SIAM J. OPTIM.
Vol. 23, No. 3, pp. 1672–1688

STABILITY AND SENSITIVITY OF STOCHASTIC DOMINANCE CONSTRAINED OPTIMIZATION MODELS*

DARINKA DENTCHEVA† AND WERNER RÖMISCH‡

Abstract. We consider convex optimization problems with kth 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 of the optimization problems in terms of a suitably selected probability metrics. Moreover, we provide conditions ensuring Hadamard directional differentiablity of the optimal value function. We introduce the notion of a shadow utility, which determines the changes of the optimal value when the underlying random variables are perturbed. Finally, we derive a limit theorem for the optimal values of empirical (Monte Carlo, sample average) approximations of dominance constrained optimization models.

Key words. stochastic order, risk, higher order stochastic dominance, shadow utility, empirical approximation

AMS subject classifications. Primary, 90C31; Secondary, 91B70, 90C15

DOI. 10.1137/120886790

1. Introduction.

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

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

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

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

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, the latter playing 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 because 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 quantitative stability results for the optimal value

*Received by the editors August 2, 2012; accepted for publication (in revised form) May 20, 2013; published electronically August 22, 2013.
http://www.siam.org/journals/siopt/23-3/88679.html

†Department of Mathematical Sciences, Stevens Institute of Technology, Hoboken, NJ 07030 (darinka.dentcheva@stevens.edu). This author was supported by NSF CMMI award 0965702.

‡Department of Mathematics, Humboldt University Berlin, 10099 Berlin, Germany (romisch@math.hu-berlin.de). This author was supported by DFG Research Center MATHEON, Berlin (http://www.matheon.de).

and the optimal solutions. Furthermore, we analyze the sensitivity of the optimal value and prove 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

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

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.

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 [21, 2]. Numerous studies on statistical inference for stochastic dominance are available; we refer to [4, 18] and their references. The stochastic dominance relation of order two 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 [6] and continued in several papers, e.g., [7, 9]. An optimization model with stochastic dominance constraints has been applied to financial optimization in [8] and to electricity market models in [1, 14, 15]. In [14, 15], 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 [5] for optimization problems with first order stochastic dominance constraints. The recent paper [22] 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 [22], we show stability with respect to distances of probability distributions. Which are associated with the models under consideration in a natural way. As in [5], 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 us 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 kth 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 kth 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 are 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 is discussed in section 5.

2. Stochastic dominance constraints and related 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.,

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

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

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

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

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

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 those functions are finite due to the representation

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

Recall that the norm $\| \cdot \|_k$ in $\mathcal{L}_k$ is defined by

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

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 [6]).

In this paper we consider distances of random variables having the property that the distance of two random variables is equal to zero exactly when their probability distributions coincide. Hence, such distances may equivalently be defined on spaces of probability measures rather than on spaces of random variables.

In section 3, we use distances of real random variables that are closely related to kth degree stochastic dominance constraints. They are called *Rachev metrics* (see [34, section 4.4]) and are defined by

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

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

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

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 $g : \mathbb{R} \to \mathbb{R}$ that have measurable kth order derivatives $g^{(k)}$ on $\mathbb{R}$ such that

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

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

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

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 $p \in [1, \infty]$:

$$\zeta_{k,p}(X,Y) \le \zeta_{1,p}(X,Y), \tag{2.7}$$

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

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

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 [19, section 3.10], and (2.9) is proved 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 metrics*, respectively.

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

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

$$\left| \mathbb{E}(X^k) - \mathbb{E}(Y^k) \right| \le k! \, \zeta_{k,1}(X,Y) \tag{2.11}$$

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 kth moments converge due to (2.11).

Finally, we mention that so-called quasi-semidistances are developed in Chapter 8 of the recent monograph [35] which are relevant to the kth order stochastic dominance relation as well. 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)

$$\min\left\{f(x) : x \in D,\, F^{(k)}_{G(x,\xi)}(\eta) \le F^{(k)}_Y(\eta)\,\forall \eta \in \mathbb{R}\right\} \tag{3.1}$$

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*

$$\mathcal{X}(\xi, Y) = \left\{x \in D :\, F^{(k)}_{G(x,\xi)}(\eta) \le F^{(k)}_Y(\eta)\,\forall \eta \in \mathbb{R}\right\} \tag{3.2}$$

is closed and convex.

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

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

where $Z_x = \max\{0, \eta - G(x,\xi)\}$. For a convergent sequence $x^n \in \mathcal{X}(\xi, Y)$ and $\lim_{n\to\infty} x^n = \bar{x}$, 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 the $(k-1)$th mean, we obtain that

$$\lim_{n\to\infty} F^{(k)}_{G(x^n,\xi)}(\eta) = F^{(k)}_{G(\bar{x},\xi)}(\eta) \quad \forall \eta \in \mathbb{R}.$$

Thus, $F^{(k)}_{G(\bar{x},\xi)}(\eta) \le F^{(k)}_Y(\eta)$ and $\bar{x} \in \mathcal{X}(\xi, Y)$. □

Following [6], we focus on the relaxed problem

$$\min\left\{f(x) : x \in D,\, F^{(k)}_{G(x,\xi)}(\eta) \le F^{(k)}_Y(\eta)\,\forall \eta \in I\right\}, \tag{3.3}$$

where $\mathbb{R}$ in (3.1) is replaced by a compact interval I. We denote the feasible set of (3.3) again by $\mathcal{X}(\xi, Y)$. Clearly, Proposition 3.1 remains valid for the relaxed problem. The optimal value of (3.3) is denoted by $v(\xi, Y)$ and its solution set by $S(\xi, Y)$. Note that all these quantities depend only on the probability distributions P_ξ and P_Y of ξ and Y, respectively. 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, where we focus on sensitivity with respect to probability distributions, we shall use the former notation.

For deriving our stability results in what follows, we utilize the *kth order uniform dominance condition* ($k$udc) at the pair (ξ, Y), which is introduced in [6] as a constraint qualification condition. Problem (3.3) satisfies $k$udc at (ξ, Y) if a point $\bar{x} \in D$ exists such that

$$\min_{\eta \in I}\left(F^{(k)}_Y(\eta) - F^{(k)}_{G(\bar{x},\xi)}(\eta)\right) > 0\,. \tag{3.4}$$

Condition (3.4) is the *Slater condition* for (3.3) (see also [16, 22]).

As shown in [7] problem (3.3) is equivalent to the *split variable* formulation

$$\min\left\{f(x) : x \in D, G(x,\xi) \geq X, F_X^{(k)}(\eta) \leq F_Y^{(k)}(\eta) \forall \eta \in I\right\} \tag{3.5}$$

of (3.3) by introducing a new stochastic variable X and a constraint $G(x,\xi) \geq X$ which holds $\mathbb{P}$-a.s. The formulation (3.5) motivates the need of two different metrics for handling the constraints $G(x,\xi) \geq X$ ($\mathbb{P}$-a.s.) and $F_X^{(k)}(\cdot) \leq F_Y^{(k)}(\cdot)$, respectively. This motivates the distance d_k on $\mathcal{L}_{k-1} \times \mathcal{L}_{k-1}$ given by

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

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 between (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 kth order uniform dominance condition is satisfied at the pair (ξ, Y).

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

$$d_{\mathrm{H}}(\mathcal{X}(\xi,Y), \mathcal{X}(\tilde{\xi},\tilde{Y})) \leq L d_k((\xi,Y),(\tilde{\xi},\tilde{Y})) \tag{3.6}$$

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

$$\operatorname{gph}\mathcal{F} = \left\{(x,r) \in D \times \mathbb{R} : \max_{\eta \in I}\left(F_{G(x,\xi)}^{(k)}(\eta) - F_Y^{(k)}(\eta)\right) \leq r\right\},$$

$$\mathcal{F}(\mathbb{R}^m) = \left[\inf_{x \in D}\ \max_{\eta \in I}\left(F_{G(x,\xi)}^{(k)}(\eta) - F_Y^{(k)}(\eta)\right), +\infty\right).$$

In particular, $\operatorname{gph}\mathcal{F}$ 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

$$d(x, \mathcal{F}^{-1}(r)) \leq (a\|x - \tilde{x}\| + b)d(r, \mathcal{F}(x)) \tag{3.7}$$

holds whenever $|r| \le \hat{\delta}$, $x \in D$ and $\tilde{x} \in \mathcal{F}^{-1}(0) = \mathcal{X}(\xi, Y)$. As D is bounded, setting $\hat{L} := a \operatorname{diam} D + b$ in (3.7) we obtain the estimate

$$d(x, \mathcal{F}^{-1}(r)) \le \hat{L} d(r, \mathcal{F}(x)) = \hat{L} \max \left\{ 0, \max_{\eta \in I} \left(F^{(k)}_{G(x,\xi)}(\eta) - F^{(k)}_Y(\eta) \right) - r \right\} \tag{3.8}$$

whenever $|r| \le \hat{\delta}$ and $x \in D$. Hence, (3.8) implies for any $x \in \mathcal{X}(\tilde{\xi}, \tilde{Y})$

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

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

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

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

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

Hence, we obtain from (3.8) for any $x \in \mathcal{X}(\xi, Y)$

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

In a final step, we derive an estimate for

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

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

$$\begin{aligned} \left| \max\{0, \eta - t\}^{k-1} - \max\{0, \eta - \tilde{t}\}^{k-1} \right| &\le (k-1) \max\{|\eta - t|, |\eta - \tilde{t}\,|\}^{k-2} |t - \tilde{t}\,| \\ &\le K_I (k-1) \max\{1, |t|, |\tilde{t}\,|\}^{k-2} |t - \tilde{t}\,| \end{aligned} \tag{3.9}$$

for some constant $K_I \ge 1$ (depending on I) and all $t, \tilde{t} \in \mathbb{R}$. Clearly, (3.9) also holds for $k = 2$. We set $g(x, \xi, \eta) := \max\{0, \eta - G(x, \xi)\}$ and obtain from (3.9) and (1.2) for all $x \in D$, $\eta \in I$ and random vectors $\tilde{\xi}$:

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

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

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

Altogether, we arrive at the estimate

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

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

$$d_{\mathrm{H}}(\mathcal{X}(\xi,Y),\mathcal{X}(\tilde{\xi},\tilde{Y})) \le \hat{L}(1+L(k))\big(1+\|\xi\|_{k-1}+\|\tilde{\xi}\|_{k-1}\big)^{k-2}d_k((\xi,Y),(\tilde{\xi},\tilde{Y}))$$

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

$$(1+L(k))\big(1+\|\xi\|_{k-1}+\|\tilde{\xi}\|_{k-1}\big)^{k-2}d_k((\xi,Y),(\tilde{\xi},\tilde{Y})) < \hat{\delta}. \tag{3.10}$$

Finally, we select $\delta>0$ such that the condition $d_k((\xi,Y),(\tilde{\xi},\tilde{Y}))<\delta$ implies that the estimate (3.10) is valid. By setting

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

we arrive at the desired estimate (3.6). ∎

It is worth noting that for large k the Lipschitz modulus $L(k)$ of the ℓ_{k-1} part of d_k in (3.6) (and later also in (3.13)) 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 and $k$udc is satisfied, the solution set $S(\xi,Y)$ of (3.3) is nonempty and convex for all pairs $(\xi,Y)\in\mathcal{L}^2_{k-1}$. In order to derive the quantitative continuity property of the solution set mapping $(\xi,Y)\to S(\xi,Y)$, a growth condition of the objective function on a neighborhood of the original solution set is needed. To this end, we consider the growth function $\psi_{(\xi,Y)}$ of the objective of (3.3) near the solution set

$$\psi_{(\xi,Y)}(\tau) := \inf\{f(x)-v(\xi,Y) : d(x,S(\xi,Y))\ge\tau,\, x\in\mathcal{X}(\xi,Y)\} \tag{3.11}$$

and the associated conditioning function

$$\Psi_{(\xi,Y)}(\theta) := \theta+\psi^{-1}_{(\xi,Y)}(2\theta) \quad (\theta\in\mathbb{R}_+), \tag{3.12}$$

where we set $\psi^{-1}_{(\xi,Y)}(t) = \sup\{\tau \in \mathbb{R}_+ : \psi_{(\xi,Y)}(\tau) \le t\}$. Clearly, the growth function $\psi_{(\xi,Y)}$ and its inverse $\psi^{-1}_{(\xi,Y)}$ are nondecreasing and lower semicontinuous, and $\Psi_{(\xi,Y)}$ is increasing and lower semicontinuous and it holds $\Psi_{(\xi,Y)}(0) = 0$ (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*

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

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

Then there exist positive constants L and δ such that

$$|v(\xi, Y) - v(\tilde{\xi}, \tilde{Y})| \le L\, d_k((\xi, Y), (\tilde{\xi}, \tilde{Y})) \tag{3.13}$$

$$\sup_{x \in S(\tilde{\xi}, \tilde{Y})} d(x, S(\xi, Y)) \le \Psi_{(\xi,Y)}(L\, d_k((\xi, Y), (\tilde{\xi}, \tilde{Y}))) \tag{3.14}$$

whenever $d_k((\xi, Y), (\tilde{\xi}, \tilde{Y})) < \delta$. The function $\Psi_{(\xi,Y)}$ is defined in (3.12).

The ideas for proving (3.13) and (3.14) go back to [20] and [37, Chapter 7.J] (see also [32, Theorem 2.3 and 2.4], [38, Theorem 9]). For the convenience of the reader a proof of Theorem 3.3 is provided in the appendix of this paper.

In order to get an impression of the growth function $\psi_{(\xi,Y)}$ and hence of $\Psi_{(\xi,Y)}$, let us consider the linear case in more detail and assume that f and G are affine functions and D is a polyhedral set. If, in addition, the pair (ξ, Y) has a finite number of realizations and $k = 2$, then the optimization model is polyhedral. In that case, the function $\Psi_{(\xi,Y)}$ is linear in a neighborhood of zero. In its entire range, the function $\Psi_{(\xi,Y)}$ for such a polyhedral problem is nonlinear but it has a linear majorant. In that case, the estimate (3.14) means upper Lipschitz continuity of S locally around (ξ, Y). If, however, $G(x, \xi)$ and Y have densities, the constraint $F^{(k)}_{G(x,\xi)}(\eta) \le F^{(k)}_Y$ may induce a nonlinear convex constraint for x and hence $\psi_{(\xi,Y)}$ may have polynomial (e.g., quadratic) growth. In those cases one obtains upper Hölder continuity of S locally around (ξ, Y).

We note that the results of this section extend in a straightforward way to the case of a finite number of kth order dominance constraints, i.e.,

$$F^{(k)}_{G_j(x,\xi)}(\eta) \le F^{(k)}_{Y_j}(\eta) \quad (\forall \eta \in I, j = 1, \dots, J,\ J \in \mathbb{N})$$

in (3.3). Proposition 3.2 and Theorem 3.3 remain valid if the extended uniform dominance condition [7]

$$\min_{j=1,\dots,J} \min_{\eta \in I} \left(F^{(k)}_{Y_j}(\eta) - F^{(k)}_{G_j(\bar{x},\xi)}(\eta) \right) > 0 \tag{3.15}$$

is valid for some $\bar{x} \in D$. Moreover, Theorem 3.3 also extends to the case that the objective f in (3.3) 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). To this end, we consider the Banach space $\mathcal{Y} = \mathcal{C}(I)$, together with its topological dual $\mathcal{Y}^*$, and the closed convex cone

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

The dual $\mathcal{Y}^*$ is isometrically isomorph to the space $\mathbf{rca}(I)$ of regular countably additive measures μ on I having finite total variation $|\mu|(I)$ [10, Chapt. IV, p. 5], and the dual pairing is defined by

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

We consider the mapping $\mathcal{G} : \mathbb{R}^m \times \mathcal{P}(\Xi) \times \mathcal{P}(\mathbb{R}) \to \mathcal{Y}$ given by

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

for every $x \in \mathbb{R}^m$ and $\eta \in I$. We 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.

We define $\mathcal{U}_{k-1}$ to be the set of all functions $u \in \mathcal{C}^{k-2}(\mathbb{R})$ such that its $(k-1)$th derivative exists almost everywhere and for which a nonnegative, nonincreasing, left-continuous, bounded function $\varphi : I \to \mathbb{R}$ exists 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 ith derivative of u. In particular, the functions $u \in \mathcal{U}_{k-1}$ are nondecreasing and concave on $\mathbb{R}$.

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*

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

holds for every $X \in \mathcal{L}_{k-1}$.

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 [6, 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 [6, 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. ∎

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

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

We formulate optimality conditions for problem (3.3).

THEOREM 4.2. *Let $k \geq 2$. Assume the kth order uniform dominance condition for problem* (3.3) *at* (ξ, Y). *If a feasible point $\bar{x}$ is an optimal solution of* (3.3), *then a function $\bar{u} \in \mathcal{U}_{k-1}$ exists so that*

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

$$\int_\Xi \bar{u}(G(\bar{x}, z)) P_\xi(dz) = \int_\mathbb{R} \bar{u}(t) P_Y(dt). \tag{4.4}$$

If $\bar{x} \in D$ satisfies the dominance constraint and (4.3)–(4.4) *hold for some function $\bar{u} \in \mathcal{U}_{k-1}$, then $\bar{x}$ solves* (3.3). *Furthermore, the dual problem to* (3.3) *at* (ξ, Y) *is*

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

and the duality relation holds.

Proof. We note that problem (3.3) 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 (3.3) in such a setting can be formulated as follows:

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

The optimality conditions for problem (3.3) state that if a feasible point $\bar{x}$ is an optimal solution, then a measure $\bar{\mu} \in K^-$ exists, so that

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

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

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

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

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

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

whenever $x \in D$. The optimality conditions and the dual problem are reformulated using $\bar{u}$ and the new Lagrangian has the desired form. The duality relation holds due to the convexity of the problem and the uniform dominance condition. ∎

Let $\mathcal{U}^*_{k-1}$ denote the solution set to (4.5). Any element of $\mathcal{U}^*_{k-1}$ will be called the *shadow utility* function. We define the infimal mapping $v : \mathcal{C}(D) \to \mathbb{R}$ by

$$v(g) = \inf_{x \in D} g(x) \qquad (g \in \mathcal{C}(D)) \tag{4.9}$$

and consider the function $g_{\bar{u}}(x) = \mathfrak{L}(x, \bar{u}; P_\xi, P_Y)$ for some shadow utility $\bar{u} \in \mathcal{U}_{k-1}^*$. Clearly, it holds that $v(g_{\bar{u}}) = \inf_{x\in D} \mathfrak{L}(x, \bar{u}; P_\xi, P_Y) = v(P_\xi, P_Y)$.

THEOREM 4.3. *Let $k \geq 2$ and D be compact. Assume the kth order uniform dominance condition for problem* (3.3) *at* (ξ, Y).

Then the optimal value $v(P_\xi, P_Y)$ is Hadamard directionally differentiable on $\mathcal{C}(D)$ and the directional derivative $v'(P_\xi, P_Y; d)$ is of the form

$$v'(P_\xi, P_Y; d) = \min\left\{d(x) : x \in S(P_\xi, P_Y)\right\} \tag{4.10}$$

for any direction $d \in \mathcal{C}(D)$. In particular, it holds that

$$\begin{aligned} v'(P_\xi, P_Y; \bar{d}) &= \lim_{t\to 0+} t^{-1}(v(P_\xi + tP_\zeta, P_Y + tP_Z) - v(P_\xi, P_Y)) \\ &= \min\left\{\mathbb{E}[\bar{u}(Z)] - \mathbb{E}[\bar{u}(G(x,\zeta))] : x \in S(P_\xi, P_Y)\right\} \end{aligned} \tag{4.11}$$

for the direction $\bar{d} = \mathbb{E}[\bar{u}(Z)] - \mathbb{E}[\bar{u}(G(\cdot,\zeta))]$, any $\bar{u} \in \mathcal{U}_{k-1}^*$, and any pair (ζ, Z) of random variables having moments of order $k-1$.

Proof. The infimal mapping v is concave and Lipschitz continuous (with modulus 1) on $\mathcal{C}(D)$ equipped with $\|\cdot\|_\infty$ and hence Hadamard directionally differentiable at each $g \in \mathcal{C}(D)$ (see [3, Prop. 2.49], [41]). The directional derivative is of the form

$$v'(g; d) = \min\left\{d(x) : x \in \arg\min_{y\in D} g(y)\right\} \tag{4.12}$$

for any direction $d \in \mathcal{C}(D)$ (see, e.g., [42, p. 165]). Setting $g = g_{\bar{u}}$ for some $\bar{u} \in \mathcal{U}_{k-1}^*$ proves (4.10). Now, let (ζ, Z) be a pair of random variables having moments of order $k-1$ and let $\bar{u} \in \mathcal{U}_{k-1}^*$. We consider the specific direction $\bar{d} = \mathbb{E}[\bar{u}(Z)] - \mathbb{E}[\bar{u}(G(\cdot\,,\zeta))]$. Then (4.10) implies

$$v'(P_\xi, P_Y; \bar{d}) = \min\{\mathbb{E}[\bar{u}(Z)] - \mathbb{E}[\bar{u}(G(x,\zeta))] : x \in S(P_\xi, P_Y)\}.$$

The directional derivative $v'(P_\xi, P_Y; \bar{d})$ is just the left-hand side of (4.11). ∎

We note that (4.11) is very similar to the directional differentiability result [5, Corollary 3.7] for optimal values of optimization problems with first order dominance constraints. Note also that Theorem 4.3 extends in a straightforward way to the case of a finite number of kth order dominance constraints if the extended uniform dominance condition (3.15) is valid.

5. Empirical approximations of optimization models with kth 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 $\Xi \times \mathbb{R}$ is 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

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

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)

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

for any $n \in \mathbb{N}$. We note that the dominance constraints in (5.1) for $k = 2$ may be reformulated as in [7, section 5] or in [24]. 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,

$$\mathcal{E}_n g := \sqrt{n}\big(P_\xi^{(n)} \times P_Y^{(n)} - P_\xi \times P_Y\big)g = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(\delta_{\xi_i} \times \delta_{Y_i} - P_\xi \times P_Y\big)g$$

$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(g(\xi_i, Y_i) - \int_\Xi\int_\mathbb{R} g(z,t)P_\xi(dz)P_Y(dt)\Big), \tag{5.2}$$

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, denote $P = P_\xi \times P_Y$. The *bracketing number* $N_{[\,]}(\varepsilon, \Gamma, \mathcal{L}_2(\Xi \times \mathbb{R}, P))$ is the minimal number of *brackets* $[l, u] = \{f \in \mathcal{L}_2(\Xi \times \mathbb{R}, P) : l \le f \le u\}$ with $\|l - u\|_2 < \varepsilon$ needed to cover the class Γ.

Under the assumptions of Theorem 4.2, a shadow utility $\hat{u} \in \mathcal{U}_{k-1}^*$ exists and it holds that

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

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

$$\Gamma_k = \Big\{g_x : g_x(z,t) = f(x) + \hat{u}(G(x,z)) - \hat{u}(t)\big),\ (z,t) \in \Xi \times \Upsilon,\ x \in D\Big\}. \tag{5.3}$$

Proposition 5.1. *Let $k \ge 2$. Assume the kth order uniform dominance condition for problem* (3.3) *at (ξ, Y) and assume that the function $G(\cdot, z)$ is Lipschitz continuous with a uniform modulus L_G (not depending on $z \in \Xi$). Let D and the supports $\Xi = \operatorname{supp}(P_\xi)$ and $\Upsilon = \operatorname{supp}(P_Y)$ 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 tight limit process $\mathbb{G}$ in the space $\ell^\infty(\Gamma_k)$ (of bounded functions on Γ_k) equipped with the supremum norm. The process $\mathbb{G}$ is Gaussian with zero mean and covariances $\mathbb{E}[\mathbb{G}g\,\mathbb{G}\tilde{g}] = \mathbb{E}_P[g\tilde{g}] - \mathbb{E}_P[g]\mathbb{E}_P[\tilde{g}]$ for $g, \tilde{g} \in \Gamma_k$.*

Proof. All functions g_x are real-valued and bounded for every $x \in D$ 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 D with Lipschitz constant L_f. Owing to the continuity of the function $G(x, \cdot)$ and the compactness of Ξ, the image $G(D, \Xi)$ 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:

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

Using [44, Example 19.7], we infer that

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

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

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

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

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, $h : B_1 \to B_2$ be a mapping and (τ_n) be a sequence of positive numbers tending to infinity as $n \to \infty$. 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 θ, it holds that*

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

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 Proposition* 5.1*, the optimal values $v(P_\xi^{(n)}, P_Y^{(n)})$, $n \in \mathbb{N}$, satisfy the limit theorem*

$$\sqrt{n}\big(v(P_\xi^{(n)}, P_Y^{(n)}) - v(P_\xi, P_Y)\big) \xrightarrow{d} \min\{\mathbb{G}(x) : x \in S(P_\xi, P_Y)\},$$

where $\mathbb{G}$ is a Gaussian process with zero mean and covariances $\mathbb{E}[\mathbb{G}(x)\,\mathbb{G}(\tilde{x})] = \mathbb{E}_P[g_x g_{\tilde{x}}] - \mathbb{E}_P[g_x]\mathbb{E}_P[g_{\tilde{x}}]$ for $x, \tilde{x} \in S(P_\xi, P_Y)$.

If $S(P_\xi, P_Y)$ is a singleton, i.e., $S(P_\xi, P_Y) = \{\bar{x}\}$, the limit $\mathbb{G}(\bar{x})$ is normal with zero mean and variance $\mathbb{E}_P[g_{\bar{x}}^2] - (\mathbb{E}_P[g_{\bar{x}}])^2$.

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 $v : \mathcal{C}(D) \to \mathbb{R}$ defined by (4.9). According to the proof of Theorem 4.3 v is Hadamard directionally differentiable and its directional derivative is given by (4.12). The limit theorem follows from Propositions 5.1 and 5.2, where the random variable X_n (with values in $\mathcal{C}(D)$) 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 $\tau_n = \sqrt{n}$ for any $x \in D$, $n \in \mathbb{N}$ and g_x defined in (5.3). The role of X in Proposition 5.2 is played by the Gaussian limit process $\mathbb{G}$ in Proposition 5.1. ∎

The latter result allows us to apply resampling techniques to determine asymptotic confidence intervals for the optimal value $v(P_\xi, P_Y)$ of (3.3). As discussed in [11, section 5], the classical bootstrapping method does not work if the directional derivative is not linear in the direction, i.e., if the solution to (3.3) is not unique. The subsampling method (cf. [29]), however, applies to this situation, too. We refer to [11, section 5] for further details of applying subsampling to determine asymptotic confidence intervals for the optimal value of mixed-integer two-stage stochastic programs.

Appendix. *Proof of Theorem* 3.3. Let the pair $(\tilde{\xi}, \tilde{Y}) \in \mathcal{L}^2_{k-1}$ 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 $\hat{x} \in \mathcal{X}(\xi, Y)$ such that

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

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

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

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 (3.13) is valid with $L := L_f L_H$.

To derive the estimate (3.14), let the pair $(\tilde{\xi}, \tilde{Y}) \in \mathcal{L}^2_{k-1}$ 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 $\|\tilde{x} - x\| \le L_H \hat{\delta}$. According to the definition of the growth function $\psi_{(\xi,Y)}$ we have

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

Furthermore, we obtain the following chain of estimates:

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

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

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

This completes the proof. □

Acknowledgment. This research was partly carried out during the second author's visit to the Stevens Institute of Technology at Hoboken, NJ.

REFERENCES

[1] D. Berleant, M. Dancre, J. P. Argaud, and G. Sheblé, *Electric company portfolio optimization under interval stochastic dominance constraints*, 4th International Symposium on Imprecise Probabilities and Their Applications, Pittsburgh, 2005.

[2] D. Blackwell, *Equivalent comparisons of experiments*, Ann. Math. Statist., 24 (1953), pp. 265–272.

[3] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[4] R. Davidson and J.-Y. Duclos, *Statistical inference for SD and for measurement of poverty and inequality*, Econometrica, 68 (2000), pp. 1435–1464.

[5] D. Dentcheva, R. Henrion, and A. Ruszczyński, *Stability and sensitivity of optimization problems with first order stochastic dominance constraints*, SIAM J. Optim., 18 (2007), pp. 322–337.

[6] D. Dentcheva and A. Ruszczyński, *Optimization with stochastic dominance constraints*, SIAM J. Optim., 14 (2003), pp. 548–566.

[7] D. Dentcheva and A. Ruszczyński, *Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints*, Math. Programm., 99 (2004), pp. 329–350.

[8] D. Dentcheva and A. Ruszczyński, *Portfolio optimization with stochastic dominance constraints*, J. Banking Finance, 30 (2006), pp. 433–451.

[9] D. Dentcheva and A. Ruszczyński, *Composite semi-infinite optimization*, Control Cybernet., 36 (2007), pp. 633–646.

[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*, Math. Oper. Res., 32 (2007), pp. 118–135.

[12] P. C. Fishburn, *Decision and Value Theory*, Wiley, New York, 1964.

[13] H. Gerber, *An Introduction to Mathematical Risk Theory*, Huebner Foundation, Philadelphia, 1981.

[14] 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, M. Bertocchi, G. Consigli, and M. A. H. Dempster, eds., Springer, New York, 2011, pp. 253–271.

[15] R. Gollmer, F. Neise, and R. Schultz, *Stochastic programs with first-order dominance constraints induced by mixed-integer linear recourse*, SIAM J. Optim., 19 (2008), pp. 552–571.

[16] M. Gugat, *Error bounds for infinite systems of convex inequalities without Slater's condition*, Math. Program., 88 (2000), pp. 255–275.

[17] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, UK, 1934.

[18] L. Horvath, P. Kokoszka, and R. Zitikis, *Testing for Stochastic dominance using the weighted McFaden statistic*, J. Econometrics, 133 (2006), pp. 191–205.

[19] V. V. Kalashnikov and S. T. Rachev, *Mathematical Methods for Construction of Queuing Models*, Wadsworth & Brooks, Pacific Grove, CA, 1990.

[20] D. Klatte, *On quantitative stability for non-isolated minima*, Control Cybernet., 23 (1994), pp. 183–200.

[21] E. Lehmann, *Ordered families of distributions*, Ann. Math. Statist., 26 (1955), pp. 399–419.

[22] Y. Liu and H. Xu, *Stability analysis of stochastic programs with second order dominance constraints*, Math. Program., Ser. A, DOI: 10.1007/s10107-012-0585-0.

[23] M. O. Lorenz, *Methods of measuring the concentration of wealth*, Publ. Amer. Statist. Assoc., 9 (1905), pp. 209–219.

[24] J. Luedtke, *New formulations for optimization under stochastic dominance constraints*, SIAM J. Optim., 19 (2008), pp. 1433–1450.

[25] H. B. Mann and D. R. Whitney, *On a test of whether one of two random variables is stochastically larger than the other*, Ann. Math. Statist., 18 (1947), pp. 50–60.

[26] K. Mosler and M. Scarsini, eds., *Stochastic Orders and Decision Under Risk*, Institute of Mathematical Statistics, Hayward, CA, 1991.

[27] A. Müller and D. Stoyan, *Comparison Methods for Stochastic Models and Risks*, Wiley, Chichester, UK, 2002.

[28] W. Ogryczak and A. Ruszczyński, *On consistency of stochastic dominance and mean-semideviation models*, Math. Programm., 89 (2001), pp. 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, Netherlands, 1993.

[31] S. T. Rachev, *Probability Metrics and the Stability of Stochastic Models*, Wiley, New York, 1991.

[32] S. T. Rachev and W. Römisch, *Quantitative stability in stochastic programming: The method of probability metrics*, Math. Oper. Res., 27 (2002), pp. 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, NJ, 2008.

[35] S. T. Rachev, S. V. Stoyanov, and F. J. Fabozzi, *A Probability Metrics Approach to Financial Risk Measures*, Wiley-Blackwell, London, 2011.
[36] R. T. Rockafellar, *Conjugate Duality and Optimization*, CBMS-NSF Regional Conf. Ser. Appl. Math. 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., Handb. Oper. Res. Management Sci. 10, Elsevier, Amsterdam, 2003, pp. 483–554.
[39] W. Römisch, *Delta method, infinite dimensional*, in Encyclopedia of Statistical Sciences, S. Kotz, C. B. Read, N. Balakrishnan, and B. Vidakovic, eds., 2nd ed., Wiley, New York, 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*, J. Optim. Theory Appl., 66 (1990), pp. 477–487.
[42] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming*, MPS/SIAM Ser. Optim., SIAM, Philadelphia, 2009.
[43] A. F. Shorrocks, *Ranking income distributions*, Economica, 50 (1983), pp. 3–17.
[44] A. W. van der Vaart, *Asymptotic Statistics*, Cambridge University Press, Cambridge, UK, 1998.
[45] S. S. Wang, V. R. Yong, and H. H. Panjer, *Axiomatic characterization of insurance prices*, Insurance Math. Econom., 21 (1997), pp. 173–183.
[46] S. S. Wang and V. R. Yong, *Ordering risks: Expected utility versus Yaari's dual theory of risk*, Insurance Math. Econom., 22 (1998), pp. 145–161.
[47] V. M. Zolotarev, *Probability metrics*, Theory Probab. Appl., 28 (1983), pp. 278–302.
[48] V. M. Zolotarev, *Contemporary Theory of Summation of Independent Random Variables* (in Russian), Nauka, Moscow, 1986.