Time inconsistency of optimal policies of distributionally robust inventory models

Alexander Shapiro
School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332
ashapiro@isye.gatech.edu

Linwei Xin
Booth School of Business, University of Chicago, Chicago, Illinois 60637
Linwei.Xin@chicagobooth.edu

In this paper, we investigate optimal policies of distributionally robust (risk averse) inventory models. We demonstrate that if the respective risk measures are not strictly monotone, then there may exist infinitely many optimal policies which are not base-stock and not time consistent. This is in a sharp contrast with the risk neutral formulation of the inventory model where all optimal policies are time consistent. This also extends previous studies of time inconsistency in the robust setting.

Key words: inventory model, base-stock policy, time consistency, distributional robustness, moment constraints, spectral risk measures

---

1. Introduction

Consider the following classical inventory model:

$$\begin{array}{ll}\min\limits_{x_t \geq y_t} & \mathbb{E}\left[\sum_{t=1}^{T} c_t(x_t - y_t) + b_t[D_t - x_t]_+ + h_t[x_t - D_t]_+\right] \\ \text{s.t.} & y_{t+1} = x_t - D_t,\ t = 1, ..., T-1,\end{array} \tag{1}$$

where $x_+ := \max\{x, 0\}$, $D_1, ..., D_T$ is a (random) demand process, c_t, b_t, h_t are the ordering, backorder penalty and holding costs per unit, respectively, y_t is the inventory level and $x_t - y_t$ is the order quantity at time t. The decisions x_t are viewed as control variables and y_t as state variables, where the initial inventory level y_1 is given. We use capital letter for D_t viewed as a random variable and d_t for its particular realization through the paper. Assume that $b_t > c_t \geq 0$, $h_t > 0$, $t = 1, ..., T$.

Assuming further that the demand process is stagewise independent[1], it is well known that the base-stock policy is optimal for Problem (1) (e.g., Zipkin 2000). These base-stock policies are time consistent in the sense that *the optimal policy computed at the initial period of the decision process, before any realization of the demand process became available, remains optimal at the later periods.*

Recently considerable attention was attracted to risk averse and distributionally robust formulations of stochastic programs, e.g., Delage and Ye (2010), Wiesemann, Kuhn and Rustem (2013), Wiesemann, Kuhn and Sim (2014), de Ruiter, Brekelmans and den Hertog (2016), Jiang and Guan (2016), Esfahani and Kuhn (2018), and references therein. In the distributionally robust approach it is argued that the "true" distribution of the data process is never known exactly and this motivates us to consider the worst-distribution approach for a specified family of probability distributions (probability measures). In the risk averse approach the expectation operator is replaced by a risk functional (risk measure) defined on an appropriate space of random variables. In the influential paper by Artzner et al. (1999), it was suggested that a "good" risk measure should satisfy certain natural conditions (axioms) and such risk measures were called coherent. By duality arguments distributionally robust and risk averse approaches, with coherent risk measures, in a sense are equivalent to each other (e.g., Section 6.3 of Shapiro, Dentcheva and Ruszczyński 2014).

In a pioneering paper, Scarf (1958) gave an elegant solution for the worst-distribution formulation in case of the static inventory model with $T=1$, when only first and second order moments of the demand distribution are specified. An extension of such distributionally robust approach to the multi-period setting, when $T>1$, is delicate.

When the employed risk functional has a nested form, an optimal policy is time consistent if and only if (iff) it satisfies the respective dynamic programming equations. Therefore time consistent policies always exist provided the dynamic programming equations have solutions. If an optimal policy is not time consistent, then it is inferior to the time consistent policies in the sense that for some realizations of the random data process it is strictly worse than the corresponding time consistent policy although both policies have the same optimal value from the first stage point of view (e.g., Section 5.2 of Shapiro 2018). In the risk neutral setting this does not happen since the expectation operator is strictly monotone. On the other hand, in distributionally robust multistage settings it could happen that there are many optimal solutions which do not satisfy the dynamic programming equations and are not time consistent.

In the framework of a robust objective, i.e. when using the max-type risk functional, analysis of the inventory model was presented in the pioneering paper by Ben-Tal, Goryashko, Guslitzer and Nemirovski (2004), where it was suggested to use affine policies. It was shown in Bertsimas, Iancu and Parrilo (2010) that in certain settings affine policies (decision rules) are optimal, but do not satisfy the dynamic programming equations and are not time consistent. Such a phenomenon of time inconsistency is claimed to be "by no means an exception, but rather a general fact, intrinsic in any robust multi-stage decision model" (e.g., Delage and Iancu 2015). Time inconsistent policies were also explicitly constructed in an inventory setting in Delage and Iancu (2015). In the robust setting[2] a concept essentially equivalent to the time consistency was referred to as Pareto Efficiency in Iancu and Trichakis (2014).

In this paper, we extend the previous study of time inconsistency to risk averse (distributionally robust) models and further show that time inconsistency is not unique to robust multistage decision making, but may happen for a large class of risk averse/distributionally robust settings. It is somewhat surprising since time inconsistency can happen in seemingly natural formulations already in the two period (i.e., $T = 2$) setting. This could happen when the associated risk functionals are not *strictly* monotone. This was mentioned in Section 6.8.5 of Shapiro, Dentcheva and Ruszczyński (2014) and investigated further in Shapiro (2017). Note that the discussion in Shapiro (2017) is generic, while in this paper we present a detail investigation of this phenomenon in a setting of the inventory model. In addition, in Shapiro (2017) the focus is on the discussion of risk functionals instead of the associated optimization problems. This paper takes a further step and investigates risk averse optimization problems associated with non-strictly monotone risk measures, which are more complex to study. For example, it is possible that such optimization problems still have only time consistent optimal policies (see Corollary 3 below). This paper aims to demonstrate existence of an infinite number of time inconsistent optimal policies in the setting of distributionally robust inventory models defined by the first n moment constraints (Section 3.1) and risk averse inventory models associated with spectral risk measure functionals (Section 3.2), which are widely accepted models of ambiguity. This is in a sharp contrast with the risk neutral formulation of the inventory model where all optimal policies are time consistent. This also extends previous studies of time inconsistency in the robust setting.

This paper focuses on distributionally robust and risk averse models involving nested risk functionals, where time consistent optimal policies always exist. More precisely, in our inventory setting, there always exists an optimal time consistent base-stock policy. In addition, all time consistent policies are of base-stock form. The goal of the paper is to understand conditions when time inconsistent optimal policies may also exist and shed light on what distributionally robust models generate such time inconsistencies. We aim at pointing to possible existence of such inconsistent policies rather than telling to the decision maker how to use this.

We also note that there is no unique way to define a multi-period distributionally robust inventory model. For example, a different version of distributionally robust inventory model, defined by the first and second order moment constraints, was studied in Xin and Goldberg (2013). The key difference from our setting in Section 3.1 is that our risk function $\mathcal{R}$ to be defined in Section 2, is represented in a nested form such that there always exists an optimal base-stock policy that is time consistent. By contrast, the risk functional defined in Xin and Goldberg (2013) may not have a nested form. As a consequence, the problem may have no optimal policy of base-stock form and may have no time consistent optimal policy.

The rest of the paper is organized as follows. In Section 2, we give a general discussion of risk averse (distributionally robust) inventory models. Section 3 is devoted to specific constructions of such time inconsistent optimal policies. We provide concluding remarks in Section 4 and defer all the proofs to the technical appendix in Section 5.

2. Distributionally robust inventory model

In this section, we provide a general discussion of risk averse (distributionally robust) inventory models of the form

$$\begin{aligned} \min_{x_t \geq y_t} \ & \mathcal{R}\left[\sum_{t=1}^T c_t(x_t - y_t) + b_t[D_t - x_t]_+ + h_t[x_t - D_t]_+\right] \\ \text{s.t.} \ & y_{t+1} = x_t - D_t,\ t = 1, ..., T-1. \end{aligned} \tag{2}$$

Here $\mathcal{R}(\cdot)$ is a real valued functional defined on a space of random variables. If $\mathcal{R}$ is the expectation operator, i.e., $\mathcal{R} := \mathbb{E}$, then (2) becomes the risk neutral formulation (1). Optimization in (2) is performed over policies $x_t = x_t(d_{[t-1]})$, $t = 1, ..., T$, which are nonanticipative functions of the demand process[3].

Assuming that the functional $\mathcal{R}$ can be represented in a nested form (see equation (6) below), it is possible to write the following dynamic programming equations for problem (2). At each time period $t = T, ..., 2$, the cost-to-go function $V_t(y_t)$ is given by the optimal value of problem

$$\min_{x_t \geq y_t} \left\{ c_t(x_t - y_t) + \rho_t\left[\psi_t(x_t, D_t) + V_{t+1}(x_t - D_t)\right] \right\}, \tag{3}$$

with $V_{T+1}(\cdot) \equiv 0$ and

$$\psi_t(x_t, d_t) := b_t[d_t - x_t]_+ + h_t[x_t - d_t]_+. \tag{4}$$

For $t = 1$, Problem (3) represents the first period optimization problem and its optimal value coincides with the optimal value of Problem (2). Consider intervals $[\alpha_t, \beta_t] \subset \mathbb{R}_+$, $t = 1, ..., T$, and denote by $\mathfrak{P}_t$ the set of all (Borel) probability measures on $[\alpha_t, \beta_t]$. The functionals ρ_t are assumed to be of the form

$$\rho_t(Z) := \sup_{P_t \in \mathfrak{M}_t} \mathbb{E}_{P_t}[Z],\ t = 1, ..., T, \tag{5}$$

where $\mathfrak{M}_t$ is a nonempty subset of $\mathfrak{P}_t$.

We consider two basic ways for constructing sets $\mathfrak{M}_t$. In one approach we assume existence of a reference probability measure and the set $\mathfrak{M}_t$ consisting of probability measures absolutely continuous with respect to the reference measure. In another approach the sets $\mathfrak{M}_t$ are defined by moment constraints. In that case there is no reference measure. For the rest of the paper, by writing that a property holds for **every realization** of a considered random variable we mean that it holds with probability one, or almost everywhere, with respect to the reference probability measure in the first setting, while in the second setting it supposed to hold for **all** possible realizations. In particular by writing that $Z \geq Z'$, for random variables Z, Z', we mean that this inequality holds with probability one in the first setting, while in the second setting it supposed to hold for all possible realizations.

The risk functional $\mathcal{R}$ of Problem (2), associated with the dynamic equations (5), has the nested form

$$\mathcal{R}(\cdot) = \rho_1 \left(\rho_{2|D_1} \big(\cdots \rho_{T|D_{[T-1]}}(\cdot) \big) \right), \tag{6}$$

where $\mathbb{E}_{P_t|D_{[t-1]}}$ denotes the conditional expectation with respect to the distribution P_t conditional on $D_{[t-1]}$ and

$$\rho_{t|D_{[t-1]}}(\cdot) := \sup_{P_t \in \mathfrak{M}_t} \mathbb{E}_{P_t|D_{[t-1]}}[\,\cdot\,],\ t = 2, ..., T.$$

Such nested risk averse formulations of multistage programs are well discussed in the literature in, e.g., Ruszczyński and Shapiro (2006), Ruszczyński (2010), Shapiro, Dentcheva and Ruszczyński (2014), Iancu, Petrik and Subramanian (2015), and for inventory models in, e.g., Ahmed, Cakmak and Shapiro (2007). In view of the nested form (6) of the considered risk functional, it is natural here to consider conditional optimality criteria of the nested form

$$\mathcal{R}_{t|D_{[t-1]}}(\cdot) := \rho_{t|D_{[t-1]}} \big(\cdots \rho_{T|D_{[T-1]}}(\cdot) \big),\ t = 2, ..., T.$$

This leads to the following notion of time consistency of an optimal policy.

Definition 1. Let $\tilde{\pi} = \{\tilde{x}_1, ..., \tilde{x}_T(d_{[T-1]})\}$ be an optimal solution (optimal policy) of Problem (2). It is said that policy $\tilde{\pi}$ is *time consistent* if for every time period $t = 2, ..., T$, the remaining policy $\{\tilde{x}_t(d_{[t-1]}), ..., \tilde{x}_T(d_{[T-1]})\}$ is optimal for the problem

$$\begin{aligned} \min_{x_\tau \geq y_\tau} \ & \mathcal{R}_{t|D_{[t-1]}} \left[\textstyle\sum_{\tau=t}^T c_\tau(x_\tau - y_\tau) + b_\tau[D_\tau - x_\tau]_+ + h_\tau[x_\tau - D_\tau]_+ \right] \\ \text{s.t.} \ & y_{\tau+1} = x_\tau - D_\tau, \ \tau = t, ..., T-1, \end{aligned}$$

conditional on the observed realization $D_{[t-1]} = d_{[t-1]}$ of the demand process.

Time consistency is naturally related to the respective dynamic programming equations.

Definition 2. It is said that policy $\tilde{\pi} = \{\tilde{x}_1, ..., \tilde{x}_T(d_{[T-1]})\}$ satisfies the dynamic programming equations if for $t = 1, ..., T$,

$$\tilde{x}_t \in \arg\min_{x_t \geq y_t} \left\{ c_t(x_t - y_t) + \rho_t \left[\psi_t(x_t, D_t) + V_{t+1}(x_t - D_t) \right] \right\},$$

conditional on $D_{[t-1]} = d_{[T-1]}$ for every realization of the demand process.

That is, at time $t = 2, ..., T$ we observed the realization of the demand process, up to this time, and made decision $\tilde{x}_{t-1}$ according to the considered policy $\tilde{\pi}$. Time consistency of $\tilde{\pi}$ means that, looking into the future, conditional on the demand realization and our decision, policy $\tilde{\pi}$ remains optimal with respect to the nested tail $\mathcal{R}_{t|D_{[t-1]}}$ of our optimization criterion. Note that this tail optimization criterion is conditional on the observed realization of the demand process. In the risk neutral case when $\mathcal{R} = \mathbb{E}$, the corresponding tail risk measure is given by the respective conditional expectation. We refer to Section 6.8.5 of Shapiro, Dentcheva and Ruszczyński (2014) for a further discussion of these concepts in risk averse settings. A natural question, which we address here, is to understand what risk averse models generate time inconsistency.

2.1. Strict monotonicity

Note that each functional ρ_t, defined in (5), is convex, positively homogeneous, translation equivariant and *monotone*. Monotonicity means that if $Z \geq Z'$, then $\rho_t(Z) \geq \rho_t(Z')$. We will need a stronger notion of monotonicity.

Definition 3. We say that functional ρ_t, of the form (5), is *strictly* monotone if $Z \geq Z'$ and $Z \neq Z'$ imply $\rho_t(Z) > \rho(Z')$.

For convex functionals, necessary and sufficient conditions for strict monotonicity are given in Shapiro (2017) in terms of the respective subdifferentials. In particular, if the set $\mathfrak{M}_t$ is defined by a finite number of moment constraints (in particular if $\mathfrak{M}_t = \mathfrak{P}_t$), and the interval $[\alpha_t, \beta_t]$ is non-degenerate (i.e., $\alpha_t < \beta_t$), then the corresponding functional is not strictly monotone. Also, a popular coherent risk measure, the Average Value-at-Risk ($\mathsf{AV@R}$), is not strictly monotone.

Any policy satisfying the dynamic programming equations is time consistent (in the sense of Definition 1); and conversely if an optimal policy is time consistent, then it satisfies the dynamic programming equations[4]. Assuming that the considered risk functionals are monotone, it is shown in Proposition 6.80 in Shapiro, Dentcheva and Ruszczyński (2014) that if the considered multi-stage problem has a *unique* optimal solution (optimal policy), then this policy is time consistent and satisfies the respective dynamic programming equations. If the employed risk functionals are strictly monotone, then every optimal policy is time consistent and satisfies the respective dynamic programming equations. On the other hand, as it was discussed in Shapiro (2017), if the risk functionals are not *strictly* monotone, then there may exist optimal policies which are not time consistent and do not satisfy the dynamic programming equations. In the robust setting, when $\mathfrak{M}_t = \mathfrak{P}_t$, this was pointed out earlier in Bertsimas, Iancu and Parrilo (2010) and such an example was explicitly constructed in Delage and Iancu (2015). In the next section, we are going to demonstrate in the setting of distributionally robust inventory models that time inconsistent optimal policies, which are no longer of base-stock form, exist already in the two period case, in a general framework of moment constraints (Section 3.1) and spectral risk measures (Section 3.2).

3. Examples of time inconsistent optimal policies

In this section, we investigate existence of time inconsistent optimal policies in the case of two-period inventory model when $T = 2$. We assume that the considered distributions of the demand vector (D_1, D_2) are of the form $P_1 \times P_2$ with $P_1 \in \mathfrak{M}_1$ and $P_2 \in \mathfrak{M}_2$ being marginal distributions of D_1 and D_2 on intervals $[\alpha_1, \beta_1]$ and $[\alpha_2, \beta_2]$ respectively. We consider two settings for defining the set $\mathfrak{M}_1$ and functional $\mathcal{R}$. In Section 3.2, we assume existence of a reference probability measure on the interval $[\alpha_1, \beta_1]$. As it was already mentioned in the previous section, in that case probabilistic statements are made with respect to the reference probability measure and the constraints in Problem (7) below are understood to hold for almost every D_1. In Section 3.1, we assume that the set $\mathfrak{M}_1$ consists of all probability measures on $[\alpha_1, \beta_1]$ satisfying the specified moment constraints. In that section the constraints should be satisfied for all $D_1 \in [\alpha_1, \beta_1]$. In both cases the set $\mathfrak{M}_2$ does not play an essential role and can be arbitrary.

Before stating our main results, let us discuss the setup of our two-period risk averse (distributionally robust) inventory model. For $T = 2$ the corresponding Problem (2) becomes (up to the constant $-c_1 y_1$)

$$\begin{aligned} \min_{x_1, x_2(\cdot)} \; & (c_1 - c_2)x_1 + \mathcal{R}\left[\psi_1(x_1, D_1) + c_2(x_2(D_1) + D_1) + \psi_2(x_2(D_1), D_2)\right] \\ \text{s.t.} \; & x_1 \geq y_1, \; x_2(D_1) \geq x_1 - D_1. \end{aligned} \tag{7}$$

Since $\mathcal{R}(\cdot) = \rho_1(\rho_{2|D_1}(\cdot))$ with functionals ρ_1 and ρ_2 of the form (5), the objective function in (7) can be written as

$$(c_1 - c_2)x_1 + \sup_{P_1 \in \mathfrak{M}_1} \mathbb{E}_{P_1}\left[\psi_1(x_1, D_1) + c_2(x_2(D_1) + D_1) + \sup_{P_2 \in \mathfrak{M}_2} \mathbb{E}_{P_{2|D_1}}\left[\psi_2(x_2(D_1), D_2)\right]\right].$$

The second stage cost-to-go function $V_2(x_1, d_1)$ here is given by the optimal value of the problem

$$\min_{x_2} \left\{ c_2 x_2 + \sup_{P_2 \in \mathfrak{M}_2} \mathbb{E}_{P_2}\left[\psi_2(x_2, D_2)\right] \right\} \text{ s.t. } x_2 \geq x_1 - d_1, \tag{8}$$

and the first stage problem is

$$\min_{x_1 \geq y_1} (c_1 - c_2)x_1 + \sup_{P_1 \in \mathfrak{M}_1} \mathbb{E}_{P_1}\left[c_2 D_1 + \psi_1(x_1, D_1) + V_2(x_1, D_1)\right]. \tag{9}$$

Let ϑ^* be its optimal cost, $\bar{x}_1$ be an optimal solution of the first stage optimization problem, and x_2^* be an optimal solution of the unconstrained version (i.e., after removing the feasibility constraint $x_2 \ge x_1 - d_1$) of Problem (8). If the unconstrained version of Problem (8) has multiple optimal solutions, we assume that x_2^* is the largest one when constructing time inconsistent policies. Then $\bar{x}_2(d_1) := \max\{x_2^*, \bar{x}_1 - d_1\}$ is an optimal solution of the second stage problem, which is a base-stock policy. Therefore, there always exists a time consistent optimal policy. Note that every time consistent policy is of base-stock form. In addition, the time consistent policy is unique if the first stage optimal solution $\bar{x}_1$ is unique and the unconstrained version of Problem (8) has a unique optimal solution. Any time inconsistent optimal policy (if it exists) is inferior to the time consistent solutions in the sense that for *some* realizations of the demand D_1 the corresponding second stage cost is strictly worse than the one of time consistent policy. There does not exist a time inconsistent optimal policy if the functional ρ_1 is *strictly* monotone. Without strict monotonicity, we are going to demonstrate in Sections 3.1 and 3.2 that there may exist many time inconsistent optimal policies.

Consider function

$$\Psi(x_2, d_1) := \psi_1(\bar{x}_1, d_1) + c_2 d_1 + c_2 x_2 + \sup_{P_2 \in \mathfrak{M}_2} \mathbb{E}_{P_2}\big[\psi_2(x_2, D_2)\big],$$

where $\psi_t(\cdot,\cdot)$ are defined in (4). Then an optimal policy of the respective problem (7) $(\bar{x}_1, \hat{x}_2(d_1))$ is time consistent iff $\hat{x}_2(d_1)$ is an optimal solution of the problem

$$\min_{x_2 \ge \bar{x}_1 - d_1} \Psi(x_2, d_1) \tag{10}$$

for *any realization* $d_1 \in [\alpha_1, \beta_1]$. Note that base-stock policy $\bar{x}_2(d_1) = \max\{x_2^*, \bar{x}_1 - d_1\}$ is the optimal solution of (10) for any $d_1 \in [\alpha_1, \beta_1]$ and

$$\vartheta^* = \mathcal{R}\left[\min_{x_2 \ge \bar{x}_1 - D_1} \Psi(x_2, D_1)\right] = \mathcal{R}[\Psi(\bar{x}_2(D_1), D_1)],$$

where $\mathcal{R}(Z) = \sup_{P \in \mathfrak{M}_1} \mathbb{E}_P[Z]$. Hence in order to show that a policy $(\bar{x}_1, \hat{x}_2(d_1))$ is optimal time inconsistent, we need to prove that $\vartheta^* = \mathcal{R}[\Psi(\hat{x}_2(D_1), D_1)]$ and $\hat{x}_2(\cdot)$ is not optimal to Problem (10) for some realization. We additionally define

$$\bar{\varphi}(d_1) := \min_{x_2 \ge \bar{x}_1 - d_1} \Psi(x_2, d_1) = \Psi(\bar{x}_2(d_1), d_1) = \psi_1(\bar{x}_1, d_1) + c_2 d_1 + \phi(d_1), \tag{11}$$

where

$$\phi(d_1) := \min_{x_2 \geq \bar{x}_1 - d_1} \left\{ c_2 x_2 + \sup_{P_2 \in \mathfrak{M}_2} \mathbb{E}_{P_2}\left[\psi_2(x_2, D_2)\right] \right\}.$$

It is possible to show that if there exists one optimal time inconsistent policy, then there is an infinite number of such time inconsistent policies. In the rest of the section, we demonstrate existence of time inconsistent optimal policies for some examples of non-strictly monotone functional $\mathcal{R}$. Again $\mathfrak{M}_2$ can be arbitrary and we will focus on the discussion of $\mathfrak{M}_1$.

3.1. Moment constraints

In this section we discuss existence of time inconsistent policies when the set $\mathfrak{M}_1$ of probability measures is defined by the first n moment constraints for *any* $n \geq 0$. Consider function $\bar{\varphi}(\cdot)$, defined in (11), and the following problem of moments

$$\max_{P \in \mathfrak{S}} \mathbb{E}_P[\bar{\varphi}(D)] \text{ subject to } \mathbb{E}_P[D^k] = \mu_k,\ k = 0, ..., n, \tag{12}$$

where $\mathfrak{S}$ denotes the set of measures on the interval $[\alpha_1, \beta_1]$. Let $\mathfrak{C}_{n+1}$ be the set of points $(\mu_0, ..., \mu_n) \in \mathbb{R}^{n+1}$ for which there exists at least one measure $P \in \mathfrak{S}$ satisfying constraints of problem (12). Note that $\mathfrak{C}_{n+1}$ includes the null vector corresponding to the zero measure and is a convex closed cone, and for $\mu_0 = 1$ the corresponding measure becomes a probability measure. We assume that the considered $\vec{\mu} := (\mu_0, \mu_1, \ldots, \mu_n) \in \mathbb{R}^{n+1}$ with $\mu_0 = 1$, is an interior point of $\mathfrak{C}_{n+1}$. This assumption essentially requires that Problem (12) has a feasible solution, i.e., $\vec{\mu} \in \mathfrak{C}_{n+1}$, and $\vec{\mu}$ is not degenerate, i.e., not a boundary point of $\mathfrak{C}_{n+1}$. We refer the reader to the Appendix for a further (self-contained) discussion of the set $\mathfrak{C}_{n+1}$.

Problem (12) has a dual that can be written as the following semi-infinite programming problem (e.g., Section 6.6 of Shapiro, Dentcheva and Ruszczyński 2014):

$$\begin{aligned} \min_{z \in \mathbb{R}^{n+1}} \ & z_0 + \mu_1 z_1 + ... + \mu_n z_n \\ \text{s.t.} \ & \bar{\varphi}(t) \leq z_0 + t z_1 + ... + t^n z_n,\ t \in [\alpha_1, \beta_1]. \end{aligned} \tag{13}$$

Since the interval $[\alpha_1, \beta_1]$ is a compact set, there is no duality gap between the primal problem (12) and its dual problem (13). Moreover, the set of optimal solutions of the dual problem (13) is nonempty and bounded iff $\vec{\mu}$ is an interior point of $\mathfrak{C}_{n+1}$ (e.g., Proposition 3.4 in Shapiro 2001). For a feasible point z of Problem (13), consider the index set of active constraints:

$$\Delta(z) := \{t \in [\alpha_1, \beta_1] : \bar{\varphi}(t) = z_0 + tz_1 + ... + t^n z_n\}. \tag{14}$$

Now we state our first main result of the paper, which provides a necessary and sufficient condition for the existence of time inconsistent optimal policies in the moment setting.

THEOREM 1. *Suppose that $\vec{\mu}$ is an interior point of $\mathfrak{C}_{n+1}$. Then there exists an infinite number of optimal time inconsistent policies if and only if the dual problem* (13) *has an optimal solution $\bar{z}$ such that*

$$\Delta(\bar{z}) \neq [\alpha_1, \beta_1]. \tag{15}$$

We would like to point out that the duality argument in the proof of Theorem 1 is quite general and does not rely on the specific inventory setting.

Note that if $\Delta(\bar{z}) = [\alpha_1, \beta_1]$, then the function $\bar{\varphi}(\cdot)$ defined in (11), is a polynomial with a degree of no more than n on the interval $[\alpha_1, \beta_1]$. Condition (15) ensures that this does not happen. Condition (15) could be violated only in rather specific cases. The intuition behind this condition is as follows. Suppose that condition (15) holds. It follows that there exist $\epsilon > 0$ and $d_0 \in [\alpha_1, \beta_1]$ such that $\bar{\varphi}(d_0) + \epsilon < \bar{z}_0 + d_0\bar{z}_1 + ... + d_0^n \bar{z}_n$. Then we can construct another policy π_0 that is identical to the optimal time consistent base-stock policy $(\bar{x}_1, \bar{x}_2(\cdot))$ except π_0 ordering an additional ϵ amount of inventory in period 2 when demand is realized as d_0 in period 1. In that case, π_0 is also optimal by the duality, but is no longer time consistent. Note that the above procedure provides a way to derive optimal inconsistent policies by using the optimal solution $\bar{z}$ to problem (13) for the moment setting.

In the robust setting, which can be viewed as $n=0$, problem (13) becomes the problem of minimization of z_0 subject to $z_0 \geq \bar{\varphi}(t)$, $t \in [\alpha_1, \beta_1]$. Since function $\bar{\varphi}(\cdot)$ is not constant on the interval $[\alpha_1, \beta_1]$ condition (15) always holds, and hence in the robust setting there always exists an infinite number of time inconsistent policies (cf., Delage and Iancu 2015).

Corollary 1. *In the robust setting (i.e., $n=0$), there exists an infinite number of time inconsistent policies.*

Suppose now that $n \geq 1$. By verifying (15), the following corollary states that optimal time inconsistent policies always exist if the first stage problem solution $\bar{x}_1$ belongs to (α_1, β_1).

Corollary 2. *Suppose that $n \geq 1$ and the first stage problem* (9) *has an optimal solution $\bar{x}_1$ such that $\alpha_1 < \bar{x}_1 < \beta_1$. Then there exists an infinite number of optimal time inconsistent policies.*

Theorem 1 and Corollary 2 together suggest that only when $\bar{x}_1$ is equal to one of the edge points of the interval $[\alpha_1, \beta_1]$ or falls outside of the interval, it could happen that the problem does not have time inconsistent policies. We demonstrate this point through the following result. This shows a difference from the robust setting in which optimal time inconsistent policies always exist.

Corollary 3. *Suppose that $n \geq 1$ and $\beta_1 \leq \bar{x}_1 \leq x_2^*$. Then there are no optimal time inconsistent policies.*

The intuition behind the assumptions in Corollary 3 is as follows. If $\bar{x}_1$ is sufficiently high (e.g., higher than β_1, the upper bound of the support) possibly due to high initial inventory level or high ordering cost in period 2, then there will be leftover inventory in the end of period 1 almost surely no matter which P_1 is chosen from $\mathfrak{M}_1$. Hence the expected cost incurred in period 1 is the same for all P_1 as they have the same mean. Therefore, any policy whose inventory level after ordering deviates from the optimal base-stock level x_2^* in period 2 for some $d_1 \in [\alpha_1, \beta_1]$, is no longer optimal. It follows that all optimal policies must be time consistent.

3.2. Spectral risk measures

In this section we discuss existence of time inconsistent policies when the functional $\mathcal{R}$ is a spectral risk measure. We assume existence of a reference probability measure $\mathbb{P}$, associated with demand D_1, and that *the support of* $\mathbb{P}$ *is the interval* $[\alpha_1, \beta_1]$, i.e., $\mathbb{P}([\alpha_1, \beta_1]) = 1$ and for any nonempty open set $A \subset [\alpha_1, \beta_1]$ it follows that $\mathbb{P}(A) > 0$.

The spectral risk measure functional is defined as

$$\mathcal{R}(Z) := \int_0^1 \sigma(t) F_Z^{-1}(t) dt, \tag{16}$$

where $F_Z(z) = \mathbb{P}(Z \le z)$ is the cumulative distribution function (cdf) of Z (with respect to the reference measure $\mathbb{P}$), $F_Z^{-1}(t) = \inf\{z : F_Z(z) \ge t\}$ is the respective quantile, and $\sigma : [0,1) \to \mathbb{R}_+$ is a monotonically nondecreasing, right side continuous function such that $\int_0^1 \sigma(t) dt = 1$. For example[5], for $\sigma(\cdot) := (1-\gamma)^{-1} \mathbf{1}_{[\gamma,1]}(\cdot)$ with $\gamma \in [0,1)$, this becomes the so-called Average Value-at-Risk measure $\mathsf{AV@R}_\gamma$, which is defined as $\mathsf{AV@R}_\gamma(Z) := \inf_{t \in \mathbb{R}} \left\{ t + (1-\gamma)^{-1} \mathbb{E}\left[(Z-t)_+\right] \right\}$. It is known that spectral risk measure (16) is strictly monotone iff the corresponding spectral function $\sigma(t)$ is strictly positive for all $t \in (0,1)$ (e.g., Shapiro, Dentcheva and Ruszczyński 2014). In other words, spectral risk measure (16) is not strictly monotone iff there exists $\gamma \in (0,1)$ such that $\sigma(\gamma) = 0$. In particular, the $\mathsf{AV@R}_\gamma$ risk measure is not strictly monotone for any $\gamma \in (0,1)$.

Now we state our second main result of the paper, which provides general sufficient conditions for time inconsistency when the spectral risk measure functional is not strictly monotone. Recall that function $\bar{\varphi}(\cdot)$, defined in (11), is convex and continuous. We make the following assumption.

Assumption 1. *The function* $\bar{\varphi}(\cdot)$ *is not constant on any subinterval of the interval* $[\alpha_1, \beta_1]$.

Theorem 2. *Consider spectral risk measure* (16). *Suppose that* $\sigma(\gamma) = 0$ *for some* $\gamma \in (0,1)$, *and Assumption* 1 *holds. Then there exists an infinite number of optimal time inconsistent policies.*

Assumption 1 is a mild condition and it could be violated in rather specific cases. For example, this assumption holds if $c_2 = 0$ and the first stage problem (9) has an optimal solution $\bar{x}_1$ such that $\bar{x}_1 \geq \alpha_1$. Indeed, under these conditions, one can check that $\bar{\varphi}(d_1) = h_1(\bar{x}_1 - d_1)_+ + b_1(d_1 - \bar{x}_1)_+ + \phi(d_1)$, where $\phi(d_1)$ is non-increasing and $\phi(d_1) = \phi(\bar{x}_1)$ for all $d_1 \geq \bar{x}_1$. Hence Assumption 1 is satisfied. This assumption also holds if $\bar{x}_1 \leq x_2^*$ and $h_1 > c_2$. Indeed, under these conditions, one can check that $\bar{\varphi}(d_1) = h_1(\bar{x}_1 - d_1)_+ + b_1(d_1 - \bar{x}_1)_+ + c_2 d_1 + \phi(0)$, which satisfies Assumption 1.

Let us use $\mathsf{AV@R}_\gamma$ as an example to explain the intuition behind the result in Theorem 2. Note that the optimal cost $\mathsf{AV@R}_\gamma(\bar{\varphi}(D_1))$ can be interpreted as the average cost on the worst $(1-\gamma)$ fraction of the possible outcomes $\bar{\varphi}(\cdot)$ under the optimal time consistent base-stock policy $(\bar{x}_1, \bar{x}_2(\cdot))$. Similar to the intuition behind Theorem 1, starting from this optimal base-stock policy, we can construct another policy π_0 by ordering an additional ϵ amount of inventory in period 2 for some demand realization that falls in the best γ fraction of the possible outcomes $\bar{\varphi}(\cdot)$. In this case, the worst $(1-\gamma)$ fraction of the possible outcomes remains the same such that π_0 is also optimal. Note that the above procedure provides a way to derive optimal inconsistent policies for the spectral risk measure setting. Moreover, the above explanation can be possibly extended to any spectral risk measure that is not strictly monotone.

4. Conclusion

In this paper, we investigated optimal policies of distributionally robust (risk averse) inventory models. We demonstrated through several examples that if the respective risk measures are not strictly monotone, then there may exist infinitely many optimal policies which are not base-stock and not time consistent. This is in a sharp contrast with the risk neutral formulation of the inventory model where all optimal policies are time consistent. This extends previous studies of time inconsistency in the robust setting.

5. Appendix

5.1. Proofs

In this section, we provide all the proofs to the results in the main body of the paper. We first prove the following auxiliary result about the set $\Delta(z)$ defined in (14).

Lemma 1. *If $\Delta(\bar z) = [\alpha_1, \beta_1]$ for some optimal solution $\bar z$ of Problem* (13)*, then any $P \in \mathfrak{M}_1$ is an optimal solution of Problem* (12)*.*

Proof of Lemma 1. Suppose that $\Delta(\bar z) = [\alpha_1, \beta_1]$ for some optimal solution $\bar z$. Then for any $P \in \mathfrak{M}_1$ we have that

$$\mathbb{E}_P[\bar\varphi(D)] = \mathbb{E}_P[\bar z_0 + \bar z_1 D + \dots + \bar z_n D^n] = \bar z_0 + \bar z_1 \mu_1 + \dots + \bar z_n \mu_n.$$

That is, $\mathbb{E}_P[\bar\varphi(D)]$ is the same for every $P \in \mathfrak{M}_1$, and hence every $P \in \mathfrak{M}_1$ is an optimal solution of Problem (12). □

We now proceed and present all the proofs.

Proof of Theorem 1. Let $\vec\mu$ be an interior point of the cone $\mathfrak{C}_{n+1}$, and hence the dual problem (13) possesses an optimal solution $\bar z$. Suppose that condition (15) holds. Note that since the objective function of the dual problem is linear, the corresponding set $\Delta(\bar z)$ of active constraints is nonempty. By the first order optimality conditions we have that a feasible point $\bar z$ of Problem (13) is optimal iff there exist points $t_i \in \Delta(\bar z)$, and $\lambda_i \ge 0$, $i = 1, \dots, k$, with $k \le n+1$, such that

$$\vec\mu - \sum_{i=1}^{k} \lambda_i g_i = 0,$$

where $g_i := (1, t_i, \dots, t_i^n) \in \mathbb{R}^{n+1}$. Let us note that the set $\Delta(\bar z)$ is closed, and by the assumption (15) we can choose $\bar z$ such that $\Delta(\bar z) \ne [\alpha, \beta]$. So we can choose a nonempty closed interval $I \subset [\alpha, \beta] \setminus \Delta(\bar z)$.

Consider a policy $\tilde x_2(\cdot)$ such that $\tilde x_2(d_1) := \bar x_2(d_1)$ for all $d_1 \in [\alpha_1, \beta_1] \setminus I$, and $\tilde x_2(d_1) := \bar x_2(d_1) + \varepsilon$ for $d_1 \in I$ and some $\varepsilon > 0$. Clearly this policy is feasible and hence the value

$$\sup_{P \in \mathfrak{M}_1} \mathbb{E}_P \left[\Psi(\tilde x_2(D_1), D_1) \right] \tag{17}$$

is not less than ϑ^*. Let us show that for $\varepsilon > 0$ small enough actually the value of (17) is ϑ^*, and hence the policy with $\bar x_2(\cdot)$ replaced by $\tilde x_2(\cdot)$ is also optimal.

Consider function $\tilde{\varphi}(d_1) := \Psi(\tilde{x}_2(d_1), d_1)$ and the dual semi-infinite program of Problem (17). We have that $\tilde{\varphi}(d_1) = \bar{\varphi}(d_1)$ for $d_1 \in [\alpha_1, \beta_1] \setminus I$, and $\sup_{d_1 \in I} |\tilde{\varphi}(d_1) - \bar{\varphi}(d_1)|$ can be made arbitrarily small for sufficiently small $\varepsilon > 0$. Since $\bar{\varphi}(t) < \bar{z}_0 + t\bar{z}_1 + ... + t^n \bar{z}_n$ for all $t \in I$, we can choose $\varepsilon > 0$ small enough such that $\tilde{\varphi}(t) < \bar{z}_0 + t\bar{z}_1 + ... + t^n \bar{z}_n$ for all $t \in I$, and hence $\tilde{\varphi}(t) \leq \bar{z}_0 + t\bar{z}_1 + ... + t^n \bar{z}_n$ for all $t \in [\alpha_1, \beta_1]$. It follows that $\bar{z}$ is a feasible point of the dual semi-infinite program of Problem (17). Moreover, the corresponding set of active constraints, which is obtained by replacing $\bar{\varphi}$ in (14) with $\tilde{\varphi}$, does not change. By the first order optimality conditions the optimal solution $\bar{z}$ of Problem (13) is also an optimal solution of the dual of Problem (17), and hence these two semi-infinite programs have the same optimal value ϑ^*. It follows that the optimal value of Problem (17) is also ϑ^*. This shows existence of an optimal solution different from $\bar{x}_2(\cdot)$.

Conversely suppose that condition (15) does not hold, which implies that $\Delta(\bar{z}) = [\alpha_1, \beta_1]$ for some dual optimal solution $\bar{z}$. Then by Lemma 1, any $P \in \mathfrak{M}_1$ is optimal for Problem (12). We argue now by a contradiction. Suppose that there is an optimal time inconsistent policy $(\bar{x}_1, \tilde{x}_2(\cdot))$. It follows that there exists $d_0 \in [\alpha_1, \beta_1]$ such that $\bar{x}_2(d_0) \neq \tilde{x}_2(d_0)$ and $\Psi(\bar{x}_2(d_0), d_0) < \Psi(\tilde{x}_2(d_0), d_0)$. Also by optimality of $\bar{x}_2(\cdot)$ we have that $\Psi(\bar{x}_2(d), d) \leq \Psi(\tilde{x}_2(d), d)$ for all $d \in [\alpha_1, \beta_1]$. Moreover, by Theorem 3(iv), there exists $P_0 \in \mathfrak{M}_1$ that has a positive mass at the point d_0. Hence

$$\vartheta^* = \mathbb{E}_{P_0}\left[\Psi(\bar{x}_2(D_1), D_1)\right] < \mathbb{E}_{P_0}\left[\Psi(\tilde{x}_2(D_1), D_1)\right] \leq \sup_{P \in \mathfrak{M}_1} \mathbb{E}_P\left[\Psi(\tilde{x}_2(D_1), D_1)\right].$$

This implies that the policy $(\bar{x}_1, \tilde{x}_2(\cdot))$ is not optimal, giving the desired contradiction. □

Proof of Corollary 2. The subdifferential of the function $\psi_1(\bar{x}_1, \cdot)$ is not a singleton at $d_1 = \bar{x}_1$. By (11) it follows that $\bar{\varphi}(d_1)$ is not differentiable at $d_1 = \bar{x}_1$. Since $\bar{x}_1$ is an interior point of the interval $[\alpha_1, \beta_1]$, it follows that $\bar{\varphi}(\cdot)$ cannot coincide with a polynomial on the interval $[\alpha_1, \beta_1]$. The proof can be concluded by applying Theorem 1. □

Proof of Corollary 3. Under the conditions, we have

$$\bar{\varphi}(d_1) = h_1(\bar{x}_1 - d_1) + c_2 d_1 + \phi(0), \tag{18}$$

which is linear in d_1. Hence there exists an optimal solution of Problem (13) that exactly coincides $\bar{\varphi}(d_1)$. Therefore, Assumption (15) is violated. $\square$

Proof of Theorem 2. Let $\bar{F}(z) := \mathbb{P}(\bar{\varphi}(D_1) \le z)$ be the cdf of random variable $\bar{\varphi}(D_1)$. Consider $z^* := \bar{F}^{-1}(\gamma)$. We have that $\mathbb{P}(\bar{\varphi}(D_1) \le z^*) \ge \gamma$ and $\mathbb{P}(\bar{\varphi}(D_1) < z^*) \le \gamma$. Since $\bar{\varphi}(\cdot)$ is not constant on the interval $\{d_1 \in [\alpha_1, \beta_1] : \bar{\varphi}(d_1) \le z^*\}$, it follows that the set $\{d_1 \in [\alpha_1, \beta_1] : \bar{\varphi}(d_1) < z^*\}$ is nonempty. Since the function $\bar{\varphi}$ is convex continuous, this set is an open interval. Let A be a non-degenerate closed subinterval of $\{d_1 \in [\alpha_1, \beta_1] : \bar{\varphi}(d_1) < z^*\}$. Since $[\alpha_1, \beta_1]$ is the support of $\mathbb{P}$, it follows that $\mathbb{P}(A) > 0$.

We have then that there exists (measurable) function $\tilde{x}_2 : [\alpha_1, \beta_1] \to \mathbb{R}$ such that $\tilde{x}_2(d_1) = \bar{x}_2(d_1)$ for all $d_1 \in [\alpha_1, \beta_1] \setminus A$, and $\tilde{x}_2(d_1) > \bar{x}_2(d_2)$ and $\Psi(\tilde{x}_2(d_1), d_1) < z^*$ for all $d_1 \in A$. Consider the cdf $\tilde{F}(\cdot)$ of the random variable $\Psi(\tilde{x}_2(D_1), D_1)$, i.e., $\tilde{F}(z) := \mathbb{P}(\Psi(\tilde{x}_2(D_1), D_1) \le z)$. We have then that $\tilde{F}^{-1}(t) = \bar{F}^{-1}(t)$ for all $t \in [\gamma, 1]$. Combining with the monotonicity of $\sigma(t)$, it follows that $\mathcal{R}(\bar{Z}) = \mathcal{R}(\tilde{Z})$, where $\bar{Z} := \bar{\varphi}(D_1)$ and $\tilde{Z} := \Psi(\tilde{x}_2(D_1), D_1)$. Thus the policy $(\bar{x}_1, \tilde{x}_2(d_1))$ has the same value as the policy $(\bar{x}_1, \bar{x}_2(d_1))$, and hence is also optimal. $\square$

5.2. The problem of moments

In this section, we review some classical results for the problem of moments. In order to simplify notation, we drop the subscripts and write $[\alpha, \beta]$ for the first stage interval. We address the question for what points $\vec{\mu} := (\mu_0, \mu_1, \ldots, \mu_n) \in \mathbb{R}^{n+1}$ there exists a measure P on the interval $[\alpha, \beta]$ satisfying the moment constraints

$$\int_{\alpha_1}^{\beta_1} t^k dP(t) = \mu_k,\ k = 0, ..., n. \tag{19}$$

If such a (nonnegative) measure P exists we say that it is a *representing* measure. By the Richter - Rogosinski Theorem (e.g., Theorem 7.37 in Shapiro, Dentcheva and Ruszczyński 2014), we have that if such representing measure exists, then there exists a representing measure P having support of no more than $n+1$ points, i.e., $P = \sum_{i=1}^{\ell} p_i \delta_{t_i}$ for some[6] $t_i \in [\alpha, \beta]$ and $\ell \le n+1$. Note that μ_0

is the normalization constant. That is, if P is a positive representing measure, then $\mu_0 > 0$ and $Q := \mu_0^{-1} P$ is a probability measure satisfying the moment constraints $\mathbb{E}_Q[D^k] = \mu_k/\mu_0$, $k = 1, ..., n$.

By $\mathfrak{C}_{n+1}$ we denote the set of points $\vec{\mu} \in \mathbb{R}^{n+1}$ for which there exists at least one representing measure. Note that $\mathfrak{C}_{n+1}$ includes the null vector corresponding to the zero measure and is a convex closed cone. For $\vec{\mu} \in \mathbb{R}^{n+1}$ consider the following so-called Hankel matrices. For an even $n = 2m$ consider matrices defined as

$$\underline{H}_{2m}(\vec{\mu}) = [\mu_{i+j}]_{i,j=0}^{m} \text{ and } \overline{H}_{2m}(\vec{\mu}) = [(\alpha+\beta)\mu_{i+j+1} - \mu_{i+j+2} - \alpha\beta\mu_{i+j}]_{i,j=0}^{m-1},$$

and for an odd $n = 2m+1$ consider matrices defined as

$$\underline{H}_{2m+1}(\vec{\mu}) = [\mu_{i+j+1} - \alpha\mu_{i+j}]_{i,j=0}^{m} \text{ and } \overline{H}_{2m+1}(\vec{\mu}) = [\beta\mu_{i+j} - \mu_{i+j+1}]_{i,j=0}^{m}.$$

Note that matrices $\underline{H}_{2m}(\vec{\mu})$ and $\overline{H}_{2m}(\vec{\mu})$ are of the respective orders $(m+1)\times(m+1)$ and $m \times m$; and matrices $\underline{H}_{2m+1}(\vec{\mu})$ and $\overline{H}_{2m+1}(\vec{\mu})$ are of order $(m+1)\times(m+1)$.

The following theorem summarizes classical results relevant to our problem, and we refer to Chapter 10 of Schmüdgen (2017) for a further discussion. Recall that the boundary of the convex closed cone $\mathfrak{C}_{n+1}$ is the set $\mathfrak{C}_{n+1} \setminus \text{int}(\mathfrak{C}_{n+1})$, where $\text{int}(\mathfrak{C}_{n+1})$ denotes the interior of $\mathfrak{C}_{n+1}$.

THEOREM 3. (i) *Point $\vec{\mu} \in \mathbb{R}^{n+1}$ belongs to the cone $\mathfrak{C}_{n+1}$ iff the matrices $\underline{H}_n(\vec{\mu})$ and $\overline{H}_n(\vec{\mu})$ are positive semidefinite.* (ii) *Point $\vec{\mu} \in \mathbb{R}^{n+1}$ belongs to the interior of $\mathfrak{C}_{n+1}$ iff the matrices $\underline{H}_n(\vec{\mu})$ and $\overline{H}_n(\vec{\mu})$ are positive definite.* (iii) *If $\vec{\mu} \in \mathbb{R}^{n+1}$ is a boundary point of $\mathfrak{C}_{n+1}$, then the corresponding representing measure is unique.* (iv) *If $\vec{\mu}$ is an interior point of $\mathfrak{C}_{n+1}$, then for any $d \in [\alpha, \beta]$ there exists a representing measure P having a positive mass at the point d.*

For example consider $n = 1$. Then Hankel matrices are 1×1 matrices $[\mu_1 - \alpha\mu_0]$ and $[\beta\mu_0 - \mu_1]$. Hence the corresponding cone $\mathfrak{C}_2 = \{(\mu_0, \mu_1) : \mu_1 - \alpha\mu_0 \geq 0,\ \beta\mu_0 - \mu_1 \geq 0\}$, which implies that $\mu_0 \geq 0$. For $\mu_0 = 1$ the corresponding representing set of probability measures is nonempty iff

$\mu_1 \in [\alpha, \beta]$. The interior of $\mathfrak{C}_2$ is defined by the constraints $\mu_1 - \alpha\mu_0 > 0$ and $\beta\mu_0 - \mu_1 > 0$, and the boundary is attained when either $\mu_1 - \alpha\mu_0 = 0$ or $\beta\mu_0 - \mu_1 = 0$.

For $n = 2$ the corresponding Hankel matrices are

$$\underline{H}_2(\vec{\mu}) = \begin{bmatrix} \mu_0 & \mu_1 \\ \mu_1 & \mu_2 \end{bmatrix} \text{ and } \overline{H}_2(\vec{\mu}) = [(\alpha+\beta)\mu_1 - \mu_2 - \alpha\beta\mu_0]$$

and hence

$$\mathfrak{C}_3 = \{(\mu_0, \mu_1, \mu_2) : \mu_0 \geq 0,\ \mu_0\mu_2 - \mu_1^2 \geq 0,\ (\alpha+\beta)\mu_1 - \mu_2 - \alpha\beta\mu_0 \geq 0\}.$$

For $\mu_0 = 1$, in terms of the variance $\sigma^2 := \mu_2 - \mu_1^2$, the corresponding set of feasible solutions is defined by the inequalities $\sigma^2 \geq 0$ and $\sigma^2 \leq (\beta - \mu_1)(\mu_1 - \alpha)$. Boundary solutions happen if $\mu_1 \in [\alpha, \beta]$ and $\sigma^2 = 0$ or $\sigma^2 = (\beta - \mu_1)(\mu_1 - \alpha)$.

Endnotes

1. It is said that the demand process is stagewise independent if D_{t+1} is independent of $(D_1, ..., D_t)$, $t = 1, ..., T-1$.

2. By "robust setting" we mean the worst case setting, i.e., the max-type objective functional is used.

3. We denote by $d_{[t]} := (d_1, ..., d_t)$ history of the demand process up to time t.

4. By saying that policy is optimal we mean that it is an optimal solution of the corresponding multistage problem.

5. By $\mathbf{1}_A$ we denote the indicator function of set A.

6. By δ_t we denote measure of mass one at point t.

Acknowledgments

Work of the first author was partly supported by NSF grant 1633196 and DARPA EQUiPS program, grant SNL 014150709.

References

S. Ahmed, U. Cakmak, A. Shapiro. 2007. Coherent risk measures in inventory problems. *European Journal of Operational Research.* **182** 226-238.

P. Artzner, F. Delbaen, J.-M. Eber, D. Heath. 1999. Coherent measures of risk. *Mathematical Finance.* **9** 203–228.

A. Ben-Tal, A. Goryashko, E. Guslitzer, A. Nemirovski. 2004. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming.* **99**(2) 351-376.

D. Bertsimas, D. Iancu, P. Parrilo. 2010. Optimality of Affine Policies in Multistage Robust Optimization. *Mathematics of Operations Research.* **35** 363-394.

F. J. C. T. de Ruiter, R.C. M. Brekelmans, D. den Hertog. 2016. The impact of the existence of multiple adjustable robust solutions. *Mathematical Programming.* **160** 531-545.

E. Delage, D. Iancu. 2015. Robust Multistage Decision Making. *INFORMS Tutorials in Operations Research.* `http://dx.doi.org/10.1287/educ.2015.0139`.

E. Delage, Y. Ye. 2010. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research.* **58** 595–612.

P.M. Esfahani, D. Kuhn. 2018. Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming.* **171** 115-166.

D.A. Iancu, M. Petrik, D. Subramanian. 2015. Tight approximations of dynamic risk measures. *Mathematics of Operations Research.* **40**(3) 655-682.

D.A. Iancu, N. Trichakis. 2014. Pareto Efficiency in Robust Optimization. *Management Science.* **60**(1) 130–147.

R. Jiang, Y. Guan. 2016. Data-driven chance constrained stochastic program. *Mathematical Programming.* **158** 291-327.

A. Ruszczyński, A. Shapiro. 2006. Conditional risk mappings. *Mathematics of Operations Research.* **31** 544–561.

A. Ruszczyński. 2010. Risk-averse dynamic programming for Markov decision processes. *Mathematical Programming.* **125** 235–261.

H. Scarf. 1958. A min-max solution of an inventory problem. *Studies in the Mathematical Theory of Inventory and Production* (Chapter 12).

K. Schmüdgen. 2017. *The Moment Problem.* Springer International Publishing AG, Switzerland.

A. Shapiro. 2001. On duality theory of conic linear problems, in *Semi-Infinite Programming: Recent Advances,* Miguel A. Goberna and Marco A. Lopez, Eds., Kluwer Academic Publishers, pp. 135 – 165.

A. Shapiro. 2016. Rectangular sets of probability measures. *Operations Research.* **64** 528-541.

A. Shapiro. 2017. Interchangeability principle and dynamic equations in risk averse stochastic programming. *Operations Research Letters.* **45** 377-381.

A. Shapiro. 2018. Tutorial on risk neutral, distributionally robust and risk averse multistage stochastic programming. `http://www.optimization-online.org/DB_FILE/2018/02/6455.pdf`.

A. Shapiro, D. Dentcheva, A. Ruszczyński. 2014. *Lectures on Stochastic Programming: Modeling and Theory.* Second Edition, SIAM, Philadelphia.

W. Wiesemann, D. Kuhn, B. Rustem. 2013. Robust Markov decision processes. *Mathematics of Operations Research.* **38** 153–183.

W. Wiesemann, D. Kuhn, M. Sim. 2014. Distributionally robust convex optimization. *Operations Research.* **62** 1358–1376.

L. Xin, D.A. Goldberg. 2013. Time (in) consistency of multistage distributionally robust inventory models with moment constraints. `arXiv preprint arXiv:1304.3074`.

P.H. Zipkin. 2000. *Foundations of Inventory Management.* McGraw-Hill, Boston.