

STABILITY AND SENSITIVITY OF STOCHASTIC DOMINANCE CONSTRAINED OPTIMIZATION MODELS

DARINKA DENTCHEVA* AND WERNER RÖMISCH†

Abstract. We consider convex optimization problems with k th order stochastic dominance constraints for $k \geq 2$. We discuss distances of random variables that are relevant for the dominance relation and establish quantitative stability results for optimal values and solution sets in terms of a suitably selected probability metrics. Moreover, we provide conditions ensuring that the optimal value function is Hadamard directionally differentiable. Finally, we discuss some implications of the results for empirical (Monte Carlo, sample average) approximations of dominance constrained optimization models.

1. Introduction. We analyze convex optimization models with stochastic dominance constraints of second and higher order formulated as follows:

$$(1.1) \quad \min\{f(x) : x \in D, G(x, \xi) \succeq_{(k)} Y\}.$$

Here D is a nonempty closed convex subset of \mathbb{R}^m , Ξ is a closed subset of \mathbb{R}^s , $f : \mathbb{R}^m \rightarrow \mathbb{R}$ is a convex function. The mapping $G : \mathbb{R}^m \times \mathbb{R}^s \rightarrow \mathbb{R}$ is continuous, concave with respect to the first argument and satisfies the linear growth condition

$$(1.2) \quad |G(x, z)| \leq K(B) \max\{1, \|z\|\} \quad (x \in B, z \in \Xi)$$

for every bounded subset $B \subset \mathbb{R}^m$ and some constant $K(B)$ (depending on B). Further, ξ denotes a s -dimensional random vector with support Ξ and Y is a real random variable on some probability space. The relation $\succeq_{(k)}$ is the stochastic dominance of order k , $k \in \mathbb{N}$. The constraint $G(x, \xi) \succeq_{(k)} Y$ is our main focus. It indicates that the random variable $G(x, \xi)$ is stochastically larger than the random variable Y , which plays the role of a benchmark outcome with an acceptable probability distribution. We assume that $k \geq 2$ and that both ξ and Y have finite moments of order $k - 1$. The stability properties of such models with respect to perturbations of the underlying probability distributions is an important issue as in many practical situations the distributions of Y and ξ are modeled on the basis of observations or experiments. Our goal is to study stability and sensitivity of the optimal value and the optimal solutions of these problems when the probability distributions involved are approximated. We establish conditions for stability and sensitivity of the optimal value, as well as limit theorems of its empirical estimates.

The relation of *stochastic dominance* is a fundamental concept of statistics, decision theory, and economics. A random variable X *dominates* another random variable Y in the k -order, which we write $X \succeq_{(k)} Y$, if

$$(1.3) \quad \mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$$

for every nondecreasing function $u(\cdot)$ from a certain set of functions, called a generator of the order $\succeq_{(k)}$. The precise definition and equivalent characterizations of the relation are given in Section 2.

*Stevens Institute of Technology, Department of Mathematical Sciences, Hoboken, NJ 07030, U.S.A.

†Humboldt-University Berlin, Department of Mathematics, 10099 Berlin, Germany

Stochastic dominance originated from the theory of majorization [17]. This relation was introduced to statistics in [25] and further developed in the context of statistical inference in [24, 2]. Numerous studies on statistical inference for stochastic dominance are available; we refer to [9, 18] and the references *ibid.* The stochastic dominance relation expresses risk-aversion, which is the basis of its popularity in economic studies. It plays a fundamental role in the inequality and poverty analysis for comparison of income distributions and investment decisions (see, e.g., [12, 23, 26, 43]). In [30, 45, 46], stochastic dominance has been applied in the area of agriculture and insurance. We refer the reader to the monographs [27, 40] for an overview on stochastic orders and the stochastic dominance relation in particular.

The study of optimization problems with stochastic dominance constraints has been initiated in [5] and continued in several papers, e.g., [6], [8]. An optimization model with stochastic dominance constraints has been applied to financial optimization in [7] and to electricity market models in [1, 15, 14]. In [15, 14], two-stage problems with stochastic ordering constraints on the recourse function were considered. Stability results with respect to perturbations of the underlying probability distribution are obtained in [4] for optimization problems with first order stochastic dominance constraints. The recent paper [19] studies stability of optimization problems with second order stochastic dominance constraints and, in particular, the behavior of empirical approximations for such models.

In the present paper, we consider second and higher order dominance constraints simultaneously and, in contrast to [19], we show stability with respect to distances of probability distributions that are associated with such models in a natural way. As in [4], we also study sensitivity of the optimal value function with respect to the probability measures involved. More precisely, we establish Hadamard directional differentiability of the optimal value function with respect to the underlying measures. This property allows to apply the delta method for deriving limit theorems of optimal values if limit theorems for the random inputs are available. This approach is used to derive a limit theorem for empirical (Monte Carlo or sample average) approximations of k th order stochastic dominance models.

Throughout the paper, \mathbb{N} denotes the set of natural numbers and \mathbb{R}^n stands for the n -dimensional real Euclidean space, $n \in \mathbb{N}$. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the space of random variables with finite k -th moments defined on Ω is denoted by $\mathcal{L}_k(\Omega, \mathcal{F}, \mathbb{P})$. For a closed subset Ξ of a Euclidean space, $\mathcal{P}(\Xi)$ denotes the set of all Borel probability measures on Ξ . The expected value of a real random variable X is denoted by $\mathbb{E}(X)$. The Banach space of real-valued continuous functions over a compact set B equipped with the norm $\|\cdot\|_\infty$ is denoted by $\mathcal{C}(B)$, where $\|f\|_\infty = \sup_{x \in B} |f(x)|$ for $f \in \mathcal{C}(B)$. The notation $\mathcal{C}^k(\mathbb{R}^n)$ stands for the set of k -times continuously differentiable real-valued functions defined on \mathbb{R}^n .

Our paper is organized as follows. In section 2 we review the definitions of stochastic dominance, suitable probability metrics and some relations among them. In section 3, we establish Lipschitz-continuity properties of the optimal value, the feasible set, and the set of optimal solutions of the dominance constrained optimization problem, when the distributions subjected to perturbation. Section 4 contains optimality conditions for convex optimization problems with dominance constraints, which are used to establish Hadamard directional differentiability of the optimal value mapping. The implication of our results for the empirical approximations of optimization problems with dominance constraints are discussed in section 5.

2. Stochastic dominance constraints and metrics. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a scalar random variable X defined on it, the real function $F_X^{(1)} = F_X$ denotes the probability distribution function of X , i.e.,

$$(2.1) \quad F_X^{(1)}(\eta) = \mathbb{P}(\{X \leq \eta\}) = \int_{-\infty}^{\eta} P_X(dt) \quad (\forall \eta \in \mathbb{R}).$$

Here $P_X = \mathbb{P} \circ X^{-1}$ denotes the probability measure on \mathbb{R} induced by X . For any $k \in \mathbb{N}$, we define recursively

$$(2.2) \quad F_X^{(k+1)}(\eta) = \int_{-\infty}^{\eta} F_X^{(k)}(t) dt \quad (\forall \eta \in \mathbb{R}).$$

The k th degree stochastic dominance (kSD) relation $\succeq_{(k)}$ is defined by

$$(2.3) \quad X \succeq_{(k)} Y \quad \Leftrightarrow \quad F_X^{(k)}(\eta) \leq F_Y^{(k)}(\eta) \quad (\forall \eta \in \mathbb{R})$$

for any pair (X, Y) of real random variables for which $F_X^{(k)}$ and $F_Y^{(k)}$ are finite. If X and Y belong to $\mathcal{L}_{k-1}(\Omega, \mathcal{F}, \mathbb{P})$, then these functions are finite due to the representation

$$(2.4) \quad F_X^{(k+1)}(\eta) = \frac{1}{k!} \int_{-\infty}^{\eta} (\eta - t)^k P_X(dt) = \frac{1}{k!} \|\max\{0, \eta - X\}\|_k^k \quad (\forall \eta \in \mathbb{R}),$$

which is valid for any $X \in \mathcal{L}_k(\Omega, \mathcal{F}, \mathbb{P})$ and $k \geq 1$. The norm $\|\cdot\|_k$ in \mathcal{L}_k is defined by

$$\|X\|_k = (\mathbb{E}(|X|^k))^{\frac{1}{k}} \quad (\forall k \geq 1).$$

Moreover, the function $F_X^{(k)}$ is nondecreasing for $k \geq 1$ and convex for $k \geq 2$. These and further properties of $F_X^{(k)}$ are discussed, e.g., in [28, 35].

For every $k \in \mathbb{N}$, the stochastic dominance relation $\succeq_{(k)}$ introduces a partial order in $\mathcal{L}_{k-1}(\Omega, \mathcal{F}, \mathbb{P})$, which is not generated by a convex cone in that space (see [5]).

In Section 3, we use distances of real random variables that are relevant for k th degree stochastic dominance constraints. Suitable distances are the *Rachev metrics* (see [34, Section 4.4]) defined by

$$(2.5) \quad \mathbb{D}_{k,p}(X, Y) := \begin{cases} \left(\int_{\mathbb{R}} |F_X^{(k)}(\eta) - F_Y^{(k)}(\eta)|^p d\eta \right)^{\frac{1}{p}} & \text{for } 1 \leq p < \infty \\ \sup_{\eta \in \mathbb{R}} |F_X^{(k)}(\eta) - F_Y^{(k)}(\eta)| & \text{for } p = \infty \end{cases}$$

for all X and Y in $\mathcal{L}_{k-1}(\Omega, \mathcal{F}, \mathbb{P})$ and $k \in \mathbb{N}$. Moreover, it holds

$$(2.6) \quad \mathbb{D}_{k,p}(X, Y) = \zeta_{k,p}(X, Y) := \sup_{g \in \mathcal{D}_{k,p}} \left| \int_{\mathbb{R}} g(t) P_X(dt) - \int_{\mathbb{R}} g(t) P_Y(dt) \right|$$

if $\mathbb{E}(X^i) = \mathbb{E}(Y^i)$, $i = 1, \dots, k-1$ ([31, Lemma 17.1.1]). Here, $\mathcal{D}_{k,p}$ denotes the set of continuous functions $g: \mathbb{R} \rightarrow \mathbb{R}$ that have measurable k th order derivatives $g^{(k)}$ on \mathbb{R} such that

$$\int_{\mathbb{R}} |g^{(k)}(x)|^{\frac{p}{p-1}} dx \leq 1 \quad (p \in (1, \infty]) \quad \text{or} \quad \text{ess sup}_{x \in \mathbb{R}} |g^{(k)}(x)| \leq 1 \quad (p = 1).$$

We note that the condition $\mathbb{E}(X^i) = \mathbb{E}(Y^i)$, $i = 1, \dots, k-1$, is implied by the finiteness of $\zeta_{k,p}(X, Y)$, since $\mathcal{D}_{k,p}$ contains all polynomials of degree $k-1$. Conversely, if X and Y belong to \mathcal{L}_{k-1} and $\mathbb{E}(X^i) = \mathbb{E}(Y^i)$, $i = 1, \dots, k-1$, holds, then the distance $\mathbb{D}_{k,p}(X, Y)$ is finite.

The metric $\zeta_{k,p}$ is an ideal metric of order $r = k-1 + \frac{1}{p}$, which means that

$$\zeta_{k,p}(cX, cY) \leq |c|^r \zeta_{k,p}(X, Y)$$

holds for all real random variables X and Y and real numbers $c \neq 0$. The following estimates are known for the metrics $\zeta_{k,p}$, $k \in \mathbb{N}$ and $p \in [1, \infty]$:

$$(2.7) \quad \zeta_{k,p}(X, Y) \leq \zeta_{1,p}(X, Y)$$

$$(2.8) \quad \zeta_{1,p}(X, Y) \leq c_{k,p} \zeta_{k,p}(X, Y)^{\frac{1}{p(k-1)+1}} \quad (p < \infty)$$

$$(2.9) \quad \zeta_{1,\infty}(X, Y) \leq C_k \zeta_{k,\infty}(X, Y)^{\frac{1}{k}},$$

where $c_{k,p}$ (only depending on k and p) and C_k (only depending on k) are positive constants. The estimate (2.7) follows by definition, (2.8) is proved as Theorem 9 in [20, Sect. 3.10] and (2.9) in [31, Lemma 17.1.8].

The distance $\mathbb{D}_{k,\infty}$ is known as *stop-loss metric of order k* in risk theory (e.g., [13]) and the distance $\zeta_{k,1}$ was introduced in [47, 48] and further discussed in [31, Part IV] and [33, Chapter 6]. We note that $\zeta_{1,1}$ and $\zeta_{1,\infty}$ correspond to the first order *Fortet-Mourier* and *Kolmogorov metric*, respectively.

The following estimates are valid for the ideal metrics $\zeta_{k,1}$ (see [48, Section 1.4]):

$$(2.10) \quad \pi(P_X, P_Y)^{1+k} \leq \hat{C}_k \zeta_{k,1}(X, Y)$$

$$(2.11) \quad |\mathbb{E}(X^k) - \mathbb{E}(Y^k)| \leq k! \zeta_{k,1}(X, Y)$$

for some constant $\hat{C}_k > 0$, where π denotes the Prohorov distance metrizing the topology of weak convergence of probability measures on \mathbb{R} . Hence, a sequence of random variables converges with respect to $\zeta_{k,1}$ iff their probability distributions converge weakly according to (2.10) and their k th moments converge due to (2.11).

Finally, we mention that so-called quasi-semidistances are developed in Chapter 8 of the recent monograph [35] that metrize the dual of the k th order stochastic dominance. The distances considered here appear in [35] as upper bounds of the relevant quasi-semidistances.

3. Stability results. If the pair (ξ, Y) belongs to $\mathcal{L}_{k-1} \times \mathcal{L}_{k-1}$ the stochastic dominance relation (2.3) allows the following equivalent reformulation of the optimization model (1.1)

$$(3.1) \quad \min \left\{ f(x) : x \in D, F_{G(x,\xi)}^{(k)}(\eta) \leq F_Y^{(k)}(\eta), \forall \eta \in \mathbb{R} \right\}$$

as semi-infinite optimization problem. The feasible set is convex and closed as we verify next.

PROPOSITION 3.1. *The general assumptions in Section 1 imply that the feasible set of (3.1) defined as*

$$(3.2) \quad \mathcal{X}(\xi, Y) = \left\{ x \in D : F_{G(x,\xi)}^{(k)}(\eta) \leq F_Y^{(k)}(\eta), \forall \eta \in \mathbb{R} \right\}$$

is closed and convex.

Proof. We notice that $F_X^{(k)}$ monotonically decreases when X increases a.s. Furthermore, the mapping $X \mapsto F_X^{(k)}$ is convex by virtue of [28, Proposition 3]. Therefore, the composition mapping $x \mapsto F_{G(x,\xi)}^{(k)}$ is convex due to the concavity of the function $G(\cdot, \xi)$. This proves the convexity of $\mathcal{X}(\xi, Y)$. We recall that

$$F_{G(x,\xi)}^{(k)}(\eta) = \frac{1}{(k-1)!} \|\max\{0, \eta - G(x, \xi)\}\|_{k-1}^{k-1} = \frac{1}{(k-1)!} \|Z_x\|_{k-1}^{k-1},$$

where $Z_x = \max\{0, \eta - G(x, \xi)\}$. For a convergent sequence $x^n \in \mathcal{X}(\xi, Y)$ and $\lim_{n \rightarrow \infty} x^n = \bar{x}$, we obtain that Z_{x^n} converges a.s. to $Z_{\bar{x}} = \max\{0, \eta - G(\bar{x}, \xi)\}$. As almost sure convergence together with (1.2) implies convergence in $(k-1)$ th mean, we obtain that

$$\lim_{n \rightarrow \infty} F_{G(x^n, \xi)}^{(k)}(\eta) = F_{G(\bar{x}, \xi)}^{(k)}(\eta) \quad \text{for all } \eta \in \mathbb{R}.$$

Thus, $F_{G(\bar{x}, \xi)}^{(k)}(\eta) \leq F_Y^{(k)}(\eta)$ and $\bar{x} \in \mathcal{X}(\xi, Y)$. \square

Clearly, Proposition 3.1 remains valid if \mathbb{R} in (3.1) is replaced by a compact interval I of \mathbb{R} . Following [5], we focus on the relaxed problem

$$(3.3) \quad \max \left\{ f(x) : x \in D, F_{G(x,\xi)}^{(k)}(\eta) \leq F_Y^{(k)}(\eta), \forall \eta \in I \right\}.$$

We denote the optimal value of (3.3) by $v(\xi, Y)$ and its solution set by $S(\xi, Y)$. Note that all these quantities depend only on the probability distribution P_ξ and P_Y of ξ and Y . The latter means that we may write $\mathcal{X}(P_\xi, P_Y)$, $v(P_\xi, P_Y)$ and $S(P_\xi, P_Y)$ instead of $\mathcal{X}(\xi, Y)$, $v(\xi, Y)$ and $S(\xi, Y)$. In this section we prefer the latter notation, while in Section 4 we focus on sensitivity with respect to probability distributions.

For deriving our stability results in what follows, we utilize the *kth order uniform dominance condition (kudc)* at the pair (ξ, Y) , which is introduced in [5] as a constraint qualification condition. Problem (3.3) satisfies kudc at (ξ, Y) if a point $\bar{x} \in D$ exists such that

$$(3.4) \quad \min_{\eta \in I} \left(F_Y^{(k)}(\eta) - F_{G(\bar{x}, \xi)}^{(k)}(\eta) \right) > 0.$$

Condition (3.4) is the *Slater condition* for (3.3) (see also [16, 19]).

We use the following distances for measuring perturbations in $\mathcal{L}_{k-1} \times \mathcal{L}_{k-1}$. For two pairs (ξ, Y) and $(\tilde{\xi}, \tilde{Y})$, we define

$$d_k((\xi, Y), (\tilde{\xi}, \tilde{Y})) = \ell_{k-1}(\xi, \tilde{\xi}) + \mathbb{D}_{k,\infty}(Y, \tilde{Y}),$$

where $k \geq 2$ is the degree of the stochastic dominance relation and ℓ_{k-1} is the \mathcal{L}_{k-1} -minimal distance (or *Wasserstein distance of order $k-1$*) defined by

$$\ell_{k-1}(\xi, \tilde{\xi}) := \inf \left\{ \|\zeta - \tilde{\zeta}\|_{k-1} : P_\zeta = P_\xi, P_{\tilde{\zeta}} = P_{\tilde{\xi}} \right\}.$$

We may write $\ell_{k-1}(P_\xi, P_{\tilde{\xi}})$ and $\mathbb{D}_{k,\infty}(P_Y, P_{\tilde{Y}})$ instead of $\ell_{k-1}(\xi, \tilde{\xi})$ and $\mathbb{D}_{k,\infty}(Y, \tilde{Y})$, respectively.

Our first stability result states a quantitative continuity property of the feasible set mapping with respect to the distance d_k , where the Pompeiu-Hausdorff metric (denoted d_H) is employed to measure the distance of (bounded) sets.

PROPOSITION 3.2. *Let D be compact and assume that the function G satisfies*

$$|G(x, z) - G(x, \tilde{z})| \leq L_G \|z - \tilde{z}\|$$

for all $x \in D$, $z, \tilde{z} \in \Xi$ and some constant $L_G > 0$. Furthermore, we assume that the k th order uniform dominance condition is satisfied at the pair (ξ, Y) .

Then there exist constants $L > 0$ and $\delta > 0$ such that

$$(3.5) \quad d_{\mathbb{H}}(\mathcal{X}(\xi, Y), \mathcal{X}(\tilde{\xi}, \tilde{Y})) \leq L d_k((\xi, Y), (\tilde{\xi}, \tilde{Y})),$$

whenever the pair $(\tilde{\xi}, \tilde{Y})$ is chosen such that $d_k((\xi, Y), (\tilde{\xi}, \tilde{Y})) < \delta$.

Proof. First we consider the set-valued mapping \mathcal{F}

$$\mathcal{F}(x) = \begin{cases} \left\{ r \in \mathbb{R} : F_{G(x, \xi)}^{(k)}(\eta) \leq F_Y^{(k)}(\eta) + r, \forall \eta \in I \right\} & \text{if } x \in D \\ \emptyset & \text{if } x \notin D \end{cases}$$

from \mathbb{R}^m to \mathbb{R} . The graph of \mathcal{F} and the range of \mathcal{F} are of the form

$$\begin{aligned} \text{gph } \mathcal{F} &= \left\{ (x, r) \in D \times \mathbb{R} : \max_{\eta \in I} \left(F_{G(x, \xi)}^{(k)}(\eta) - F_Y^{(k)}(\eta) \right) \leq r \right\} \\ \mathcal{F}(\mathbb{R}^m) &= \left[\inf_{x \in D} \max_{\eta \in I} \left(F_{G(x, \xi)}^{(k)}(\eta) - F_Y^{(k)}(\eta) \right), +\infty \right). \end{aligned}$$

In particular, $\text{gph } \mathcal{F}$ is convex and closed and 0 belongs to the interior of $\mathcal{F}(\mathbb{R}^m)$ due to the uniform dominance condition. The Robinson-Ursescu theorem [37, Theorem 9.48] then implies the existence of $\hat{\delta} > 0$ and $a, b \in \mathbb{R}_+$ such that

$$(3.6) \quad d(x, \mathcal{F}^{-1}(r)) \leq (a\|x - \tilde{x}\| + b)d(r, \mathcal{F}(x))$$

holds whenever $|r| \leq \hat{\delta}$, $x \in D$ and $\tilde{x} \in \mathcal{F}^{-1}(0) = \mathcal{X}(\xi, Y)$. As D is bounded, setting $\hat{L} := a \text{diam } D + b$ in (3.6) we obtain the estimate

$$(3.7) \quad d(x, \mathcal{F}^{-1}(r)) \leq \hat{L} d(r, \mathcal{F}(x)) = \hat{L} \max \left\{ 0, \max_{\eta \in I} \left(F_{G(x, \xi)}^{(k)}(\eta) - F_Y^{(k)}(\eta) \right) - r \right\}$$

whenever $|r| \leq \hat{\delta}$ and $x \in D$. Hence, (3.7) implies for any $x \in \mathcal{X}(\tilde{\xi}, \tilde{Y})$

$$\begin{aligned} d(x, \mathcal{X}(\xi, Y)) &= d(x, \mathcal{F}^{-1}(0)) \leq \hat{L} \max \left\{ 0, \max_{\eta \in I} \left(F_{G(x, \xi)}^{(k)}(\eta) - F_Y^{(k)}(\eta) \right) \right\} \\ &\leq \hat{L} \max \left\{ 0, \min_{\eta \in I} \left(F_{\tilde{Y}}^{(k)}(\eta) - F_{G(x, \tilde{\xi})}^{(k)}(\eta) \right) + \max_{\eta \in I} \left(F_{G(x, \xi)}^{(k)}(\eta) - F_Y^{(k)}(\eta) \right) \right\} \\ &\leq \hat{L} \left(\max_{\eta \in I} \left| F_{G(x, \tilde{\xi})}^{(k)}(\eta) - F_{G(x, \xi)}^{(k)}(\eta) \right| + \mathbb{D}_{k, \infty}(Y, \tilde{Y}) \right). \end{aligned}$$

Furthermore, the equivalence of $x \in \mathcal{F}^{-1}(r(x))$ and $x \in \mathcal{X}(\tilde{\xi}, \tilde{Y})$ holds if

$$r(x) := \max_{\eta \in I} \left(F_{G(x, \xi)}^{(k)}(\eta) - F_Y^{(k)}(\eta) \right) - \max_{\eta \in I} \left(F_{G(x, \tilde{\xi})}^{(k)}(\eta) - F_{\tilde{Y}}^{(k)}(\eta) \right).$$

For any $x \in D$, we have the estimate

$$|r(x)| \leq \max_{\eta \in I} \left| F_{G(x, \xi)}^{(k)}(\eta) - F_{G(x, \tilde{\xi})}^{(k)}(\eta) \right| + \mathbb{D}_{k, \infty}(Y, \tilde{Y}).$$

Hence, we obtain from (3.7) for any $x \in \mathcal{X}(\xi, Y)$

$$\begin{aligned} d(x, \mathcal{X}(\tilde{\xi}, \tilde{Y})) &\leq \hat{L} \max \left\{ 0, \max_{\eta \in I} \left(F_{G(x, \xi)}^{(k)}(\eta) - F_{\tilde{Y}}^{(k)}(\eta) \right) - r(x) \right\} \\ &\leq \hat{L} \max \left\{ 0, \max_{\eta \in I} \left(F_{G(x, \tilde{\xi})}^{(k)}(\eta) - F_{\tilde{Y}}^{(k)}(\eta) \right) \right\} \\ &\leq \hat{L} \left(\max_{\eta \in I} \left| F_{G(x, \tilde{\xi})}^{(k)}(\eta) - F_{G(x, \xi)}^{(k)}(\eta) \right| + \mathbb{D}_{k, \infty}(Y, \tilde{Y}) \right). \end{aligned}$$

In a final step, we derive an estimate for

$$\max_{\eta \in I} \left| F_{G(x, \tilde{\xi})}^{(k)}(\eta) - F_{G(x, \xi)}^{(k)}(\eta) \right|.$$

for any $x \in D$. To this end, we consider the real function $t \rightarrow \max\{0, \eta - t\}^{k-1}$ for fixed $\eta \in I$. The function is differentiable for $k > 2$ and we obtain from the mean value theorem for any $\eta \in I$:

$$(3.8) \quad \begin{aligned} |\max\{0, \eta - t\}^{k-1} - \max\{0, \eta - \tilde{t}\}^{k-1}| &\leq (k-1) \max\{|\eta - t|, |\eta - \tilde{t}|\}^{k-2} |t - \tilde{t}| \\ &\leq K_I (k-1) \max\{1, |t|, |\tilde{t}|\}^{k-2} |t - \tilde{t}| \end{aligned}$$

for some constant $K_I \geq 1$ (depending on I) and all $t, \tilde{t} \in \mathbb{R}$. Clearly, (3.8) also holds for $k = 2$. We set $g(x, \xi, \eta) := \max\{0, \eta - G(x, \xi)\}$ and obtain from (3.8) and (1.2) for all $x \in D$, $\eta \in I$ and random vectors $\tilde{\xi}$:

$$\begin{aligned} |F_{G(x, \xi)}^{(k)}(\eta) - F_{G(x, \tilde{\xi})}^{(k)}(\eta)| &\leq \frac{1}{(k-1)!} \left| \|g(x, \xi, \eta)\|_{k-1}^{k-1} - \|g(x, \tilde{\xi}, \eta)\|_{k-1}^{k-1} \right| \\ &\leq \frac{1}{(k-1)!} \mathbb{E} |g(x, \xi, \eta)^{k-1} - g(x, \tilde{\xi}, \eta)^{k-1}| \\ &\leq \frac{K_I}{(k-2)!} \mathbb{E} [\max\{1, |G(x, \xi)|, |G(x, \tilde{\xi})|\}^{k-2} |G(x, \xi) - G(x, \tilde{\xi})|] \\ &\leq \frac{L_G K_I}{(k-2)!} \mathbb{E} [(K(D) \max\{1, \|\xi\|, \|\tilde{\xi}\|\})^{k-2} \|\xi - \tilde{\xi}\|] \end{aligned}$$

Next we use Hölder's inequality (with $p = k-1$ and $q = \frac{k-1}{k-2}$) to obtain

$$\begin{aligned} \mathbb{E} [\max\{1, \|\xi\|, \|\tilde{\xi}\|\}^{k-2} \|\xi - \tilde{\xi}\|] &\leq \left(\mathbb{E} [\max\{1, \|\xi\|, \|\tilde{\xi}\|\}^{k-1}] \right)^{\frac{k-2}{k-1}} \left(\mathbb{E} [\|\xi - \tilde{\xi}\|^{k-1}] \right)^{\frac{1}{k-1}} \\ &= \|\max\{1, \|\xi\|, \|\tilde{\xi}\|\}\|_{k-1}^{k-2} \|\xi - \tilde{\xi}\|_{k-1} \\ &\leq (1 + \|\xi\|_{k-1} + \|\tilde{\xi}\|_{k-1})^{k-2} \|\xi - \tilde{\xi}\|_{k-1}. \end{aligned}$$

Altogether, we arrive at the estimate

$$\left| F_{G(x, \xi)}^{(k)}(\eta) - F_{G(x, \tilde{\xi})}^{(k)}(\eta) \right| \leq \frac{L_G K_I (K(D))^{k-2}}{(k-2)!} (1 + \|\xi\|_{k-1} + \|\tilde{\xi}\|_{k-1})^{k-2} \|\xi - \tilde{\xi}\|_{k-1},$$

Let $L(k)$ denote the leading constant at the right-hand side of the final estimate. Then the Pompeiu-Hausdorff distance allows the estimate

$$d_H(\mathcal{X}(\xi, Y), \mathcal{X}(\tilde{\xi}, \tilde{Y})) \leq \hat{L} (1 + L(k)) (1 + \|\xi\|_{k-1} + \|\tilde{\xi}\|_{k-1})^{k-2} d_k((\xi, Y), (\tilde{\xi}, \tilde{Y}))$$

if the pair $(\tilde{\xi}, \tilde{Y}) \in \mathcal{L}_{k-1} \times \mathcal{L}_{k-1}$ satisfies the inequality

$$(3.9) \quad (1 + L(k))(1 + \|\xi\|_{k-1} + \|\tilde{\xi}\|_{k-1})^{k-2} d_k((\xi, Y), (\tilde{\xi}, \tilde{Y})) < \hat{\delta}.$$

Finally, we select $\delta > 0$ such that the condition $d_k((\xi, Y), (\tilde{\xi}, \tilde{Y})) < \delta$ implies that the estimate (3.9) is valid. By setting

$$L := \hat{L}(1 + L(k))(1 + \|\xi\|_{k-1} + \delta)^{k-2}$$

we arrive at the desired estimate (3.5). \square

It is worth noting that for large k the Lipschitz modulus $L(k)$ of the ℓ_{k-1} part of d_k in (3.5) (and later also in (3.12)) gets smaller if $\|\xi\|_{k-1}$ grows at most exponentially with k . Hence, higher order stochastic dominance constraints may have improved stability properties.

If D is compact, the solution set $S(\xi, Y)$ of (3.1) is nonempty for all pairs $(\xi, Y) \in \mathcal{L}_{k-1}^2$. In order to derive quantitative continuity property of the solution set mapping $(\xi, Y) \rightarrow S(\xi, Y)$, a growth condition of the objective function on a neighborhood of the original solution set is needed. To this end, we consider the (nondecreasing) growth function $\psi_{(\xi, Y)}$ of the objective of (3.3)

$$(3.10) \quad \psi_{(\xi, Y)}(\tau) := \inf\{f(x) - v(\xi, Y) : d(x, S(\xi, Y)) \geq \tau, x \in \mathcal{X}(\xi, Y)\}$$

and the associated function

$$(3.11) \quad \Psi_{(\xi, Y)}(\theta) := \theta + \psi_{(\xi, Y)}^{-1}(2\theta) \quad (\theta \in \mathbb{R}_+),$$

where we set $\psi_{(\xi, Y)}^{-1}(t) = \sup\{\tau \in \mathbb{R}_+ : \psi_{(\xi, Y)}(\tau) \leq t\}$. Clearly, $\Psi_{(\xi, Y)}$ is increasing and $\Psi_{(\xi, Y)}(0) = 0$ (see also [37, Theorem 7.64]). The following is the main stability result of our paper.

THEOREM 3.3. *Let D be compact and assume that the function G satisfies*

$$|G(x, z) - G(x, \tilde{z})| \leq L_G \|z - \tilde{z}\|$$

for all $x \in D$, $z, \tilde{z} \in \Xi$ and some constant $L_G > 0$. Furthermore, we assume that the k th order uniform dominance condition is satisfied at (ξ, Y) .

Then there exist positive constants L and δ such that

$$(3.12) \quad |v(\xi, Y) - v(\tilde{\xi}, \tilde{Y})| \leq L d_k((\xi, Y), (\tilde{\xi}, \tilde{Y}))$$

$$(3.13) \quad \sup_{x \in S(\tilde{\xi}, \tilde{Y})} d(x, S(\xi, Y)) \leq \Psi_{(\xi, Y)}(L d_k((\xi, Y), (\tilde{\xi}, \tilde{Y})))$$

whenever $d_k((\xi, Y), (\tilde{\xi}, \tilde{Y})) < \delta$. The function $\Psi_{(\xi, Y)}$ is defined in (3.11).

The ideas for proving (3.12) and (3.13) go back to [21] and [37, Chapter 7.J] (see also [38, Theorem 9]). For the convenience of the reader a proof of Theorem 3.3 is provided in the appendix of this paper.

We note that the results may be extended in a straightforward way to the case that a finite number of k th order dominance constraints, i.e.,

$$F_{G_j(x, \xi)}^{(k)}(\eta) \leq F_{Y_j}^{(k)}(\eta) \quad (\forall \eta \in \mathbb{R}, j = 1, \dots, J, J \in \mathbb{N})$$

are present in (3.1). The stability results remain valid if the following extended uniform dominance condition ([6])

$$(3.14) \quad \min_{j=1, \dots, J} \min_{\eta \in I} \left(F_Y^{(k)}(\eta) - F_{G(\bar{x}, \xi)}^{(k)}(\eta) \right) > 0$$

is valid for some $\bar{x} \in D$. Moreover, Theorem 3.3 also extends to the case that the objective f in (3.1) is replaced by an expectation function of the form $\mathbb{E}(g(\cdot, \xi))$ where g is a real-valued function defined on $\mathbb{R}^m \times \mathbb{R}^s$, convex in the first variable and Lipschitz continuous with respect to ξ .

4. Optimal value sensitivity. Next, we study directional differentiability properties of the optimal value of (3.3) with respect to the probability distributions P_ξ and P_Y of ξ and Y , respectively. To this end, we consider the Banach space $\mathcal{Y} = \mathcal{C}(I)$, the closed convex cone

$$K = \{y \in \mathcal{Y} : y(\eta) \geq 0, \forall \eta \in I\} \subset \mathcal{Y},$$

the convex-analysis-indicator functions I_D and I_K of D and K , and the function $\varphi : \mathbb{R}^m \times \mathcal{Y} \times \mathcal{P}(\Xi) \times \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{R}$ given by

$$\varphi(x, y; P_\xi, P_Y) := f(x) + I_D(x) + I_K(\mathcal{G}(x; P_\xi, P_Y) + y),$$

where $\mathcal{G} : \mathbb{R}^m \times \mathcal{P}(\Xi) \times \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{Y}$ is given by

$$(4.1) \quad \begin{aligned} \mathcal{G}(x; P_\xi, P_Y)(\eta) &= F_Y^{(k)}(\eta) - F_{G(x, \xi)}^{(k)}(\eta) \\ &= \int_{\mathbb{R}} \frac{\max\{0, \eta - t\}^{k-1}}{(k-1)!} P_Y(t) - \int_{\Xi} \frac{\max\{0, \eta - G(x, z)\}^{k-1}}{(k-1)!} P_\xi(dz) \end{aligned}$$

for every $x \in \mathbb{R}^m$ and $\eta \in I$. Note that the domain of $\varphi(\cdot, \cdot; P_\xi, P_Y)$ is nonempty and that $\varphi(\cdot, \cdot; P_\xi, P_Y)$ is lower semicontinuous and convex.

We also note that the definition of $\mathcal{G}(x, \cdot, \cdot)$ can be extended from pairs of probability measures to pairs of finite nonnegative measures having finite moments of order $k-1$, and even to pairs of finite signed measures via the Hahn-Jordan decomposition. Hence, we may also extend the domain of $\varphi(x, y; \cdot, \cdot)$ to such pairs of measures for every $(x, y) \in \mathbb{R}^m \times \mathcal{Y}$.

Denote dual pairing between of \mathcal{Y} and \mathcal{Y}^* by $\langle \cdot, \cdot \rangle$. Here \mathcal{Y}^* is isometrically isomorph to the space $\mathbf{rca}(I)$ of regular countably additive measures μ on I having finite total variation $|\mu|(I)$, and the dual pairing is given by

$$\langle \mu, y \rangle = \int_I y(\eta) \mu(d\eta) \quad (\forall y \in \mathcal{Y}, \mu \in \mathbf{rca}(I)).$$

The next result enables us to state our main result.

LEMMA 4.1. *Let $k \geq 2$ and $I = [a, b]$. For each nonnegative $\mu \in \mathbf{rca}(I)$ there exists $u \in \mathcal{U}_{k-1}$ such that the identity*

$$(4.2) \quad \langle \mu, F_X^{(k)} \rangle = \int_I F_X^{(k)}(\eta) \mu(d\eta) = -\mathbb{E}[u(X)]$$

holds for every $X \in \mathcal{L}_{k-1}$. Here, \mathcal{U}_{k-1} denotes the set of all functions $u \in \mathcal{C}^{k-1}(\mathbb{R})$, for which there exists a nonnegative, nonincreasing, left-continuous, bounded function $\varphi : I \rightarrow \mathbb{R}$ such that

$$\begin{aligned} u^{(k-1)}(t) &= (-1)^k \varphi(t) && \text{for almost all } t \in I = [a, b], \\ u^{(k-1)}(t) &= (-1)^k \varphi(a) && \text{for } t < a, \\ u(t) &= 0 && \text{for } t \geq b, \\ u^{(i)}(b) &= 0 && \text{for } i = 1, \dots, k-2, \end{aligned}$$

where the symbol $u^{(i)}$ denotes the i th derivative of u . In particular, the functions $u \in \mathcal{U}_{k-1}$ are nondecreasing and concave on \mathbb{R} .

Proof. Let $\mu \in \mathbf{rca}(I)$ be nonnegative. Then μ is extended to the Borel field of \mathbb{R} by assigning measure 0 to Borel sets not intersecting I . The function $u \in \mathcal{U}_{k-1}$ is then defined as in the proof of [5, Theorem 7.1] by putting

$$\begin{aligned} u^{(k-1)}(t) &= (-1)^k \mu([t, b]) && \text{for almost all } t \leq b, \\ u(t) &= 0, && \text{for } t \geq b, \\ u^{(i)}(b) &= 0, && \text{for } i = 1, \dots, k-2. \end{aligned}$$

As in the proof of [5, Theorem 7.1] one obtains for any $X \in \mathcal{L}_{k-1}$

$$\langle \mu, F_X^{(k)} \rangle = (-1)^k \int_{-\infty}^b F_X^{(k)}(\eta) du^{(k-1)}(t) = - \int_{-\infty}^b u(t) dF_X(t) = -\mathbb{E}[u(X)]$$

via integration by parts $k-1$ times. \square

Define the Lagrange-like function: $\mathfrak{L} : \mathbb{R}^m \times \mathcal{U}_{k-1} \rightarrow \mathbb{R}$ as follows:

$$\mathfrak{L}(x, u; P_\xi, P_Y) := f(x) - \int_{\Xi} u(G(x, z)) P_\zeta(dz) + \int_{\mathbb{R}} u(t) P_Z(dt)$$

We formulate optimality conditions for problem (3.1).

THEOREM 4.2. *Let $k \geq 2$ and $I = [a, b]$. Assume the k th order uniform dominance condition for problem (3.1) at (ξ, Y) . If a feasible point $\hat{x} \in \mathbb{D}$ is an optimal solution of (3.1) then a function $\hat{u} \in \mathcal{U}_{k-1}$ exists so that*

$$(4.3) \quad \mathfrak{L}(\hat{x}, \hat{u}; P_\xi, P_Y) = \min_{x \in D} \mathfrak{L}(x, \hat{u}, P_\xi, P_Y)$$

$$(4.4) \quad \int_{\Xi} \hat{u}(G(\hat{x}, z)) P_\zeta(dz) = \int_{\mathbb{R}} \hat{u}(t) P_Z(dt)$$

If \hat{x} satisfies the dominance constraint and (4.3)-(4.4) for some function $\hat{u} \in \mathcal{U}_{k-1}$, then \hat{x} solves (3.1). Furthermore, the dual problem to (3.1) at (ξ, Y) is

$$(4.5) \quad \max_{u \in \mathcal{U}_{k-1}} \left[\inf_{x \in D} [f(x) + \mathbb{E}(u(G(x; \xi))) - \mathbb{E}(u(Y))] \right]$$

and the duality relation holds.

Proof. We note that problem (3.1) can be cast in the setting of Example 4 on p. 26 of [36]) by considering the dominance constraint as a constraint in the space of continuous functions: $\mathcal{G}(x; P_\xi, P_Y) \in K \subset \mathcal{Y}$. The polar cone of K is

$$K^- = \{\mu \in \mathbf{rca}(I) : \langle \mu, y \rangle \leq 0, \forall y \in K\} = \{\mu \in \mathbf{rca}(I) : \mu \leq 0\}.$$

The Lagrangian Λ associated with problem (3.1) can be formulated as follows:

$$\Lambda(x, \mu; P_\xi, P_Y) = \begin{cases} f(x) + \langle \mu, \mathcal{G}(x; P_\xi, P_Y) \rangle & \text{if } x \in D, \mu \in K^-, \\ -\infty & \text{if } x \in D, \mu \notin K^-, \\ +\infty & \text{if } x \notin D. \end{cases}$$

The optimality conditions for problem (3.1) state that if a feasible point \hat{x} is an optimal solution, then a measure $\hat{\mu} \in K^-$ exists, so that

$$(4.6) \quad \Lambda(\hat{x}, \hat{\mu}; P_\xi, P_Y) = \min_{x \in D} \Lambda(x, \hat{\mu}; P_\xi, P_Y)$$

$$(4.7) \quad \langle \hat{\mu}, \mathcal{G}(\hat{x}; P_\xi, P_Y) \rangle = 0.$$

The dual problem has the following form (cf. [36, (5.13)])

$$(4.8) \quad \max \left\{ \inf_{x \in D} \{f(x) + \langle \mu, \mathcal{G}(x; P_\xi, P_Y) \rangle\} : \mu \in K^- \right\},$$

Using Lemma 4.1, we associate a function $\hat{u} \in \mathcal{U}_{k-1}$ with the measure $\hat{\mu}$ and reformulate the Lagrangian Λ to the following form:

$$\Lambda(x, \hat{\mu}; P_\xi, P_Y) = \mathfrak{L}(x, \hat{u}; P_\xi, P_Y) = f(x) + \int_{\Xi} u(G(x, z)) P_\zeta(dz) - \int_{\mathbb{R}} u(t) P_Z(dt)$$

whenever $x \in D$ and $\mu \in K^-$. The optimality conditions and the dual problem are reformulated using \hat{u} and the new Lagrangian has the desired form. The duality relation holds due to the convexity of the problem. \square

The optimal value of problem (3.3) has the representation

$$(4.9) \quad v(P_\xi, P_Y) = \inf_{x \in \mathbb{R}^m} \varphi(x, 0; P_\xi, P_Y).$$

Next, we derive conditions such that the optimal value function v given by (4.9) is Hadamard directionally differentiable at (P_ξ, P_Y) , where we equip the space $\mathcal{P}(\Xi) \times \mathcal{P}(\mathbb{R})$ with the distance

$$d_k((P_{\bar{\xi}}, P_{\bar{Y}}), (P_\xi, P_Y)) = \mathbb{D}_{k, \infty}(P_\xi, P_{\bar{\xi}}) + \mathbb{D}_{k, \infty}(P_Y, P_{\bar{Y}}).$$

More precisely, we consider the subset of $\mathcal{P}(\Xi) \times \mathcal{P}(\mathbb{R})$, where the distance d_k is finite. The following is the main result of this section.

THEOREM 4.3. *Let D be compact and $k \geq 2$. Assume that the k th order uniform dominance condition is satisfied at (ξ, Y) and that the solution of problem (3.1) is unique. Then the optimal value function v has the following directional derivative of Hadamard type at (P_ξ, P_Y) into direction (P_ζ, P_Z)*

$$v'(P_\xi, P_Y; P_\zeta, P_Z) = \lim_{n \rightarrow \infty} \frac{1}{t_n} (v(P_\xi + t_n P_{\zeta_n}, P_Y + t_n P_{Z_n}) - v(P_\xi, P_Y))$$

for all sequences (t_n) and $((P_{\zeta_n}, P_{Z_n}))$ converging to $0+$ and (P_ζ, P_Z) (with respect to d_k), respectively. The Hadamard directional derivative is of the form

$$(4.10) \quad v'(P_\xi, P_Y; P_\zeta, P_Z) = \sup_{u \in \mathcal{U}_{k-1}^*} \left\{ \int_{\mathbb{R}} u(t) P_Z(dt) - \int_{\Xi} u(G(\bar{x}, z)) P_\zeta(dz) \right\},$$

where $\{\bar{x}\} = S(P_\xi, P_Y)$ and \mathcal{U}_{k-1}^* denotes the set of solutions to problem (4.5).

Proof. We consider the following auxiliary function optimal value function

$$(4.11) \quad w(y) := \inf_{x \in \mathbb{R}^m} \varphi(x, y; P_\xi, P_Y)$$

is convex on \mathcal{Y} . As the set D is nonempty and compact, φ is proper and the function w is proper and lower semicontinuous as well. Furthermore, for any $x \in D$ a point \bar{y} exists such that $\mathcal{G}(x; P_\xi, P_Y) = \bar{y} \in \text{int } K$. Thus, $\text{int dom}(w)$ is nonempty. Consequently, w is continuous on $\text{int dom}(w)$. Hence, for any $\bar{y} \in \text{int dom}(w)$, the convex subdifferential $\partial w(\bar{y})$ is nonempty, convex, bounded and weakly* compact in the topological dual \mathcal{Y}^* of \mathcal{Y} . Furthermore, w is Hadamard directionally differentiable at \bar{y} and the following formula holds (see, [36, Theorem 17]):

$$\lim_{\substack{t_n \rightarrow 0+ \\ d_n \rightarrow d}} \frac{1}{t_n} (w(\bar{y} + t_n d_n) - w(\bar{y})) = w'(\bar{y}, d) = \sup_{\mu \in \partial w(\bar{y})} \langle \mu, d \rangle \quad (\forall d \in \mathcal{Y}).$$

Here, $\partial w(\bar{y})$ is the solution set $S^*(P_\xi, P_Y)$ of the dual (4.8).

Now, let (t_n) be a sequence tending to $0+$, and P_ζ, P_Z be contamination measures such that for any random variables ζ and Z having probability distributions P_ζ and P_Z respectively, it holds $\zeta, Z \in \mathcal{L}_{k-1}$. Additionally, we require

$$(4.12) \quad \lim_{n \rightarrow \infty} d_k((P_{\zeta_n}, P_{Z_n}), (P_\zeta, P_Z)) = \lim_{n \rightarrow \infty} \mathbb{D}_{k, \infty}(P_{\zeta_n}, P_\zeta) + \mathbb{D}_{k, \infty}(P_{Z_n}, P_Z) = 0.$$

Using the definition (4.1) of \mathcal{G} we get for any $x \in D$

$$\begin{aligned} \varphi(x, 0, P_\xi + t_n P_{\zeta_n}, P_Y + t_n P_{Z_n}) &= f(x) + I_D(x) + I_K(\mathcal{G}(x; P_\xi + t_n P_{\zeta_n}, P_Y + t_n P_{Z_n})) \\ &= f(x) + I_D(x) + I_K(\mathcal{G}(x; P_\xi, P_Y) + t_n y_n(x)) \\ &= \varphi(x, t_n y_n(x), P_\xi, P_Y) \end{aligned}$$

where $y_n(x) = \mathcal{G}(x; P_{\zeta_n}, P_{Z_n})$ for $x \in D$. Due to (3.13) in Theorem 3.3 there exist \bar{x}_n in $S(P_\xi + t_n P_{\zeta_n}, P_Y + t_n P_{Z_n})$ such that the sequence (\bar{x}_n) converges to \bar{x} . Due to (4.12) the sequence $(y_n(\bar{x}_n))$ converges to $y = \mathcal{G}(\bar{x}; P_\zeta, P_Z) \in \mathcal{Y}$ (see the proof of Proposition 3.2). Hence, we obtain

$$(4.13) \quad \begin{aligned} &\lim_{n \rightarrow \infty} \frac{1}{t_n} (v(P_\xi + t_n P_{\zeta_n}, P_Y + t_n P_{Z_n}) - v(P_\xi, P_Z)) \\ &= \lim_{n \rightarrow \infty} \frac{1}{t_n} (w(t_n y_n(\bar{x}_n)) - w(0)) = \sup_{\mu \in \partial w(0)} \langle \mu, y \rangle \\ &= \sup_{\mu \in S^*(P_\xi, P_Y)} \langle \mu, \mathcal{G}(\bar{x}; P_\zeta, P_Z) \rangle. \end{aligned}$$

Using Lemma 4.1 again we pass to the dual problem (4.5) and transform the right-hand side of (4.13) to obtain

$$\begin{aligned} v'(P_\xi, P_Y; P_\zeta, P_Z) &= \lim_{n \rightarrow \infty} \frac{1}{t_n} (v(P_\xi + t_n P_{\zeta_n}, P_Y + t_n P_{Z_n}) - v(P_\xi, P_Y)) \\ &= \sup_{u \in \mathcal{U}_{k-1}^*} \{ \mathbb{E}[u(Z)] - \mathbb{E}[u(G(\bar{x}, \zeta))] \} \end{aligned}$$

where \mathcal{U}_{k-1}^* denotes the set of solutions to problem (4.5). \square

Theorem 4.3 is very similar to the directional differentiability result [4, Corollary 3.7] for optimal values of optimization problems with first order dominance constraints, but provides Hadamard directional differentiability in case of k th order dominance constraints with $k \geq 2$. We note that Theorem 4.3 extends in a straightforward way to the case of a finite number of k th order dominance constraints if the extended uniform dominance condition (3.14) is valid.

Under the assumptions of Theorem 4.3, the solution set of problem (4.8) is weakly* compact, which implies the existence of a measure $\tilde{\mu}$ for which the supremum at the right-hand side of (4.13) is attained. A function $\tilde{u} \in \mathcal{U}_{k-1}^*$ corresponds to this measure according to Lemma 4.1. We call this function a *shadow utility* function and obtain the following corollary.

COROLLARY 4.4. *Under the assumptions of Theorem 4.3, a shadow utility function \tilde{u} exists such that the Hadamard-directional derivative of the optimal value function v at (P_ξ, P_Y) in direction (P_ζ, P_Z) has the form*

$$v'(P_\xi, P_Y; P_\zeta, P_Z) = \int_{\mathbb{R}} \tilde{u}(t) P_Z(dt) - \int_{\Xi} \tilde{u}(G(\bar{x}, z)) P_\zeta(dz),$$

where $\{\bar{x}\} = S(P_\xi, P_Y)$ and \tilde{u} is a solution of problem (4.5).

5. Empirical approximations of optimization models with k th order dominance constraints. We assume that a sequence (ξ_n, Y_n) of independent and identically distributed random vectors on some probability space with values in $\Xi \times \mathbb{R}$ are given such that $P_{\xi_1} = P_\xi$ and $P_{Y_1} = P_Y$. Let $P_\xi^{(n)}$ and $P_Y^{(n)}$ denote the (random) empirical measures

$$P_\xi^{(n)} = \frac{1}{n} \sum_{i=1}^n \delta_{\xi_i} \quad \text{and} \quad P_Y^{(n)} = \frac{1}{n} \sum_{i=1}^n \delta_{Y_i} \quad (n \in \mathbb{N}),$$

where δ_z denotes the unit mass at z in Ξ or \mathbb{R} . Inserting the empirical measures into (3.3) instead of P_ξ and P_Y , respectively, leads to the empirical approximation of (3.3)

$$(5.1) \quad \begin{aligned} & \min f(x) \\ \text{s.t.} \quad & \sum_{i=1}^n \left[\max\{0, \eta - G(x, \xi_i(\cdot))\} \right]^{k-1} \leq \sum_{i=1}^n \left[\max\{0, \eta - Y_i(\cdot)\} \right]^{k-1}, \forall \eta \in I, \\ & x \in D. \end{aligned}$$

for any $n \in \mathbb{N}$. We note that the dominance constraints in (5.1) for $k = 2$ may be reformulated as in [6, Section 5] or in [22]. Here, we are only interested in the asymptotic behavior of the empirical approximations (5.1) for n tending to ∞ .

We begin with properties of the empirical process

$$(5.2) \quad \begin{aligned} \mathcal{E}_n g &:= \sqrt{n}(P_\xi^{(n)} \times P_Y^{(n)} - P_\xi \times P_Y)g = \frac{1}{\sqrt{n}} \sum_{i=1}^n (\delta_{\xi_i} \times \delta_{Y_i} - P_\xi \times P_Y)g \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(g(\xi_i, Y_i) - \int_{\Xi} \int_{\mathbb{R}} g(z, t) P_\xi(dz) P_Y(dt) \right) \end{aligned}$$

evaluated at g belonging to some class Γ_k of real-valued measurable functions on $\Xi \times \mathbb{R}$. Here, $P_\xi \times P_Y$ denotes the product measure. Boundedness or convergence properties of the empirical process depend on the size of the class Γ measured in terms of certain covering or bracketing numbers in $\mathcal{L}_2(\Xi \times \mathbb{R}, P_\xi \times P_Y)$. To introduce the latter concept, we consider the space \mathcal{L}_2 to be equipped with the usual norm $\|g\|_2 = (P|g|^2)^{\frac{1}{2}}$, where $P = P_\xi \times P_Y$. The *bracketing number* $N_{[\cdot]}(\varepsilon, \Gamma, \mathcal{L}_2(\Xi \times \mathbb{R}, P))$ is the minimal number of *brackets* $[l, u] = \{f \in \mathcal{L}_2(\Xi \times \mathbb{R}, P) : l \leq f \leq u\}$ with $\|l - u\|_2 < \varepsilon$ needed to cover the class Γ . We define the closed convex subset \hat{D} of D by setting

$$\hat{D} = \{x \in D : \mathcal{G}(x, P_\xi, P_Y) \in K\},$$

where the mapping \mathcal{G} is defined in (4.1).

Under the assumptions of Theorem 4.3, a shadow utility $\hat{u} \in \mathcal{U}_{k-1}$ exists due to Corollary 4.4. It holds

$$v(P_\xi, P_Y) = \inf_{x \in \hat{D}} \mathcal{L}(x, \hat{u}; P_\xi, P_Y).$$

We introduce a function class Γ_k for optimization models with k -th order dominance constraints as follows:

$$\Gamma_k = \left\{ g_x : g_x(z, t) = f(x) + \hat{u}(G(x, z)) - \hat{u}(t), (z, t) \in \Xi \times \Upsilon, x \in \hat{D} \right\},$$

where $\hat{u} \in \mathcal{U}_{k-1}^*$ is any shadow utility for this problem.

PROPOSITION 5.1. *Assume that the function $G(\cdot, z)$ is Lipschitz continuous with a uniform modulus L_G (not depending on $z \in \Xi$) and the assumptions of Theorem 4.3. Let D and the supports $\Xi = \text{supp}(P_\xi)$ and $\Upsilon = \text{supp}(P_Y)$ of P_ξ and P_Y be compact. Then the class Γ_k is a Donsker class, i.e., the empirical process $\{\mathcal{E}_n g : g \in \Gamma_k\}$ given by (5.2) converges in distribution to a Gaussian limit process \mathbb{G} (with zero mean) in the space $\ell^\infty(\Gamma_k)$ (of bounded functions on Γ_k) equipped with supremum norm.*

Proof. All functions g_x are real-valued and bounded for every $x \in \hat{D}$ due to the compactness of the set $\Xi \times \Upsilon$ and the continuity of the functions involved. The function f is convex and, therefore, it is Lipschitz continuous over the compact set \hat{D} with Lipschitz constant L_f . Owing to the continuity of the function $G(x, \cdot)$ and the compactness of Ξ , the image $G(\hat{D}, \Xi)$ is contained in a compact interval $[a, b]$. The function \hat{u} is concave and, hence, Lipschitz continuous on any compact interval $[a, b] \subset \mathbb{R}$. Denoting its Lipschitz modulus by L_u , we obtain the following sequence of inequalities.

$$\begin{aligned} |g_x(z, t) - g_{\tilde{x}}(z, t)| &\leq |f(x) + \hat{u}(G(x, z)) - \hat{u}(t) - f(\tilde{x}) - \hat{u}(G(\tilde{x}, z)) + \hat{u}(t)| \\ &\leq |f(x) - f(\tilde{x})| + |\hat{u}(G(x, z)) - \hat{u}(G(\tilde{x}, z))| \\ &\leq (L_f + L_u L_G) \|x - \tilde{x}\| \end{aligned}$$

Using [44, Example 19.7], we infer that

$$N_{[]}(\varepsilon, \Gamma_k, \mathcal{L}_2(\Xi \times \Upsilon, P)) \leq \bar{C} \varepsilon^{-m}$$

holds for some constant $\bar{C} > 0$ (depending on the diameter of D). Hence, the following bound holds for the bracketing integral:

$$\begin{aligned} J_{[]} (1, \Gamma_k, \mathcal{L}_2(\Xi \times \Upsilon, P)) &= \int_0^1 \sqrt{\log N_{[]}(\varepsilon, \Gamma_k, \mathcal{L}_2(\Xi \times \Upsilon, P))} d\varepsilon \\ &\leq \int_0^1 \sqrt{\log \bar{C} \varepsilon^{-m}} d\varepsilon \leq \sqrt{\log \bar{C}} + \sqrt{m} \int_0^1 \varepsilon^{-\frac{1}{2}} d\varepsilon. \end{aligned}$$

We conclude that the bracketing integral is finite and the result follows from [44, Theorem 19.5]. \square

The next result on the functional delta method is Theorem 7.59 in [42]. For further background on the delta method we refer to [44, Section 20] and to [39].

PROPOSITION 5.2. *Let B_1 and B_2 be Banach spaces equipped with their Borel σ -fields and B_1 be separable. Let (X_n) be random elements of B_1 , $h : B_1 \rightarrow B_2$ be a mapping and (τ_n) be a sequence of positive numbers tending to infinity as $n \rightarrow \infty$. If for some $\theta \in B_1$ the sequence $(\tau_n(X_n - \theta))$ converges in distribution to some random element X of B_1 and h is Hadamard directionally differentiable at μ , it holds*

$$\tau_n(h(X_n) - h(\theta)) \xrightarrow{d} h'(\theta; X),$$

where \xrightarrow{d} denotes convergence in distribution.

Now, we are ready to prove a limit theorem for the optimal values $v(P_\xi^{(n)}, P_Y^{(n)})$ of the empirical approximations (5.1) to (3.3).

THEOREM 5.3. *Under the assumptions of Theorem 4.3 and Proposition 5.1, the optimal values $v(P_\xi^{(n)}, P_Y^{(n)})$, $n \in \mathbb{N}$, satisfy the limit theorem*

$$\sqrt{n}(v(P_\xi^{(n)}, P_Y^{(n)}) - v(P_\xi, P_Y)) \xrightarrow{d} \mathbb{E}[\hat{u}(Z)] - \mathbb{E}[\hat{u}(G(\bar{x}, \zeta))]$$

where $\{\bar{x}\} = S(P_\xi, P_Y)$, $\hat{u} \in \mathcal{U}_{k-1}^*$ as defined in Theorem 4.2 and (ζ, Z) are normally distributed with zero mean.

Proof. We consider the separable Banach space $B_1 = \mathcal{C}(D)$ of real-valued continuous functions on D , $B_2 = \mathbb{R}$ and the infimal mapping $h : \mathcal{C}(D) \rightarrow \mathbb{R}$ defined by $h(g) = \inf_{x \in D} g(x)$. As h is concave and Lipschitz continuous (with modulus 1), it is Hadamard directionally differentiable. In particular, we consider the continuous function g ,

$$g(x) = (P_\xi \times P_Y)g_x = \mathfrak{L}(x, \hat{u}; P_\xi, P_Y)$$

and its infimum

$$v(P_\xi, P_Y) = \inf_{x \in \widehat{D}} \mathfrak{L}(x, \hat{u}; P_\xi, P_Y).$$

The Hadamard directional derivative

$$h'(g; d) = \min\{d(y); y \in \arg \min_{x \in \widehat{D}} g(x)\}$$

of h evaluated at g into direction d , $d(x) = (P_\zeta \times P_Z)g_x$ is just $v'(P_\xi, P_Y; P_\zeta, P_Z)$ and is evaluated in Theorem 4.3.

The limit theorem follows from Proposition 5.1, where the random variable X_n (with values in $\mathcal{C}(\widehat{D})$) is given by $X_n(x) = (P_\xi^{(n)} \times P_Y^{(n)})g_x$, θ is $\theta(x) = (P_\xi \times P_Y)g_x$ and $\tau_n = \sqrt{n}$. The role of X in Proposition 5.2 is played by the Gaussian limit process \mathbb{G} in Proposition 5.1. \square

The latter result allows to apply resampling techniques, in particular, subsampling (cf. [29]) to determine asymptotic confidence intervals for the optimal value $v(P_\xi, P_Y)$ in case of k th order dominance constraints (as in [11, Section 5] for the optimal value of mixed-integer two-stage stochastic programs).

Appendix. *Proof.* (Theorem 3.3)

Let the pair $(\tilde{\xi}, \tilde{Y}) \in \mathcal{L}_{k-1}^2$ be such that

$$\hat{\delta} := d_k((\xi, Y), (\tilde{\xi}, \tilde{Y})) < \delta,$$

where $\delta > 0$ is the corresponding constant from Proposition 3.2. Now, let $x \in S(\xi, Y)$ and $\tilde{x} \in S(\tilde{\xi}, \tilde{Y})$. Then there exists $\hat{x} \in S(\xi, Y)$ such that

$$\|\hat{x} - \tilde{x}\| \leq L_H \hat{\delta},$$

where L_H is the Lipschitz constant from Proposition 3.2. We obtain

$$\begin{aligned} v(\xi, Y) - v(\tilde{\xi}, \tilde{Y}) &= f(x) - f(\tilde{x}) \\ &\leq f(x) - f(\hat{x}) + f(\hat{x}) - f(\tilde{x}) \leq f(\hat{x}) - f(\tilde{x}) \\ &\leq L_f \|\hat{x} - \tilde{x}\| \leq L_f L_H \hat{\delta}, \end{aligned}$$

where L_f is the Lipschitz modulus of the function f on the compact set D . Analogously, we obtain the same estimate for $v(\tilde{\xi}, \tilde{Y}) - v(\xi, Y)$. Hence, the estimate (3.12) is valid with $L := L_f L_H$.

To derive the estimate (3.13), let the pair $(\tilde{\xi}, \tilde{Y}) \in \mathcal{L}_{k-1}^2$ be selected as above and let $\tilde{x} \in S(\tilde{\xi}, \tilde{Y})$. Then there exists $x \in \mathcal{X}(\xi, Y)$ such that $\|\tilde{x} - x\| \leq L_H \hat{\delta}$. According to the definition of the growth function $\psi_{(\xi, Y)}$ we have

$$f(x) - v(\xi, Y) \geq \psi_{(\xi, Y)}(d(x, S(\xi, Y))).$$

Furthermore, we obtain the following chain of estimates

$$\begin{aligned} 2L\hat{\delta} &\geq L_f\|\tilde{x} - x\| + L\hat{\delta} \\ &\geq f(x) - f(\tilde{x}) + v(\tilde{\xi}, \tilde{Y}) - v(\xi, Y) = f(x) - v(\xi, Y) \\ &\geq \psi_{(\xi, Y)}(d(x, S(\xi, Y))), \end{aligned}$$

where \mathbb{B} denotes the unit ball in \mathbb{R}^m . Finally, we conclude

$$\begin{aligned} d(\tilde{x}, S(\xi, Y)) &\leq L_H\hat{\delta} + d(x, S(\xi, Y)) \\ &\leq L_H\hat{\delta} + \psi_{(\xi, Y)}^{-1}(2L\hat{\delta}) \\ &\leq \Psi_{(\xi, Y)}(\max\{L_H, L\}\hat{\delta}). \end{aligned}$$

This completes the proof. \square

Acknowledgement: This research was partly carried out during the second authors visit to the Stevens Institute of Technology at Hoboken. The work of the first author is supported by the NSF CMMI award 0965702 and the second author is supported by the DFG Research Center MATHEON *Mathematics for Key Technologies* in Berlin (<http://www.matheon.de>).

REFERENCES

- [1] D. Berleant, M. Dancre, J. P. Argaud and G. Sheblé (2005), Electric company portfolio optimization under interval stochastic dominance constraints, *4th International Symposium on Imprecise Probabilities and their Applications*, Pittsburgh.
- [2] D. Blackwell (1953), Equivalent Comparisons of Experiments *Ann. Math. Statist.* 24, (1953), 265-272.
- [3] J. F. Bonnans and A. Shapiro: *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.
- [4] D. Dentcheva, R. Henrion and A. Ruszczyński: Stability and sensitivity of optimization problems with first order stochastic dominance constraints, *SIAM Journal on Optimization* 18 (2007), 322–337.
- [5] D. Dentcheva and A. Ruszczyński: Optimization with stochastic dominance constraints, *SIAM Journal on Optimization* 14 (2003), 548–566.
- [6] D. Dentcheva and A. Ruszczyński: Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints, *Mathematical Programming* 99 (2004), 329–350.
- [7] D. DENTCHEVA AND A. RUSZCZYŃSKI, Portfolio optimization with stochastic dominance constraints. *Journal of Banking and Finance* 30/2 (2006), pp. 433–451.
- [8] D. Dentcheva and A. Ruszczyński: Composite semi-infinite optimization, *Control and Cybernetics* 36 (2007), 633–646.
- [9] R. Davidson and J.-Y. Duclos: Statistical inference for SD and for measurement of poverty and inequality, *Econometrica*, 68 (2000), 1435-1464.
- [10] N. Dunford and J. T. Schwartz: *Linear Operators, Part I: General Theory*, Wiley Classics Library, New York, 1988.
- [11] A. Eichhorn and W. Römisch: Stochastic integer programming: Limit theorems and confidence intervals, *Mathematics of Operations Research* 32 (2007), 118–135.
- [12] P. C. Fishburn: *Decision and Value Theory*, John Wiley & Sons, New York, 1964.
- [13] H. Gerber: *An Introduction to Mathematical Risk Theory*, Huebner Foundation Monograph, 1981.
- [14] R. Gollmer, F. Neise and R. Schultz: Stochastic programs with first-order dominance constraints induced by mixed-integer linear recourse, *SIAM Journal on Optimization*, 19 (2008), 552-571.
- [15] R. Gollmer, U. Gotzes, F. Neise and R. Schultz: Risk modeling via stochastic dominance in power systems with dispersed generation, in *Stochastic Optimization Methods in Finance and Energy* (Eds. M. Bertocchi, G. Consigli and M. A. H. Dempster), Springer, New York, 2011, 253-271.

- [16] M. Gugat: Error bounds for infinite systems of convex inequalities without Slater's condition, *Mathematical Programming* 88 (2000), 255–275.
- [17] G. H. Hardy, J. E. Littlewood and G. Pólya: *Inequalities*, Cambridge University Press, Cambridge, MA, 1934.
- [18] L. Horvath, P. Kokoszka and R. Zitikis: Testing for Stochastic dominance using the weighted McFaden statistic, *Journal of Econometrics* 133 (2006), 191–205.
- [19] Y. Liu and H. Xu: Stability and sensitivity analysis of stochastic programs with second order dominance constraints, manuscript, University of Southampton, 2010.
- [20] V. V. Kalashnikov and S. T. Rachev: *Mathematical Methods for Construction of Queuing Models*, Wadsworth & Brooks, Pacific Grove, 1990.
- [21] D. Klatte: On quantitative stability for non-isolated minima, *Control and Cybernetics* 23 (1994), 183–200.
- [22] J. Luedtke: New formulations for optimization under stochastic dominance constraints, *SIAM Journal on Optimization* 19 (2008), 1433–1450.
- [23] M. O. Lorenz (1905): Methods of measuring the concentration of wealth, *Publications of the American Statistical Association*, 9 (1905), 209–219.
- [24] E. Lehmann: Ordered families of distributions, *Annals of Mathematical Statistics* 26 (1955), 399–419.
- [25] H. B. Mann and D. R. Whitney: On a test of whether one of two random variables is stochastically larger than the other, *Annals of Mathematical Statistics*, 18 (1947), 50–60.
- [26] K. Mosler and M. Scarsini (Eds.): *Stochastic Orders and Decision Under Risk*, Institute of Mathematical Statistics, Hayward, California, 1991.
- [27] A. Müller and D. Stoyan: *Comparison Methods for Stochastic Models and Risks*, Wiley, Chichester, 2002.
- [28] W. Ogryczak and A. Ruszczyński: On consistency of stochastic dominance and mean-semideviation models, *Mathematical Programming* 89 (2001), 217–232.
- [29] D. N. Politis, J. P. Romano and M. Wolf: *Subsampling*, Springer, New York, 1999.
- [30] J. Quiggin: *Generalized Expected Utility Theory - The Rank-Dependent Expected Utility Model*. Kluwer, Dordrecht, 1993.
- [31] S. T. Rachev: *Probability Metrics and the Stability of Stochastic Models*, Wiley, 1991.
- [32] S. T. Rachev and W. Römisch: Quantitative stability in stochastic programming: The method of probability metrics, *Mathematics of Operations Research* 27 (2002), 792–818.
- [33] S. T. Rachev and L. Rüschendorf: *Mass Transportation Problems I: Theory*, Springer, Berlin 1998.
- [34] S. T. Rachev, S. V. Stoyanov and F. J. Fabozzi: *Advanced Stochastic Models, Risk Assessment, and Portfolio Optimization*, Wiley, Hoboken, 2008.
- [35] S. T. Rachev, S. V. Stoyanov and F. J. Fabozzi: *A Probability Metrics Approach to Financial Risk Measures*, Wiley-Blackwell, 2011.
- [36] R. T. Rockafellar: *Conjugate Duality and Optimization*, CBMS-NSF Regional Conference Series in Applied Mathematics 16 SIAM, Philadelphia, 1974.
- [37] R. T. Rockafellar and R. J-B. Wets: *Variational Analysis*, Springer, Berlin, 1998.
- [38] W. Römisch: Stability of Stochastic Programming Problems, in: *Stochastic Programming* (A. Ruszczyński and A. Shapiro eds.), Handbooks in Operations Research and Management Science, Volume 10, Elsevier, Amsterdam 2003, 483–554.
- [39] W. Römisch: Delta method, infinite dimensional, in: *Encyclopedia of Statistical Sciences* (S. Kotz, C.B. Read, N. Balakrishnan, B. Vidakovic eds.), Second Edition, Wiley 2005.
- [40] M. Shaked and J. G. Shanthikumar, *Stochastic Orders and Their Applications*, Academic Press, Boston 1994.
- [41] A. Shapiro: On concepts of directional differentiability, *Journal of Optimization Theory and Applications* 66 (1990), 477–487.
- [42] A. Shapiro, D. Dentcheva and A. Ruszczyński: *Lectures on Stochastic Programming*, MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2009.
- [43] A. F. Shorrocks: Ranking income distributions, *Economica, New Series*, 50 (1983), 3–17.
- [44] A. W. van der Vaart: *Asymptotic Statistics*, Cambridge University Press, Cambridge 1998.
- [45] S. S. Wang, V. R. Yong and H.H. Panjer: Axiomatic characterization of insurance prices, *Insurance Mathematics and Economics* 21 (1997), 173–183.
- [46] S. S. Wang and V. R. Yong: Ordering risks: expected utility versus Yaari's dual theory of risk, *Insurance Mathematics and Economics* 22 (1998), 145–161.
- [47] V. M. Zolotarev: Probability metrics, *Theory of Probability and its Applications* 28 (1983), 278–302.
- [48] V. M. Zolotarev: *Contemporary Theory of Summation of Independent Random Variables* (in Russian), Nauka, Moscow, 1986.