straightforward. Although the set of feasible asset weights Λ is defined through a system of linear constraints, and thus a search over this set can be implemented as an LP problem, Definition 2 in essence yields an infinite continuum of non-linear constraints between the CDFs' of X and Y. However, Theorem 1 offers a solution strategy in which these constraints are enforced only at the extreme points of the convex hull of P. The challenge thus becomes to establish if dominance holds for all of the extreme points simultaneously. In what follows, we address this challenge for second-order stochastic dominance, while that for first-order stochastic dominance is addressed in the appendix.

Throughout this section we assume – without loss of generality – that the states are indexed in an increasing order of the benchmark portfolio returns, i.e.,

$$y_1 \leq y_2 \leq \cdots \leq y_n. \tag{15}$$

For a fixed probability vector $p \in \text{ext}(\text{conv}(P))$ the integrals of the portfolios' CDFs are convex and continuous non-decreasing piecewise-linear functions. Hence, for X to dominate Y it is necessary that $F_X^2(\cdot; p)$ remains below (or equal to) $F_Y^2(\cdot; p)$ at each point $y_1, y_2, \ldots, y_n$. This condition is also sufficient: If it were the case that F_X^2 is above F_Y^2 somewhere inside the interval (y_{i-1}, y_i), then F_X^2 is greater than F_Y^2 also at either y_{i-1} or y_i. This is because F_Y^2 is linear on this interval and F_X^2 is convex. Specifically, for fixed i and p this condition can be formalized as

$$\begin{aligned} F_Y^2(y_i; p) \quad &\geq \quad F_X^2(y_i; p) = \sum_{\substack{k \in \{1,\ldots,n\} \\ x_k \leq y_i}} (y_i - x_k) p_k \\ &= \quad \sum_{k=1}^{n} \max\{y_i - x_k, 0\} p_k = \min_{d_{i1},\ldots,d_{in}} \left\{ \sum_{k=1}^{n} d_{ik} p_k \middle| \begin{array}{c} d_{ik} \geq y_i - x_k \\ d_{ik} \geq 0 \end{array} \right\}. \end{aligned}$$

Hence, establishing SSD with set-valued probabilities is equivalent to determining if there exists a solution to a system of linear inequalities as stated by the following theorem.

Theorem 2. *$X \succeq_P^2 Y$ if and only if there exists $d \in \mathbb{R}_+^{n \times n}$ that satisfies constraints*

$$d_{ik} \quad \geq \quad y_i - x_k \;\forall\; i, k \in \{1, ..., n\}, \tag{16}$$

$$\sum_{k=1}^{n} d_{ik} p_k \quad \leq \quad F_Y^2(y_i; p) \;\forall\; i \in \{1, ..., n\}, p \in \text{ext}(\text{conv}(P)). \tag{17}$$

A similar result for a single probability vector has been presented by Dentcheva & Ruszczyński (2003) (see also Rockafellar & Uryasev 2000). Theorem 2 can be used to formulate an LP problem, which for a given benchmark portfolio Y identifies a weakly dominating portfolio $X \in \mathcal{X}$. Substituting $x_k = \sum_{j=1}^{m} \lambda_j x_{jk}$, where $\lambda \in \Lambda$, into constraint (16) and using the expected return $\mathbb{E}_{\hat{p}}[X]$ as the objective function gives

$$\max_{\substack{\lambda \in \Lambda \\ d \in \mathbb{R}_+^{n \times n}}} \quad \sum_{i=1}^{n} \hat{p}_i \sum_{j=1}^{m} \lambda_j x_{ji} \tag{18}$$

$$\sum_{j=1}^{m} \lambda_j x_{jk} + d_{ik} \geq y_i \;\forall\; i, k \in \{1, ..., n\} \tag{19}$$

$$\sum_{k=1}^{n} d_{ik} p_k \leq F_Y^2(y_i; p) \;\forall\; i \in \{1, ..., n\}, \forall p \in \text{ext}(\text{conv}(P)). \tag{20}$$

Corollary 2. *Let $\hat{p} \in \text{conv}(P)$. If there exists $X \in \mathcal{X}$ such that $X \succeq_P^2 Y$, then LP problem (18)–(20) has an optimal solution (λ^*, d^*) such that $\mathbb{E}_{\hat{p}}[\sum_{j=1}^{m} \lambda_j^* X_j] \geq \mathbb{E}_{\hat{p}}[Y]$.*

If the expected return of the optimal portfolio $X^* = \sum_j \lambda_j^* X_j$ is strictly greater than that of the benchmark portfolio Y, then X^* strictly dominates Y w.r.t. to feasible probability set P (i.e., $X^* \succ_P^2 Y$). However, if the expected returns of the two portfolios are equal, this does not suffice to conclude that a dominating portfolios is not marketed. It can be the case that Y provides the maximal expected return under probability $\hat{p}$, but it remains possible to get the same expected return with less risk using another portfolio whose return distribution is a mean-preserving anti-spread of Y's distribution (see, e.g., Kuosmanen 2004). In practice, such a situation is not common as numerical impression alone makes it unlikely that $\mathbb{E}_{\hat{p}}[X^*] = \mathbb{E}_{\hat{p}}[Y]$.

5. Application to Industrial Diversification

In this section we apply the developed models to empirical data sets. Specifically, these applications seek to answer the following questions: (i) Do actual markets include the portfolios that stochastically dominate the market portfolio for multiple state probability vectors, and (ii) if such portfolios exist, does the dominance over the market portfolio hold also out-of-sample? We also examine how the specification of the state probabilities and sample size affects these results by running the tests using several different sets of feasible probabilities.

5.1. Data

We use daily returns of the Fama/French 49 value-weighted industry portfolios to represent the base assets $X_1, \ldots, X_{49}$ and the CRSP all-share index as the benchmark market portfolio Y [1]. The data sample includes all daily return observations from the first registered trading day in January 1927 (Jan 3rd 1927) to the last one in December 2015 (Dec 31st 2015), spanning a time horizon of 89 years (i.e., 356 quarters and 23487 trading days). The industry-specific daily mean returns vary from 0.029% to 0.075% with an industry-specific standard deviation ranging from 0.92% to 3.30%. The benchmark has a daily mean return of 0.041% along with a standard deviation of 1.07%. A detailed presentation of descriptive statistics of the data set can be found in Table 1.

Table 1: Descriptive Statistics of the 49 Industry Portfolios and the Benchmark Market Portfolio

SIC	Mean	Std.	Skew.	Kurt.	Min	Max	SIC	Mean	Std.	Skew.	Kurt.	Min	Max
Agric	0.043	1.51	0.61	18.57	-15.27	23.69	Guns	0.055	1.39	-0.05	10.92	-19.49	14.92
Food	0.044	0.92	-0.06	24.73	-16.04	15.54	Gold	0.042	2.26	0.45	10.26	-23.38	25.56
Soda	0.056	1.40	-0.28	14.05	-19.22	11.68	Mines	0.043	1.53	0.21	17.24	-17.91	19.85
Beer	0.054	1.47	0.01	22.79	-24.06	19.91	Coal	0.042	2.09	0.32	15.88	-19.34	27.31
Smoke	0.052	1.20	0.16	16.18	-13.99	16.22	Oil	0.046	1.28	0.08	17.07	-19.50	19.27
Toys	0.046	2.15	0.59	29.56	-26.75	39.74	Util	0.038	1.10	0.29	26.44	-15.26	17.92
Fun	0.052	1.81	0.18	15.61	-24.11	20.81	Telcm	0.039	1.03	0.20	21.01	-16.69	15.98
Books	0.041	1.56	0.85	28.22	-19.34	33.40	PerSv	0.044	2.03	0.32	28.17	-30.99	30.61
Hshld	0.041	1.17	-0.15	34.71	-21.46	25.87	BusSv	0.050	1.98	5.29	242.90	-37.41	61.56
Clths	0.040	1.14	-0.15	26.37	-18.51	20.49	Hardw	0.054	1.54	-0.02	18.81	-23.52	21.65
Hlth	0.044	1.54	-0.19	12.89	-15.45	17.39	Softw	0.043	2.40	0.64	13.63	-20.76	24.19
MedEq	0.053	1.60	13.05	1067.48	-53.62	111.82	Chips	0.051	1.76	0.16	26.22	-30.57	37.90
Drugs	0.049	1.15	-0.24	20.24	-18.70	16.70	LabEq	0.049	1.44	-0.05	12.13	-18.78	15.93
Chems	0.046	1.28	-0.14	18.89	-18.91	16.86	Paper	0.075	3.30	8.55	303.65	-45.65	150.00
Rubbr	0.054	1.68	0.58	28.00	-19.79	26.32	Boxes	0.048	1.26	-0.18	14.25	-21.43	12.59
Txtls	0.041	1.31	0.12	19.00	-18.40	19.50	Trans	0.039	1.36	0.12	15.25	-17.56	18.49
BldMt	0.042	1.26	0.06	21.20	-17.96	22.97	Whlsl	0.040	1.63	3.12	187.53	-44.44	66.92
Cnstr	0.048	2.02	0.69	18.48	-23.81	29.35	Rtail	0.045	1.14	0.00	16.93	-18.00	17.81
Steel	0.038	1.67	0.59	29.54	-23.94	30.39	Meals	0.047	1.35	-0.03	12.88	-15.48	19.40
FabPr	0.029	1.48	-0.13	8.88	-15.45	11.44	Banks	0.052	1.48	0.31	25.88	-20.43	23.05
Mach	0.044	1.38	0.31	22.60	-18.06	26.16	Insur	0.044	1.38	0.35	21.83	-17.15	18.93
ElcEq	0.052	1.57	0.20	16.65	-19.70	24.44	RlEst	0.038	2.13	1.13	25.17	-21.23	36.78
Autos	0.047	1.58	0.36	17.92	-19.72	27.88	Fin	0.047	1.58	0.05	28.19	-28.65	23.28
Aero	0.062	1.79	0.49	22.16	-19.29	32.00	Other	0.033	1.49	-0.05	15.39	-20.26	16.84
Ships	0.042	1.52	0.09	10.72	-13.20	16.62	Bench	0.041	1.07	-0.12	19.64	-17.41	15.76

[1] All data were accessed and downloaded in June 2016 from the data library of Kenneth R. French at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. The industry portfolios have been formed by grouping each NYSE, AMEX, and NASDAQ stock according to its 4-digit Standard Industrial Classification (SIC) code. The Center for Research in Security Prices (CRSP) index is a proxy of the value-weighted return of all common stocks incorporated in the US and listed on the NYSE, AMEX, and NASDAQ exchanges.

5.2. Investment Strategy and Empirical Test Setup

We employ a standard buy-hold trading strategy with no short positions, a 12-month formation period and rebalancing after a 1-month holding period. This choice is motivated by prior research insights into industry momentum: Moskowitz & Grinblatt (1999) document the profitability of industry level momentum [2] is strongest in intermediate 6- to 12-month historical horizons for portfolio formation with relatively short holding periods. Nevertheless, our preliminary tests showed that using a 12-month historical formation period outperforms its 6-month counterpart in all portfolio performance measures. Moreover, a 12-month formation period has been commonly used to test SD models (see, e.g., Hodder et al. 2015; Post & Kopa 2017).

With this investment strategy our data sample yields 1056 overlapping formation periods (Jan 1927 – Dec 1927, Feb 1927 – Jan 1928, ..., Dec 2014 – Nov 2015). The daily returns within each 12-month formation period are used as the state-space in problem (18)–(20), which implies that the number states varies between $n \in [226, 302]$. The optimal portfolio $\lambda^*(t)$ for each period $t \in \{1, \ldots, 1056\}$ is solved using different sets of feasible probabilities P. The expectation in the objective function (18) is evaluated under equal (i.e., empirical) state probabilities (i.e., $\hat{p} = (\frac{1}{n}, \ldots, \frac{1}{n})$). The out-of-sample performance of each optimal portfolio $\lambda^*(t)$ is evaluated using daily returns from subsequent 1-month holding period, which results in a total of 1056 out-of-sample observations (Jan 1928, Feb 1928, ..., Dec 2015).

5.3. Confidence Region around Equal State Probabilities

The first test uses a set of feasible probabilities that allows variations around the vector of equal state probabilities $p = (\frac{1}{n}, \ldots, \frac{1}{n})$. Specifically, we use the set of feasible probabilities

$$P_\alpha^T = \left\{ p \in P^0 \middle| p_i \geq \frac{\alpha}{n}, \forall i \in \{1, \ldots, n\} \right\}, \tag{21}$$

where $\alpha \in [0.9, 1]$ determines the lower bound for the probability in each state. Thus, $P_1^T = \{(\frac{1}{n}, \ldots, \frac{1}{n})\}$ is the smallest set corresponding to equal state probabilities. Decreasing the lower bound increases the size of set (21) and preliminary tests on the data indicated that problem (18)–(20) becomes infeasible when α exceeds the value of 0.9.

Figure 2 provides an overview on the performance of the optimized portfolios compared to the benchmark. Table 2 presents details on the out-of-sample performances of the optimal portfolios for all probability sets P_α^T, for $\alpha \in \{1.00, 0.98, \ldots, 0.90\}$. The top panel describes excess portfolio returns over the benchmark. The annualized mean excess return (Mean) and standard deviation (Std. dev.) exhibit a consistent pattern of diminishing returns and risk with respect to the size increase of the probability set. Notably, increasing the size of the probability set leads to a substantial reduction of active risk (i.e., the standard deviation of excess returns) by 8.32%. Moreover, this also improves the skewness of mean excess returns, which is alternatively characterized by

[2] The findings of Moskowitz & Grinblatt (1999) suggest that (i) the industry momentum effect is considerably stronger at intermediate formation horizons (up to 24-month) with short holding periods, and (ii) industry momentum returns are primarily generated by longing winner industries instead of shorting loser industries.

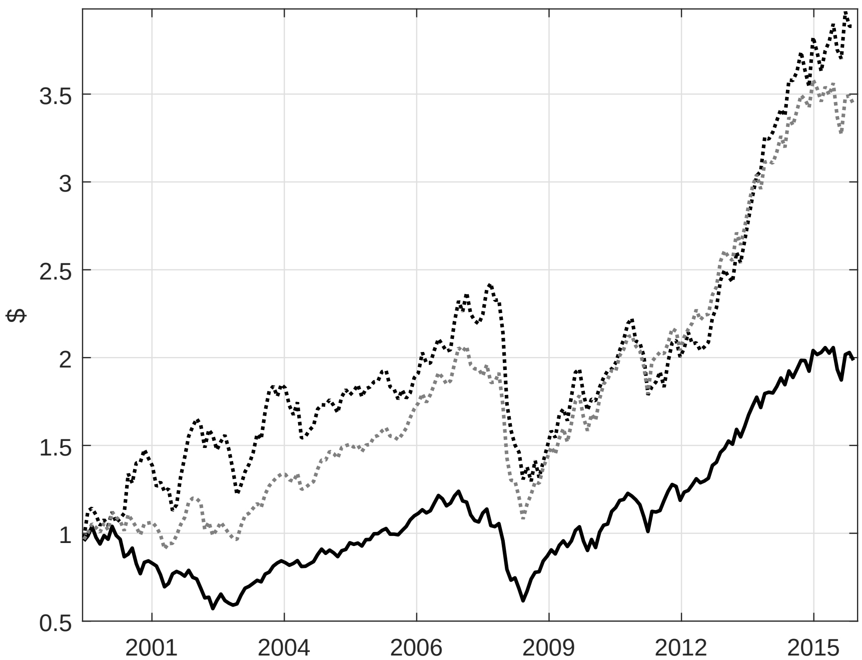


Figure 2: The solid line shows the development of benchmark portfolio value between January 2000 and December 2015. The dashed lines show the values of the portfolios optimized using probability sets $P_{1.0}^T$ (black) and $P_{0.9}^T$ (grey).

the remarkably lower downside risk for large negative returns with a significantly improved expected shortfall [3] ($\text{CVaR}_{5\%}$) of 13.13%. The certainty equivalent (CE) and risk premium (RP) [4] are evaluated using a logarithmic utility function $\ln(\cdot)$. Not surprisingly, both the CE and RP decrease rapidly. In addition, the risk-adjusted (return-to-variability) ratios (Sortino and Information) [5] imply that each unit of portfolio risk earns higher excess return as a result of increasing the size of the probability set. Portfolio turnover decreases and the average number of industries included in the optimal portfolio (# of industries) increases when the size of the probability set increases. Clearly, using a larger set P_α^T leads to a broader diversification across the base asset span. Moreover, a larger set also results in lower transaction costs as observed from the declining portfolio turnovers, which in essence indicate more cost-efficient portfolio rebalancing.

Since the objective of our model is to identify portfolios that dominate the benchmark in the sense of SSD for multiple state probability vectors in-sample, it is interesting to analyze if SSD between the two portfolios also holds when examining the realized returns during the holding period. For this purpose the bottom panel in Table

[3] Expected shortfall, also known as Conditional Value at Risk (CVaR), is the expected return conditional that the realized return belongs to the worst α%-tail of the distribution.

[4] Certainty equivalent (CE) is a certain return that yields the same expected utility as the optimal portfolio does. Risk premium (RP) is the difference between the expected portfolio return and the certainty equivalent.

[5] Sortino ratio (SR) is the ratio between the optimal portfolio's excess return over the risk-free rate and the semi-deviation of the return. Information ratio (IR) is the ratio the optimal portfolio's excess return over the benchmark portfolio and the standard deviation of the return.

Table 2: Out-of-Sample Portfolio Performance with 1-month Holding Period and Set of Feasible Probabilities P_α^T. The top panel shows the mean, standard deviation, skewness, CVaR$_5$%, certainty equivalent, risk premium and Sortino and Information ratios, of annualized excess returns to the benchmark portfolio. The middle panel reports two portfolio diversification measures, portfolio turnover and average number of industries included in the optimal portfolio. The bottom panel presents two dominance measures, almost-SSD and full SSDs over the benchmark, as well as the percentiles of distribution of p-values obtained from deploying the dominance test of Davidson (2009) in each out-of-sample period.

Measures	$P_{1.00}^T$	$P_{0.98}^T$	$P_{0.96}^T$	$P_{0.94}^T$	$P_{0.92}^T$	$P_{0.90}^T$
	Excess returns over benchmark					
Mean (%)	6.11	5.82	5.40	4.82	4.61	4.13
Std. dev. (%)	15.38	12.82	11.66	10.48	8.82	7.06
Skewness	0.89	0.89	0.98	1.13	1.10	1.10
CVaR$_{5\%}$(%)	-18.67	-15.62	-13.52	-12.21	-7.84	-5.54
CE (%)	5.07	5.09	4.79	4.33	4.26	3.90
RP (%)	1.04	0.73	0.61	0.49	0.35	0.23
Sortino ratio	1.00	1.22	1.26	1.29	1.76	2.22
Information ratio	0.40	0.45	0.46	0.46	0.52	0.58
	Portfolio diversification					
Turnover	0.76	0.75	0.73	0.71	0.69	0.64
# of industries	5.14	6.27	7.47	9.17	11.53	14.32
	Dominance over benchmark					
ε-ASSD (ε)	0.288	0.214	0.171	0.149	0.118	0.107
# of SSDs (%)	14.20	21.50	25.38	27.46	29.55	31.75
	p-values of dominance test					
1^{st} percentile	0.000	0.000	0.000	0.000	0.000	0.000
5^{th} percentile	0.002	0.000	0.000	0.000	0.000	0.000
10^{th} percentile	0.018	0.006	0.004	0.000	0.000	0.000
25^{th} percentile	0.070	0.052	0.044	0.030	0.019	0.014
50^{th} percentile	0.180	0.168	0.145	0.118	0.107	0.070
75^{th} percentile	0.266	0.268	0.251	0.218	0.208	0.183

2 reports the share of out-of-sample holding periods in which the optimal portfolio dominates the benchmark in the sense of SSD (# of SSDs). Because SSD is a binary measure (either dominance holds or not), it does not indicate how close the empirical CDFs are to satisfying the SSD condition (see Definition 1). Hence, the bottom panel in Table 2 also reports the minimum ε-value for which almost second-order stochastic dominance (ε-ASSD; Tzeng et al. 2013; see also Tsetlin et al. 2015) holds.

The results show that increasing the size of P_α^T leads to higher likelihood of obtaining full dominance by SSD over the benchmark out-of-sample. Specifically, with equal state probabilities $P = P_{1.00}^T$ out-of-sample dominance over the benchmark is obtained only in 14.20% of all holding periods. In contrast, using the largest set $P = P^T_{0.90}$ yields out-of-sample dominance over the benchmark in 31.75% of all holding periods. This finding is also supported by the ε-ASSD metric, since the average ε-values decrease as the size of the probability set increases[6].

To evaluate the statistical significance of the observed dominance relations, we apply the dominance test by Davidson (2009). This test can deal with correlated sample distributions under the null hypothesis of non-dominance, and its bootstrap algorithm requires only modest computational efforts for moderate sample sizes ($n \approx 250$). Specifically, in our setting the null hypothesis is that the optimal portfolio does *not* dominate the benchmark portfolio [7], i.e., $H_0 : \sum_{j=1}^{49} \lambda_j^* X_j \not\succ^2 Y$. Intuitively, rejecting the null implies that the alternative hypothesis $H_1 : \sum_{j=1}^{49} \lambda_j^* X_j \succ^2 Y$ holds. We test each out-of-sample dominance, and the bottom panel of Table 2 reports percentiles of the obtained distribution of p-values from the dominance test by Davidson (2009). Indeed, the p-values exhibit a consistent pattern of decline as the size of the probability set is increased.

5.4. Probability Ranking

The second test applies a set of feasible probabilities that assigns higher probabilities to the states corresponding to more recent return observations. By indexing the states in a temporal order, this type of set can be formally defined as

$$P^R_\alpha = \left\{ p \in P^0 \middle| p_i \geq p_{i-1} \geq \frac{\alpha}{n}, \forall i \in \{2, \ldots, n\} \right\}, \tag{22}$$

where parameter $\alpha \in [0, 1]$ determines the lower bound for the probability in each state. The smallest set P^R_1 consists of a single probability vector that assigns an equal probability $\frac{1}{n}$ to each state, i.e., $P^R_1 = \{(\frac{1}{n}, \ldots, \frac{1}{n})\}$. In turn, the largest set $P^R_{0.0}$ has n extreme points corresponding to the state probabilities $(0, 0, ..., 0, 1)$, $(0, 0, ..., 0, \frac{1}{2}, \frac{1}{2})$, $(0, 0, ...0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, ... , $(\frac{1}{n}, \ldots, \frac{1}{n})$.

Tables 3 represents the out-of-sample performance of the optimal portfolios obtained from problem (18)–(20) using probability sets P^R_α, for $\alpha \in \{1.0, 0.8, \ldots, 0.0\}$. Again, increasing the size of the probability set leads to a decrease in both mean excess returns and active risk. However, unlike P^T_α, using P^R_α results in only a moderate active risk decrease of 4.86%, with a slight improvement of 4.63% in the downside. Surprisingly, the Sortino and Information ratios stay roughly around 1 and 0.40, respectively. One possible explanation is that excess returns and risks (both active and downside) decline approximately at the same rate as the size of the probability set is increased.

Both portfolio turnover and the average number of industries in the optimal portfolio increase with regard to the size increase of P^R_α. On average, using the largest set $P^R_{0.0}$ requires a diversification involving two or three

---

[6] We use the almost SSD formulation by Tzeng et al. (2013): Let $[\underline{t}, \overline{t}] \subset \mathbb{R}$ be an interval containing the outcomes of portfolios X and Y. Then X dominates Y in the sense of ε-ASSD if $\mathbb{E}[X] \geq \mathbb{E}[Y]$ and $\int_{t \in S} [F^2_X(t) - F^2_Y(t)]dt \leq \varepsilon \int_{t \in [\underline{t}, \overline{t}]} |F^2_X(t) - F^2_Y(t)|dt$, where $S = \{t \in [\underline{t}, \overline{t}] \mid F^2_Y(t) < F^2_X(t)\}$.

[7] Although there are several statistical tests for stochastic dominance (see, among others, Davidson 2009; Scaillet & Topaloglou 2010; Linton et al. 2014, Post 2017, Ng et al. 2017), not all of them are well suited to deal with large pairwise correlated samples. Moreover, many tests use a null hypothesis which in our setting would be that the optimized portfolio dominates the benchmark portfolio, i.e., $H_0 : \sum_{j=1}^{49} \lambda_j^* X_j \succ^2 Y$. Rejecting such a null hypothesis would imply either $H_1 : Y \succ^2 \sum_{j=1}^{49} \lambda_j^* X_j$ or $\sum_{j=1}^{49} \lambda_j^* X_j \sim^2 Y$. To avoid difficulties in interpreting the test results, we deploy the statistical test by Davidson (2009), which offers a straightforward interpretation in our test setting.

Table 3: Out-of-Sample Portfolio Performance with 1-month Holding Period and Set of Feasible Probabilities P_α^R

Measures	$P_{1.0}^R$	$P_{0.8}^R$	$P_{0.6}^R$	$P_{0.4}^R$	$P_{0.2}^R$	$P_{0.0}^R$
	Excess returns over benchmark					
Mean (%)	6.11	5.41	5.32	4.97	5.00	4.51
Std. dev. (%)	15.38	13.40	12.32	11.71	11.10	10.52
Skewness	0.89	0.74	0.64	0.40	0.33	0.40
$\text{CVaR}_{5\%}$(%)	-18.67	-17.77	-16.05	-16.07	-14.82	-14.04
CE (%)	5.07	4.60	4.63	4.34	4.43	4.00
RP (%)	1.04	0.81	0.69	0.63	0.57	0.51
Sortino ratio	1.00	0.98	1.12	1.02	1.11	1.04
Information ratio	0.40	0.40	0.43	0.42	0.45	0.43
	Portfolio diversification					
Turnover	0.76	0.89	0.97	1.01	1.04	1.07
# of industries	5.14	5.95	6.51	7.02	7.49	7.91
	Dominance over benchmark					
ε-ASSD (ε)	0.288	0.253	0.224	0.218	0.206	0.204
# of SSDs (%)	14.20	17.23	19.51	19.03	19.22	20.83
	p-values of dominance test					
1^{st} percentile	0.000	0.000	0.000	0.000	0.000	0.000
5^{th} percentile	0.002	0.000	0.000	0.000	0.000	0.000
10^{th} percentile	0.018	0.018	0.008	0.005	0.005	0.002
25^{th} percentile	0.070	0.061	0.055	0.041	0.044	0.034
50^{th} percentile	0.180	0.161	0.169	0.152	0.143	0.141
75^{th} percentile	0.266	0.263	0.270	0.243	0.239	0.249

more industries, which gives rise to a much greater portfolio turnover of 1.07. Additionally, with the largest set $P_{0.0}^R$ out-of-sample dominance over the benchmark is obtained in 20.83% of all holding periods.

5.5. Varying Sample Size

The third test utilizes a set of feasible probabilities that assigns an equal probability to the states corresponding to the k most recent return observations while allowing k to vary between a fixed minimum sample size $\underline{n}$ and the full sample size n. Formally, this set is defined by

$$P_{\underline{n}}^S = \left\{ p \in P^0 \;\middle|\; p = (0, \ldots, 0, \underbrace{\frac{1}{k}, \ldots, \frac{1}{k}}_{k \text{ elements}}), k \in \{\underline{n}, ..., n\} \right\}, \tag{23}$$

where $\underline{n} \in \{1, ..., n\}$ is the minimum sample size. Thus, increasing $\underline{n}$ decreases the size of the feasible probability set, and when $\underline{n} = n$, the set consists of a single vector of equal state probabilities, i.e., $P_n^S = \{(\frac{1}{n}, \ldots, \frac{1}{n})\}$.

Table 4: Out-of-Sample Portfolio Performance with 1-month Holding Period and Set of Feasible Probabilities $P^S_{\underline{n}}$

Measures	P^S_1	$P^S_{0.75}$	$P^S_{0.5}$	$P^S_{0.25}$
	Excess returns over benchmark			
Mean (%)	6.11	5.81	5.44	4.90
Std. dev. (%)	15.38	14.95	13.94	12.32
Skewness	0.89	0.91	0.89	0.62
$\text{CVaR}_{5\%}$(%)	-18.67	-18.38	-16.60	-16.62
CE (%)	5.07	4.83	4.57	4.20
RP (%)	1.04	0.98	0.87	0.70
Sortino ratio	1.00	0.97	0.99	0.95
Information ratio	0.40	0.39	0.39	0.40
	Portfolio diversification			
Turnover	0.76	0.75	0.76	0.81
# of industries	5.14	5.49	5.83	6.32
	Dominance over benchmark			
ε-ASSD (ε)	0.288	0.263	0.249	0.222
# of SSDs (%)	14.20	16.29	18.37	20.27
	p-values of dominance test			
1^{st} percentile	0.000	0.000	0.000	0.000
5^{th} percentile	0.002	0.002	0.000	0.000
10^{th} percentile	0.018	0.018	0.009	0.008
25^{th} percentile	0.070	0.073	0.071	0.067
50^{th} percentile	0.180	0.179	0.166	0.170
75^{th} percentile	0.266	0.263	0.245	0.263

Table 4 reports the out-of-sample portfolio performance of problem (18)–(20) using probability sets $P^S_{\underline{n}}$, for $\underline{n} \in \{1, 0.75, \dots, 0.25\}$. Compared with P^T_α and P^R_α, using $P^S_{\underline{n}}$ offers only modest improvements with regard to active risk as well as to downside risk, 3.06% and 2.05%, respectively. Applying a larger set $P^S_{\underline{n}}$ fails to outperform the smallest set P^S_1 in terms of risk-adjusted returns evaluated by the Sortino and Information ratios. Nevertheless, increasing the size of the probability set results in a higher chance of dominating the benchmark portfolio out-of-sample.

5.6. Sensitivity Analysis

To analyze how the length of the holding period affects the results, we replicated the tests by using both a 3-month and a 6-month holding period, resulting in a total of 352 and 176 out-of-sample observations, respectively. The tests were carried out using all three types of probability sets (P^T, P^R, P^S). The probability set based on the confidence region around state probabilities (P^T) continues to produce portfolios with the best overall out-of-sample performance, and for brevity we report the results based on this probability set only. The out-of-sample performance of the optimal portfolios obtained from problem (18)–(20) using 3-month and 6-month holding periods is presented in Tables 5 and 6, respectively.

Table 5: Out-of-Sample Portfolio Performance with 3-month Holding Period and Set of Feasible Probabilities P_α^T

Measures	$P_{1.00}^T$	$P_{0.98}^T$	$P_{0.96}^T$	$P_{0.94}^T$	$P_{0.92}^T$	$P_{0.90}^T$
	Excess returns over benchmark					
Mean (%)	7.08	6.17	5.39	4.59	4.14	3.87
Std. dev. (%)	14.98	13.35	11.56	10.05	8.94	7.16
Skewness	1.10	1.10	1.13	1.25	1.08	1.22
$\text{CVaR}_{5\%}$(%)	-15.44	-16.13	-11.93	-10.47	-9.61	-6.84
CE (%)	6.12	5.39	4.81	4.14	3.77	3.64
RP (%)	0.96	0.78	0.58	0.45	0.37	0.23
Sortino ratio	1.48	1.35	1.37	1.34	1.32	1.85
Information ratio	0.47	0.46	0.47	0.46	0.46	0.54
	Portfolio diversification					
Turnover	1.15	1.11	1.09	1.06	1.03	0.97
# of industries	5.18	6.21	7.43	9.11	11.46	14.28
	Dominance over benchmark					
ε-ASSD (ε)	0.364	0.269	0.203	0.173	0.147	0.121
# of SSDs (%)	8.52	15.34	19.32	22.16	23.01	25.28
	p-values of dominance test					
1^{st} percentile	0.000	0.001	0.000	0.000	0.000	0.000
5^{th} percentile	0.004	0.007	0.002	0.001	0.002	0.000
10^{th} percentile	0.011	0.012	0.016	0.017	0.010	0.002
25^{th} percentile	0.079	0.074	0.096	0.055	0.039	0.029
50^{th} percentile	0.210	0.181	0.169	0.155	0.123	0.112
75^{th} percentile	0.339	0.268	0.307	0.259	0.218	0.207

The 3-month holding strategy with the smallest set P_1^T (i.e., equal state probabilities) earns the highest mean excess return of roughly 7%, which is consistent with the empirical findings of Post & Kopa (2017). Rebalancing the portfolio at the end of every 3 months, instead of 6 months, yields better performance with regard to risk-adjusted return measures (Sortino and Information ratios) when using a small- or medium-sized probability set (until $P_{0.94}^T$). However, longer holding periods reduce the share of periods in which the optimal portfolios dominate the benchmark. Regardless of the length of the holding period, increasing the size of the probability set increases the likelihood that the optimized portfolio will dominate the benchmark in out-of-sample comparison. Hence, the main finding, i.e., incomplete probability information can help in identifying portfolios whose dominance over the benchmark is robust, is not particularly sensitive to the specification of the holding period.

Table 6: Out-of-Sample Portfolio Performance with 6-month Holding Period and Set of Feasible Probabilities P_α^T

Measures	$P_{1.00}^T$	$P_{0.98}^T$	$P_{0.96}^T$	$P_{0.94}^T$	$P_{0.92}^T$	$P_{0.90}^T$
	Excess returns over benchmark					
Mean (%)	5.79	5.80	4.85	4.45	3.95	3.71
Std. dev. (%)	14.56	13.58	11.80	10.78	9.61	7.67
Skewness	1.26	1.53	1.19	1.07	0.82	1.00
$\text{CVaR}_{5\%}$(%)	-15.13	-13.19	-15.25	-15.04	-13.49	-8.70
CE (%)	4.89	5.02	4.23	3.93	3.53	3.44
RP (%)	0.90	0.78	0.62	0.52	0.42	0.27
Sortino ratio	1.17	1.38	1.14	1.08	1.00	1.40
Information ratio	0.40	0.43	0.41	0.41	0.41	0.48
	Portfolio diversification					
Turnover	1.45	1.42	1.37	1.35	1.32	1.22
# of industries	5.16	6.14	7.41	9.09	11.47	14.25
	Dominance over benchmark					
ε-ASSD (ε)	0.386	0.264	0.213	0.159	0.160	0.137
# of SSDs (%)	6.25	14.20	14.20	18.75	22.16	26.14
	p-values of dominance test					
1^{st} percentile	0.015	0.027	0.007	0.001	0.000	0.000
5^{th} percentile	0.022	0.044	0.025	0.003	0.003	0.000
10^{th} percentile	0.056	0.050	0.030	0.005	0.004	0.001
25^{th} percentile	0.174	0.101	0.057	0.024	0.021	0.017
50^{th} percentile	0.228	0.179	0.185	0.092	0.072	0.092
75^{th} percentile	0.324	0.281	0.327	0.276	0.246	0.220

6. Discussion and Conclusions

In this paper we have developed models to identify portfolios whose stochastic dominance over the benchmark portfolio is not sensitive to the specification of state probabilities. The key idea was to allow state probabilities to take any values within a set of feasible probabilities. We then showed that as long as the convex hull of this set has a finite number of extreme points, a portfolio that dominates a given benchmark portfolio for all probabilities can be identified by solving an LP problem. Hence, powerful commercial implementations of LP solution algorithms can be utilized to identify such optimal portfolios. Furthermore, the requirement of a polyhedral convex hull is not particularly restrictive: The set of feasible probabilities can capture, for instance, a confidence region around the vector of equal probabilities (empirical distribution) or all probability vectors satisfying some pre-specified moment conditions (empirical likelihood approach). Moreover, the use of set-valued probabilities avoids specifying exactly how many historical return observations are included in the state-space. The developed models were illustrated by applying them to empirical data of industry portfolio returns. Results from the application indicate that a portfolio that dominates the market portfolio for a set of feasible state probabilities in-sample is also more likely to dominate the market portfolio out-of-sample.

Although the models developed in this paper are relatively general and conceptually straightforward, they require *a priori* specification of a set of feasible probabilities. However, there are several techniques that help with this specification. For instance, the technique deployed by the application in Section 5 can be used in other applications as well. Historical return observations from any market can be partitioned into training and validation data sets to identify the probability set that yields the best out-of-sample performance in terms of stochastic dominance.

Determining an appropriate set of feasible probabilities can also be supported by expert judgement or empirical stylized facts about the particular application area. In portfolio selection, for instance, Post et al. (2018) suggest that such stylized facts about the financial markets can be operationalized by enforcing moment conditions based on common risk factors (Fama & French 1993). Technically, such conditions correspond to linear constraints on state probabilities and hence they can be directly used to define the set of feasible probabilities in our model. It is important to highlight that *a priori* assumptions about the probability distribution cannot be avoided, as they are necessary when deploying any formal model for portfolio selection. For instance, the common 'plug-in' approach of using empirical probabilities is based on the – often implicit – assumption that the asset returns are identically distributed and serially independent. Relaxing this assumption immediately raises the question of which probabilities should be used in the model. The strength of the model developed here is that it does not require selecting a single probability vector to produce decision recommendations.

The models developed in this paper can be applied in other MS/OR areas beyond financial portfolio diversification. For instance, the management of mixed asset portfolios consisting of both market-traded securities and in-house R&D projects can be supported by MILP models in which the uncertainties related to security prices and the success of R&D projects are modeled with a scenario tree (see, e.g., Gustafsson & Salo 2005). Specifying the probabilities for such scenario trees cannot rely solely on historical financial data but often requires judgmental estimates from multiple experts. Hence, the models developed here could be used to capture a set of feasible probabilities that contains the estimates of each expert and then identify a mixed asset portfolio that dominates a specified target distribution for all feasible probabilities. Similarly, energy investment decisions are made under uncertainties about the future price of CO_2 emissions. These decisions can be supported with MILP models in which uncertainties are captured by scenario trees (see, e.g., Kettunen et al. 2011). The models developed in this paper could be used to find energy investment portfolios that are robust to changes in the level of risk aversion as well as variations in scenario probabilities.

Appendix: Proofs

Proof of Theorem 1. '$\Rightarrow$': Assume $X \succeq_P^2 Y$. Take any $p^* \in \text{ext}(\text{conv}(P))$ and $t \in \mathbb{R}$. Then there exist $p^1, ..., p^{n+1} \in P$ and $(\alpha_1, \ldots, \alpha_{n+1}) \in \mathbb{R}_+^{n+1}$ such that $\sum_{i=1}^{n+1} \alpha_i = 1$ and $p^* = \sum_{j=1}^{n+1} \alpha_j p^j$. Evaluating the difference between the integrals of the CDFs of X and Y at $(t;p^*)$ yields

$$\begin{aligned} F_X^2(t;p^*) - F_Y^2(t;p^*) &= \sum_{i|x_i\le t} p_i^*(t-x_i) - \sum_{i|y_i\le t} p_i^*(t-y_i) = \sum_{i|x_i\le t}\sum_{j=1}^{n+1}\alpha_j p_i^j(t-x_i) - \sum_{i|y_i\le t}\sum_{j=1}^{n+1}\alpha_j p_i^j(t-y_i) \\ &= \sum_{j=1}^{n+1}\alpha_j \sum_{i|x_i\le t} p_i^j(t-x_i) - \sum_{j=1}^{n+1}\alpha_j \sum_{i|y_i\le t} p_i^j(t-y_i) \\ &= \sum_{j=1}^{n+1}\alpha_j\Bigg(\sum_{i|x_i\le t} p_i^j(t-x_i) - \sum_{i|y_i\le t} p_i^j(t-y_i)\Bigg) = \sum_{j=1}^{n+1}\underbrace{\alpha_j}_{\ge 0}\underbrace{\Big(F_X^2(t;p^j) - F_Y^2(t,p^j)\Big)}_{\le 0,\text{ since } p^j\in P} \le 0, \end{aligned}$$

i.e., $F_X^2(t;p^*) \le F_Y^2(t;p^*)$. □

(i) '⇐':Assume $F_X^2(t;p) \le F_Y^2(t;p)$ for all $t\in\mathbb{R}$ and $p\in\{p^1,\dots,p^l\} = \text{ext}(\text{conv}(P))$. Take any $p^*\in P$ and $t\in\mathbb{R}$. Then $p^*\in\text{conv}(P)$ and hence there exists $(\alpha_1,\dots,\alpha_l)\in\mathbb{R}_+^l$ such that $\sum_{i=1}^l \alpha_i = 1$ and $p^* = \sum_{p=1}^l \alpha_j p^j$. Evaluating the difference between the integrals of the CDFs of X and Y at $(t;p^*)$ yields

$$F_X^2(t;p^*) - F_Y^2(t;p^*) = \dots = \sum_{j=1}^{l}\underbrace{\alpha_j}_{\ge 0}\underbrace{\Big(F_X^2(t;p^j) - F_Y^2(t,p^j)\Big)}_{\le 0,\text{ since } p^j\in\text{ext}(\text{conv}(P))} \le 0,$$

i.e., $F_X^2(t;p^*) \le F_Y^2(t;p^*)$.

Proof of Corollary 1. Let $P = P^0$. By Theorem 1 $X \succeq_P^2 Y$ iff $F_X^2(t,p^l) \le F_Y^2(t,p^l)$ for all $t\in\mathbb{R}$ in each extreme point p^l. Evaluating F_X^2 and F_Y^2 at extreme point p^l gives

$$F_X^2(t,p^l) = \begin{cases} 0, & \text{if } t < x_l \\ t - x_l, & \text{if } t \ge x_l \end{cases}, \text{ and } F_Y^2(t,p^l) = \begin{cases} 0, & \text{if } t < y_l \\ t - y_l, & \text{if } t \ge y_l \end{cases}$$

and hence $F_X^2(t,p^l) \le F_Y^2(t,p^l)$ for all $t\in\mathbb{R}$ holds iff $x_l \ge y_l$.

Proof of Theorem 2. '⇒': Assume $X \succeq_P^2 Y$, which by Theorem 1 implies that $F_X^2(t;p) \le F_Y^2(t;p)\ \forall\ t\in\mathbb{R}, p\in\text{ext}(\text{conv}(P))$. Construct $d\in\mathbb{R}_+^{n\times n}$ such that $d_{ik} = \max\{0, y_i - x_k\}$ for each $i,k\in\{1,...,n\}$. Then d clearly satisfies constraint (16). To show that d satisfies constraint (17) we evaluate the LHS for arbitrary $i\in\{1,...,n\}, p\in\text{ext}(\text{conv}(P))$, which gives

$$\sum_{k=1}^n d_{ik}p_k = \sum_{\substack{k\in\{1,\dots,n\}\\ x_k\le y_i}}^{n} (y_i - x_k)p_k = F_X^2(y_i;p),$$

which is less or equal to the RHS $F_Y^2(y_i;p)$. □

'⇐': Assume $d^*\in\mathbb{R}_+^{n\times n}$ satisfies constraints (16) and (17). Then for any $i\in\{1,...,n\}, p\in\text{ext}(\text{conv}(P))$

$$F_Y^2(y_i;p) \ge \sum_{k=1}^n d_{ik}^* p_k \ge \sum_{k=1}^n \max\{0, y_i - x_k\}p_k = F_X^2(y_i;p).$$

Now take any $t\in\mathbb{R}$:

(i) Assume $t \le y_1$. Then $F_X^2(t;p) \le F_X^2(y_1;p) \le F_Y^2(y_1;p) = 0 = F_Y^2(t;y)$, since $F_{(\cdot)}^2$ is always non-negative and non-decreasing.

(ii) Assume $t \in [y_i, y_{i+1}]$ for some $i \in \{1, ..., n-1\}$. On this interval F_Y^2 is linear, F_X^2 is convex and on the end points F_Y is greater than or equal to F_X. Hence, $F_X^2(t;p) \leq F_Y^2(t;p)$.

(iii) Assume $t > y_n$. Then $\frac{\partial}{\partial t} F_Y^2(t;p) = F_Y(t;p) = 1$ and $\frac{\partial}{\partial t} F_X^2(t;p) = F_X(t;p) \leq 1$. Hence, $F_Y^2(y_n;p) \geq F_X^2(y_n;p)$ implies $F_Y^2(t;p) \geq F_X^2(t;p)$.

Thus, $F_Y^2(t;p) \geq F_X^2(t;p)$ for all $t \in \mathbb{R}, p \in \text{ext}(\text{conv}(P))$ which by Theorem 1 implies $X \succeq_P^2 Y$.□

Proof of Corollary 2. Assume exists $X^* = \sum_{j=1}^m \lambda_j^* X_j \in \mathcal{X}$ such that $X^* \succ_P^2 Y$. This implies that $X^* \succeq_P^2 Y$ and furthermore $\mathbb{E}_p[X^*] \geq \mathbb{E}_p[Y]$ for any $p \in P$. By Theorem 2 there exists $d^* \in \mathbb{R}^{n \times n}$ such that (λ^*, d^*) is a feasible solution to problem (18)–(20) and therefore the optimal objective function value is at least $\mathbb{E}_{\hat{p}}[X^*]$. □

Appendix: First-order Stochastic Dominance under Incomplete Probability Information

First-order stochastic dominance (FSD) between two portfolio is established by comparing the cumulative distribution functions of the portfolios' returns. In particular, portfolio X dominates portfolio Y in the sense of FSD if $F_X(t) \leq F_Y(t)$ for all $t \in \mathbb{R}$. No expected utility maximizing investor who prefers higher returns over lower ones would choose a portfolio that is dominated in the sense of FSD. FSD can be extended to admit a set of feasible probability vectors in a similar manner as SSD. This extension is formally presented by the following definition.

Definition 3. Portfolio $X \in \mathcal{X}$ dominates portfolio $Y \in \mathcal{X}$ in the sense of FSD with regard to the set of feasible probabilities $P \subseteq P^0$, denoted by $X \succeq_P^1 Y$, if

$$F_X(t;p) \leq F_Y(t;p) \ \forall \ t \in \mathbb{R}, p \in P.$$

From the definition it is clear that dominance w.r.t. some set of feasible probabilities P implies dominance w.r.t. any subset $P' \subseteq P$. Moreover, it is well known that FSD implies SSD. These two properties can be summarized by

$$\begin{array}{ccc} X \succeq_P^1 Y & \Rightarrow & X \succeq_{P'}^1 Y \\ \Downarrow & & \Downarrow \\ X \succeq_P^2 Y & \Rightarrow & X \succeq_{P'}^2 Y. \end{array}$$

A sufficient and necessary condition for FSD holding for all probability vectors in set P is that it holds everywhere on the 'border' of P's convex hull. This property is formally stated by the following theorem, which uses conv$(\cdot)$ to denote the convex hull and ext$(\cdot)$ to denote the extreme points of a set.

Theorem 3. *Let $P \subseteq P^0$. Then*

$$X \succeq_P^1 Y \Leftrightarrow F_X(t;p) \leq F_Y(t;p) \ \forall \ t \in \mathbb{R}, p \in \text{ext}(\text{conv}(P)).$$

Proof.Equivalent to proof of Theorem 1 where $F_{(\cdot)}^2$ is replaced by $F_{(\cdot)}$. □

If there are no constraints on the state probabilities, then FSD coincides with SSD and state-wise dominance, as stated by the following corollary.

Corollary 3. *Let $P = P^0$. Then*

$$X \succeq_P^1 Y \Leftrightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \geq \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \Leftrightarrow X \succeq_P^2 Y.$$

Proof. The second equivalence is implied by Corollary 1. To prove the first equivalence consider Theorem 3 which implies that $X \succeq_{P^0}^1 Y$ if and only if $F_X(t, p^l) \leq F_Y(t, p^l)$ for all $t \in \mathbb{R}, p \in$ext(conv(P^0)). Since conv(P^0) $= P^0$ these extreme points correspond to the unit vectors $p^1, ..., p^n$ such that $p_i^l = 1$ if $l = i$ and $p_i^l = 0$ if $l \neq i$. In each extreme point the CDFs of X and Y are step functions of the form

$$F_X(t, p^l) = \begin{cases} 0, \text{ if } t < x_l \\ 1, \text{ if } t \geq x_l \end{cases}, \text{ and } F_Y(t, p^l) = \begin{cases} 0, \text{ if } t < y_l \\ 1, \text{ if } t \geq y_l \end{cases}$$

and hence $F_X(t, p^l) \leq F_Y(t, p^l)$ for all $t \in \mathbb{R}$ holds iff $x_l \geq y_l$ for each $l \in \{1, \ldots, n\}$. □

With a finite state-space the CDF of any portfolio with a fixed probability vector $p \in$ ext(conv(P)) is a right continuous non-decreasing step function in which the locations of the steps correspond to the state-specific returns of the portfolio. By assuming that the states are indexed in an increasing order of the benchmark portfolio returns, i.e., $y_1 \leq y_2 \leq \cdots \leq y_n$, the CDF of Y is constant on each interval $[-\infty, y_1), [y_1, y_2), \ldots [y_{i-1}, y_i), \ldots, [y_{n-1}, y_n)$. The maximum value $F_X(t; p)$ obtains in the interval $t \in [y_{i-1}, y_i)$ is equal to the sum of probabilities of those states $k \in \{1, ..., n\}$ in which X has a state-specific return x_k strictly below y_i. These states can be identified by introducing a binary variables z_{ik} for each pair of states $i, k \in \{1, ..., n\}$ and requiring that $z_{ik} = 1$ whenever $x_k < y_i$. This requirement can be implemented as the linear constraint $x_k + z_{ik} M \geq y_i$, where M is a large positive constant. Hence, for X to dominate Y it is necessary that $\sum_{k=1}^n z_{ik} p_k \leq F_Y(y_{i-1}; p)$ for each $i \in \{1, ..., n\}$. The following theorem states that this condition is also sufficient if it holds for each extreme point p of the convex hull of P.

Theorem 4. $X \succeq_P^1 Y$ *if and only if there exists* $z \in \{0,1\}^{n \times n}$ *that satisfies constraints*

$$x_k + M z_{ik} \quad \geq \quad y_i \ \forall \ i, k \in \{1, ..., n\}, \tag{24}$$

$$\sum_{k=1}^n z_{ik} p_k \quad \leq \quad F_Y(y_{i-1}; p) \ \forall \ i \in \{1, ..., n\}, p \in \text{ext}(\text{conv}(P)) \tag{25}$$

where $F_Y(y_0; p) = 0$ *and* M *is a large positive constant.*

Proof. '⇒': Assume $X \succeq_P^1 Y$, which by Theorem 1 implies that $F_X(t; p) \leq F_Y(t; p) \ \forall \ t \in \mathbb{R}, p \in$ ext(conv(P)). Construct $z \in \{0,1\}^{n \times n}$ so that

$$z_{ik} = \begin{cases} 1, \text{ if } x_k < y_i \\ 0, \text{ otherwise} \end{cases} \quad \forall \ i, k \ \in \{1, ..., n\}.$$

Then z clearly satisfies constraint (24). To show that z satisfies constraint (25) we evaluate the LHS for arbitrary $i \in \{1, ..., n\}, p \in$ ext(conv(P)), which gives

$$\sum_{k=1}^n z_{ik} p_k = \sum_{k | x_k < y_i} p_k = \begin{cases} 0, \text{ if } \{x_k | x_k < y_i\} = \emptyset \\ F_X(\max_k \{x_k | x_k < y_i\}; p), \text{ otherwise.} \end{cases}$$

We can then use $F_X(t;p) \leq F_Y(t;p)\ \forall\ t \in \mathbb{R}$ to bound this from above:

$$0 \leq F_X(\max_k\{x_k | x_k < y_i\}; p) \leq F_Y(\max_k\{x_k | x_k < y_i\}; p) \leq \sup_{t<y_i} F_Y(t;p) = F_Y(y_{i-1};p),$$

which is equal to the RHS of constraint (25).

'⇐': We prove the contrapositive: Assume $\neg(X \succeq_P^1 Y)$, which by Theorem 1 implies that $\exists t \in [0, y_n), p \in \text{ext}(\text{conv}(P))$ such that $F_X(t;p) > F_Y(t;p)$. Take any $z \in \{0,1\}^{n\times n}$ and if violates constraint (24), then the proof is complete. In turn, if it satisfies constraint (24), then $z_{ik} = 1$ for all $i, k \in \{1, ..., n\}$ such that $x_k < y_i$. Denote $y_0 = -\infty$ and select $l \in \{1, ..., n\}$ such that $t \in [y_{l-1}, y_l)$. The LHS of constraint (25) for index $i = l$ equals

$$\sum_{k=1}^{n} z_{lk} p_k \geq \sum_{k|x_k<y_l} p_k \geq \sum_{k|x_k \leq t} p_k = F_X(t;p) > F_Y(t;p) = F_Y(y_{l-1};t),$$

which implies constraint (25) are not satisfied. □

Theorem 4 can be used to formulate a MILP problem which for a given benchmark portfolio Y identifies a dominating portfolio $X \in \mathcal{X}$ with the maximal expected return $\mathbb{E}_{\hat{p}}[X] = \sum_{i=1}^{n} p_i x_i$ under some probability vector $\hat{p} \in P$. Substituting $x_k = \sum_{j=1}^{m} \lambda_j x_{jk}$, where $\lambda \in \Lambda$, into constraint (24) and into the expected portfolio return yields

$$\max_{\substack{\lambda \in \Lambda \\ z \in \{0,1\}^{n\times n}}} \quad \sum_{i=1}^{n} \hat{p}_i \sum_{j=1}^{m} \lambda_j x_{ji} \tag{26}$$

$$\sum_{j=1}^{m} \lambda_j x_{jk} + M z_{ik} \geq y_i\ \forall\ i, k \in \{1, ..., n\} \tag{27}$$

$$\sum_{k=1}^{n} z_{ik} p_k \leq F_Y(y_{i-1};p)\ \forall\ i \in \{1, ..., n\}, p \in \text{ext}(\text{conv}(P)),. \tag{28}$$

For the above model to correctly test the existence of a strictly dominating portfolio, the probability vector used in the objective function has to belong to the relative interior of conv(P), formally defined as

$$\text{relint}(\text{conv}(P)) = \{p \in \text{conv}(P) \mid \forall p' \in \text{conv}(P)\ \exists \epsilon > 0 \text{ s.t. } \hat{p} + \epsilon(\hat{p} - p') \in \text{conv}(P)\}. \tag{29}$$

Fortunately, the relative interior is non-empty for every non-empty set (e.g., relint$(\{p\}) = \{p\}$) and hence the requirement of using a probability vector in the relative interior of the set of feasible probabilities does not limit the applicability of the following corollary.

Corollary 4. *Let $\hat{p} \in$ relint(conv(P)). There exists $X \in \mathcal{X}$ such that $X \succ_P^1 Y$ if and only if MILP problem (26)–(28) has an optimal solution (λ^*, z^*) such that $\mathbb{E}_{\hat{p}}[\sum_{j=1}^{m} \lambda_j^* X_j] > \mathbb{E}_{\hat{p}}[Y]$.*

Proof. '⇒' Assume there exists $X = \sum_{j=1}^{m} \lambda_j X_j \in \mathcal{X}$ such that $X \succ_P^1 Y$. Since $X \succeq_P^1 Y$, Theorem 4 implies that there exists $z \in \{0,1\}^{n\times n}$ such that (z, λ) is a feasible solution to problem (26)–(28). Furthermore, since $\neg(Y \sim_P^1 X)$ there must exists $p^* \in P, t^* \in \mathbb{R}$ such that $(F_X(t;p^*) < F_Y(t;p^*))$, which implies $\mathbb{E}_{p^*}[X - Y] > 0$. Since $\hat{p}$ belongs to the relative interior of conv(P), there exists $\epsilon > 0$ such that $p' = \hat{p} + \epsilon(\hat{p} - p^*)$ ∈conv(P). Rearranging this gives

$$\hat{p} = \frac{1}{1+\epsilon} p' + \frac{\epsilon}{1+\epsilon} p^*.$$

Hence, linearity of the expectation operator yields

$$\mathbb{E}_{\hat{p}}[X-Y] = \sum_{i=1}^{n} \hat{p}_i(x_i - y_i) = \sum_{i=1}^{n} \frac{1}{1+\epsilon} p'_i(x_i - y_i) + \sum_{i=1}^{n} \frac{\epsilon}{1+\epsilon} p^*_i(x_i - y_i) =$$

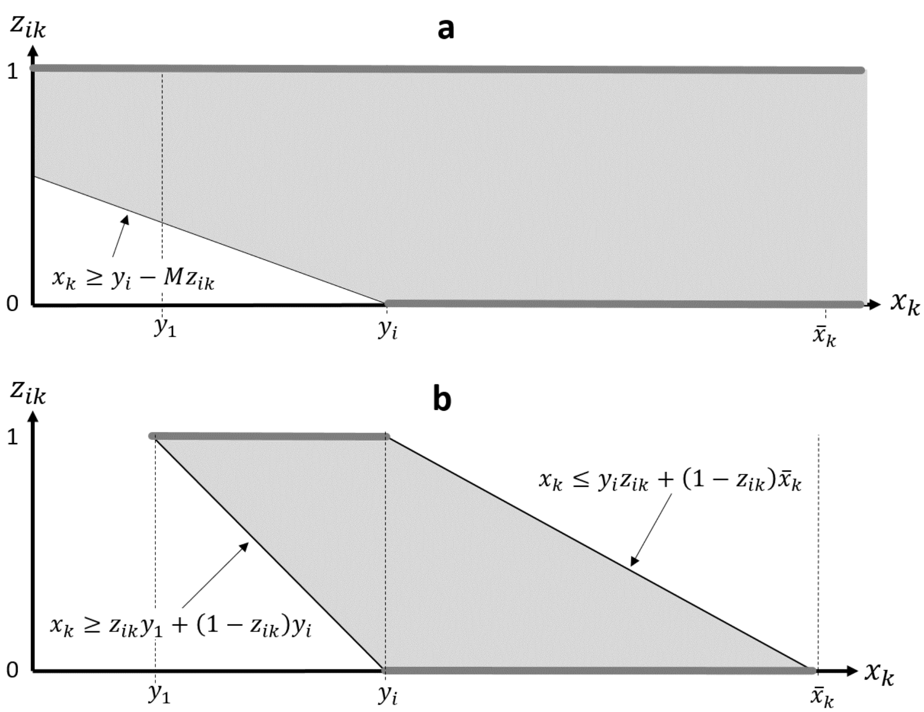
$$= \underbrace{\frac{1}{1+\epsilon}}_{>0} \underbrace{\mathbb{E}_{p'}[X-Y]}_{\geq 0,\text{ since } p' \in P} + \underbrace{\frac{\epsilon}{1+\epsilon}}_{>0} \underbrace{\mathbb{E}_{p^*}[X-Y]}_{>0} > 0,$$

which is equivalent to $\mathbb{E}_{\hat{p}}[X] > \mathbb{E}_{\hat{p}}[Y]$.

'$\Leftarrow$'. Assume (λ^*, z^*) is an optimal solution to problem (26)–(28) such that $\mathbb{E}_{\hat{p}}[\sum_{j=1}^{n} \lambda_j^* X_j] > \mathbb{E}_{\hat{p}}[Y]$. Denote $X^* = \sum_{j=1}^{m} \lambda_j^* X_j$. Since the solution satisfies constraints (27)–(28) Theorem 4 implies $X^* \succeq_P^1 Y$. Furthermore, since $\mathbb{E}_{\hat{p}}[X^*] > \mathbb{E}_{\hat{p}}[Y]$ we have $\neg(Y \succeq_P^1 X^*)$. Together these imply $X^* \succ_P^1 Y$. $\square$

Assuming that the feasible asset weights $\lambda \in \mathbb{R}^m$ are defined by a single constraint ($\sum_j \lambda_j = 1$), the number of constraints in MILP problem (26)–(28) is equal to $1 + n^2 + nq$, where q is the number of extreme points of the set conv(P). The number of decision variables is $m + n^2$, which includes m asset weights λ_j and n^2 auxiliary variables z_{ik}. The efficiency of the formulation can be improved by reducing the feasible region of the continuous relaxation as small as possible without removing the integer optimal solutions. This is exemplified by Figure 5.2(a) presenting the feasible values that constraint (27) allows for the pair of decision variables $x_k = \sum_{j=1}^{m} \lambda_j x_{jk}$ and z_{ik}. Clearly, there are feasible solutions that are known *a priori* not to be optimal.

Figure 3: Figures (a) and (b) show in dark gray the feasible values of decision variables variables $x_k = \sum_{j=1}^{m} \lambda_j x_{jk}$ and z_{ik} under constraint (27) and constraints (30)–(31), respectively. The continuous relaxation of the feasible region is in light gray.



First, Definition 3 directly implies that if any realization of X is below the lowest realization of Y, i.e. y_1, then X does not strictly stochastically dominate Y. However, rather than introducing the additional constraint

$x_k \geq y_1, k \in \{1, ..., n\}$, we instead replace constraint (27) with

$$\sum_{j=1}^{m} \lambda_j x_{jk} \geq z_{ik} y_1 + (1 - z_{ik}) y_i \; \forall \; i \in \{1, ..., n\}, k \in \{1, ..., n\}, \tag{30}$$

which leads to an even smaller feasible region (cf. Figure 5.2(b)). For $i = 1$, constraint (30) yields $x_k \geq y_1, k \in \{1, ..., n\}$, and hence there is no longer need for bookkeeping on which of the X's state-specific returns fall below y_1. Thus, the decision variables $z_{11}, ..., z_{1n}$ can be removed from the problem as well as constraint (28) for the index $i = 1$.

Second, even if x_k is above y_i setting $z_{ik} = 1$ does not violate constraint (27) nor (30), but it can violate constraint (28). If there is an optimal solution to problem (26)–(28) such that $(x_k, z_{ik}) = (x_k^*, 1)$, where $x_k^* > y_i$, then the solution $(x_k, z_{ik}) = (x_k^*, 0)$ is also feasible (cf. Figure 5.2) and it has the same objective function value. Hence, it may be beneficial to ensure that the variables z_{ik} are not set equal to one when x_k is above y_i. This can be implemented with the additional constraint

$$\sum_{j=1}^{m} \lambda_j x_{jk} \leq y_i z_{ik} + (1 - z_{ik}) \overline{x}_k \; \forall \; i \in \{2, ..., n\}, k \in \{1, ..., n\}, \tag{31}$$

where $\overline{x}_k = \max_{X \in \mathcal{X}} X(s_k) = \max_{\lambda \in \Lambda} \sum_{j=1}^{m} \lambda_j x_{jk}$ is the maximum possible return in the kth state. This parameter value can be readily solved *a priori* through linear programming although with the standard market set $\Lambda = \{\lambda \in \mathbb{R}^m \mid \sum_j \lambda_j = 1\}$ the solution is trivial: $\overline{x}_k = \max_j x_{jk}$.

Third, recall that the states are indexed in an ascending order of the benchmark portfolio's returns (cf. equation (15)). Furthermore, constraint (28) requires that $z_{ik} = 1$ for each state k in which the portfolio return $x_k = \sum_{j=1}^{m} \lambda_j x_{jk}$ is strictly below that of the benchmark portfolio in state i. If this is the case, then $x_k < y_i \leq y_{i+1}$ which implies that also $z_{i+1,k} = 1$. Hence, we may introduce the constraint

$$z_{ik} \leq z_{i+1,k} \; \forall \; i \in \{2, ..., n-1\}, k \in \{1, ..., n\} \tag{32}$$

to problem (26)–(28) without removing any feasible integer solutions. With these three modifications MILP problem (26)–(28) becomes

$$\max_{\substack{\lambda \in \Lambda \\ z \in \{0,1\}^{(n-1) \times n}}} \quad \sum_{i=1}^{n} \hat{p}_i \sum_{j=1}^{m} \lambda_j x_{ji} \tag{33}$$

$$\sum_{j=1}^{m} \lambda_j x_{jk} \geq \; y_1 \; \forall \; k \in \{1, ..., n\} \tag{34}$$

$$\sum_{j=1}^{m} \lambda_j x_{jk} \geq y_1 z_{ik} + (1 - z_{ik}) y_i \; \forall \; i \in \{2, ..., n\}, k \in \{1, ..., n\} \tag{35}$$

$$\sum_{j=1}^{m} \lambda_j x_{jk} \leq y_i z_{ik} + (1 - z_{ik}) \overline{x}_k \; \forall \; i \in \{2, ..., n\}, k \in \{1, ..., n\} \tag{36}$$

$$\sum_{k=1}^{n} z_{ik} p_k \leq F_Y(y_{i-1}; p) \; \forall \; i \in \{2, ..., n\}, p \in \text{ext}(\text{conv}(P)) \tag{37}$$

$$z_{ik} \leq z_{i+1,k} \; \forall \; i \in \{2, ..., n-1\}, k \in \{1, ..., n\}. \tag{38}$$

Corollary 5. *Let $\hat{p} \in \operatorname{relint}(\operatorname{conv}(P))$. There exists $X \in \mathcal{X}$ such that $X \succ_P^1 Y$ if and only if MILP problem (33)–(38) has an optimal solution (λ^*, z^*) such that $\mathbb{E}_{\hat{p}}[\sum_{j=1}^m \lambda_j^* X_j] > \mathbb{E}_{\hat{p}}[Y]$.*

Proof. Corollary 4 provides equivalence between the existence of a dominating portfolio and optimal objective function value of MILP problem (26)–(28). Hence, to prove this corollary, it is sufficient to show the following two results: (i) If (z^1, λ) is a feasible solution to MILP problem (26)–(28), then there exists $z^2 \in \{0,1\}^{(n-1)\times n}$ such that (z^2, λ) is a feasible solution to MILP problem (33)–(38), and (ii) if (z^2, λ) is a feasible solution to MILP problem (33)–(38), then there exists $z^1 \in \{0,1\}^{(n\times n)}$ such that (z^2, λ) is a feasible solution to problem (26)–(28). Notice that solutions (z^1, λ) and (z^2, λ) yield equal objective function values, and hence the two results together imply that (z^1, λ) is an optimal solution to problem (26)–(28) if and only if (z^2, λ) is an optimal optimal solution to problem (33)–(38). In what follows we prove these two results.

(i) Assume (z^1, λ) is a feasible solution to MILP problem (26)–(28) and denote $K = \{1, \ldots, n\}$ and $I = \{2, \ldots, n\}$. For $i = 1$, constraint (28) yields $\sum_{k=1}^n z_{1,k}^1 p_k \leq F_Y(y_0; p) = 0$, which implies that for any $k \in K$ $z_{1,k}^1 = 0$. In this case, constraint (27) reduces to $\sum_{j=1}^m \lambda_j x_{jk} \geq y_1$, $k \in K$, and hence λ satisfies constraint (34). Next we show that solution (z^2, λ), where $z^2 \in \{0,1\}^{(n-1)\times n}$ is defined as

$$z_{ik}^2 = \begin{cases} 1 & \text{if } \sum_{j=1}^m \lambda_j x_{jk} < y_i \\ 0 & \text{if } \sum_{j=1}^m \lambda_j x_{jk} \geq y_i \end{cases}, \ i \in I, k \in K, \tag{39}$$

satisfies also constraints (35)–(38).

Constraint (35): Take any $i \in I, k \in K$. If $z_{ik}^2 = 1$, then the constraint reduces to $\sum_{j=1}^m \lambda_j x_{jk} \geq y_1$, which is equivalent constraint (34), which is satisfied. If $z_{ik}^2 = 0$, the constraint reduces to $\sum_{j=1}^m \lambda_j x_{jk} \geq y_i$, which holds by equation (39).

Constraint (36): Take any $i \in I, k \in K$. If $z_{ik}^2 = 1$, by equation (39) $\sum_{j=1}^m \lambda_j x_{jk} < y_i$, and the constraint reduces to $\sum_{j=1}^m \lambda_j x_{jk} \leq y_i$. If $z_{ik}^2 = 0$, the constraint reduces to $\sum_{j=1}^m \lambda_j x_{jk} \leq \overline{x}_k = \max_{\lambda\in\Lambda} \sum_{j=1}^m \lambda_j x_{jk}$, which holds for any $\lambda \in \Lambda$.

Constraint (37): Notice that if $z_{ik}^2 = 1$ for some $i \in I, k \in K$, then $\sum_{j=1}^m \lambda_j x_{jk} < y_i$. Since z_{ik}^1 satisfies constraint (27), $z_{ik}^1 = 1$ must hold. Hence, $z_{ik}^2 \leq z_{ik}^1$ for all $i \in I, k \in K$. Now evaluating the left-hand-side of constraint (37) for any $i \in I$ gives $\sum_{k=1}^n z_{ik}^2 p_k \leq \sum_{k=1}^n z_{ik}^1 p_k \leq F_Y(y_i; p)$ since z^1 satisfies (28).

Constraint (38): If $z_{i,k}^2 = 1$ for some $i \in I \setminus \{n\}, k \in K$, then (39) implies $\sum_{j=1}^m \lambda_j x_{jk} < y_i$. By definition, $y_i \leq y_{i+1}$, and hence (39) implies that $z_{i+1,k}^2 = 1$. Thus $z_{i,k}^2 \leq z_{i+1,k}^2$ for all $i \in I \setminus \{n\}, k \in K$.

(ii) Assume (z^2, λ) is a feasible solution to MILP problem (33)–(38). Define $z^1 \in \{0,1\}^{n\times n}$ such that

$$z_{ik}^1 = \begin{cases} 0 & \text{if } i = 1, k \in K \\ z_{ik}^2 & \text{if } i \in I, k \in K \end{cases}. \tag{40}$$

The solution (z^1, λ) satisfies constraint (27) for $i = 1$ and $k \in K$ since

$$\sum_{j=1}^m \lambda_j x_{jk} + M z_{1,k}^1 = \sum_{j=1}^m \lambda_j x_{jk} \geq y_1,$$

where the last inequality holds since λ satisfies (34). This solution also satisfies constraint (27) for any $i \in I$ and $k \in K$ since

$$y_i - M z_{ik}^1 = y_i - M z_{ik}^2 \leq y_i - \underbrace{(y_i - y_1)}_{<M} z_{ik}^2 = y_1 z_{ik}^1 + (1 - z_{ik}^2) y_i \leq \sum_{j=1}^{m} \lambda_j x_{jk},$$

where the last inequality holds since (z^2, λ) satisfies (35). The solution satisfies constraint (28) since $\sum_{k=1}^{n} z_{0,k}^1 p_k = 0$ and for any $i \in I$

$$\sum_{k=1}^{n} z_{ik}^1 p_k = \sum_{k=1}^{n} z_{ik}^2 p_k \leq F_Y(y_{i-1}; p),$$

since z^2 satisfies (37). □

Acknowledgements

Peng Xu would like to gratefully acknowledge generous financial support from the Foundation for Economic Education (Liikesivistysrahasto, grant no. 150338) for this research.

References

Aboidi, R., & Thon, D. (1994). Efficient algorithms for stochastic dominance tests based on financial market data. *Management Science*, *40*, 508–515.

Baptista, S., Barbosa-Povoa, A. P., Escudero, L. F., Gomes, M. I., & Pizarro, C. (2019). On risk management of a two-stage stochastic mixed 0-1 model for the closed-loop supply chain design problem. *European Journal of Operational Research*, *274*, 91 – 107.

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*, 322–329.

Davidson, R. (2009). Testing for restricted stochastic dominance: Some further results. *Review of Economic Analysis*, *1*, 34–59.

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

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

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*, 893–900.

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

Fang, Y., & Post, T. (2017). Higher-degree stochastic dominance optimality and efficiency. *European Journal of Operational Research*, *261*, 984–993.

Gustafsson, J., & Salo, A. (2005). Contingent portfolio programming for the management of risky projects. *Operations Research*, *53*, 946–956.

Gutjahr, W. (2015). Bi-objective multi-mode project scheduling under risk aversion. *European Journal of Operational Research*, *246*, 421–434.

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

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

Huang, R. J., Tzeng, L., Wang, J.-Y., & Zhao, L. (2020). Operational asymptotic stochastic dominance. *European Journal of Operational Research*, *280*, 312 – 322.

Jiang, Y., Liang, X., Liang, H., & Yang, N. (2018). Multiple criteria decision making with interval stochastic variables: A method based on interval stochastic dominance. *European Journal of Operational Research*, *271*, 632–643.

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

Kallio, M., & Hardoroudi, N. D. (2019). Advancements in stochastic dominance efficiency tests. *European Journal of Operational Research*, *276*, 790 – 794.

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

Kettunen, J., Bunn, D., & Blyth, W. (2011). Investment propensities under carbon policy uncertainty. *The Energy Journal*, *32*.

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

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

Levy, H., & Hanoch, G. (1970). Relative effectiveness of efficiency criteria for portfolio selection. *Journal of Financial and Quantitative Analysis*, *5*, 63–76.

Levy, M. (2019). Stocks for the log-run and constant relative risk aversion preferences. *European Journal of Operational Research*, *277*, 1163 – 1168.

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.

Linton, O., Post, T., & Whang, Y.-J. (2014). Testing for the stochastic dominance efficiency of a given portfolio. *Econometrics Journal*, *17*, S59–S74.

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*, 3549–3554.

Lozano, S., & Gutiérrez, E. (2008). Data envelopment analysis of mutual funds based on second-order stochastic dominance. *European Journal of Operational Research*, *189*, 230–244.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance*, *7*, 77–91.

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

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

Ng, P., Wong, W.-K., & Xiao, Z. (2017). Stochastic dominance via quantile regression with applications to investigate arbitrage opportunity and market efficiency. *European Journal of Operational Research*, *261*, 666–678.

Pearman, A., & Kmietowicz, Z. (1986). Stochastic dominance with linear partial information. *European Journal of Operational Research*, *23*, 57 – 63.

Post, T. (2003). Empirical tests for stochastic dominance efficiency. *The Journal of Finance*, *58*, 1905–1932.

Post, T. (2017). Empirical tests for stochastic dominance optimality. *Review of Finance*, (pp. 793–810).

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

Post, T., & Kopa, M. (2013). General linear formulations of stochastic dominance criteria. *European Journal of Operational Research*, *230*, 321 – 332.

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

Post, T., & Potí, V. (2017). Portfolio analysis using stochastic dominance, relative entropy, and empirical likelihood. *Management Science*, *63*, 153–165.

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

Scaillet, O., & Topaloglou, N. (2010). Testing for stochastic dominance efficiency. *Journal of Business and Economic Statistics*, *28*, 169–180.

Shalit, H., & Yitzhaki, S. (1994). Marginal conditional stochastic dominance. *Management Science*, *40*, 670–684.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, *38*, pp. 332–382.

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

Tzeng, L., Huang, R., & Shih, P. (2013). Revisiting almost second-degree stochastic dominance. *Management Science*, *59*, 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*, 205–220.

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.