Time (in)consistency of multistage distributionally robust inventory models with moment constraints

Linwei Xin
Booth School of Business, University of Chicago, Chicago, IL 60637

David A. Goldberg
Operations Research and Information Engineering, Cornell University, Ithaca, NY 14850

Recently, there has been a growing interest in developing inventory control policies which are robust to model misspecification. One approach is to posit that nature selects a worst-case distribution for any relevant stochastic primitives from some pre-specified family. Several communities have observed that a subtle phenomena known as time inconsistency can arise in this framework. In particular, it becomes possible that a policy which is optimal at time zero may not be optimal for the associated optimization problem in which the decision-maker recomputes her policy at each point in time, which has implications for implementability. If there exists a policy which is optimal for both formulations, we say that the policy is *time consistent*, and the problem is *weakly time consistent*. If every optimal policy is time consistent, we say that the problem is *strongly time consistent*. We study these phenomena in the context of managing an inventory over time, when only the mean, variance, and support are known for the demand at each stage. We provide several illustrative examples showing that here the question of time consistency can be quite subtle. We complement these observations by providing simple sufficient conditions for weak and strong time consistency. Although a similar phenomena was previously identified by Shapiro for the setting in which only the mean and support of the demand are known, here our model is rich enough to exhibit a variety of additional interesting behaviors.

Key words: supply chain management, inventory, news vendor, multistage distributionally robust optimization, rectangularity, moment constraints, time consistency, dynamic programming, base-stock policy

1. Introduction

The news vendor problem, used to analyze the trade-offs associated with stocking an inventory, has its origin in a seminal paper by Edgeworth (1888). In its classical formulation, the problem is stated as a minimization of the expected value of the relevant ordering, backorder, and holding costs. However, in applications knowledge of the exact distribution of the demand process is rarely available. This motivates the study of minimax type (i.e. distributionally robust) formulations, where minimization is performed with respect to a worst-case distribution from some family of potential distributions. In a pioneering paper Scarf (1958) gave an elegant solution for the minimax news vendor problem when only the first and second order moments of the demand distribution are known. His work has led to considerable follow-up work (cf. Gallego et al. (1993, 1994), Gallego (1998, 2001), Popescu (2005), Yue et al. (2006), Gallego et al. (2007), Perakis et al. (2008), Chen et al. (2009), See et al. (2010), Hanasusanto et al. (2015), Zhu et al. (2013)). For a more general overview of risk analysis for news vendor and inventory models we can refer, e.g., to Ahmed et al. (2007) and Choi et al. (2011). We also note that a distributionally robust minimax approach is not the only way to model such uncertainty, and that there is a considerable literature on alternative approaches such as the robust optimization paradigm (cf. Kasugai et al. (1961), Ben-Tal et al. (2005), Bertsimas et al. (2006), Ben-Tal et al. (2009), Bertsimas et al. (2010), Carrizosaa et al. (2016), Gabrel et al. (2014)) and Bayesian approach (cf. Scarf (1959, 1960), Lovejoy (1992), Levi et al. (2015), Klabjan et al. (2013)).

In practice an inventory must often be managed over some time horizon, and the classical news vendor problem was naturally extended to the multistage setting, for which there is also a considerable literature (cf. Zipkin (2000)). Recently, distributionally robust variants of such multistage problems have begun to receive attention in the literature (cf. Gallego (2001), Ahmed et al. (2007), Choi et al. (2008), See et al. (2010), Shapiro (2012), Klabjan et al. (2013)). It has been observed that such multistage distributionally robust optimization problems can exhibit a subtle phenomenon known as time inconsistency. Over the years various concepts of time consistency have been discussed in the economics literature, in the context of rational decision making. This can be traced back at least to the work of Strotz (1955) - for a more recent overview we refer the reader to the survey by Etner et al. (2012). Questions of time consistency have also attracted attention in the mathematical finance literature and the associated theory of coherent risk measures (cf. Wang (1999), Artzner et al. (2007), Roorda et al. (2007), Cheridito et al. (2009), Ruszczyński (2010)). These concepts have also been studied from the perspective of robust control across various academic communities (cf. Hansen et al. (2001), Iyengar (2005), Nilim et al. (2005), Grunwald et al. (2011), Carpentier et al. (2012), Osogami et al. (2012), Wiesemann et al. (2013), Pflug et al. (2014), Asamov et al. (2015), De Lara et al. (2016), Shapiro et al. (2016a,b), Bielecki et al. (2018), Gérard et al. (2019)). Recently, these concepts have also begun to receive attention in the setting of inventory control (cf. Chen et al. (2007, 2012), Yang (2013), Homem-de-Mello et al. (2016), Shapiro et al. (2019)).

In this work, we will consider questions of time (in)consistency in the context of managing an inventory over time. We will give several formal definitions of time consistency, which are naturally suited to our framework, in Section 4. At this point let us provide the following high-level intuition. A multi-stage distributionally robust optimization problem can be viewed in two ways. In one formulation, the policy maker is allowed to recompute her policy choice after each stage, thus taking prior realizations of demand into consideration when performing the relevant minimax calculations at later stages. One natural way to model such a re-computation leads to a tractable dynamic programming (DP) formulation (we will refer to this as the *distributionally robust DP formulation*), for which a base-stock policy is known to be optimal. In the second formulation, the policy maker is not allowed to recompute her policy after each stage (we will refer to this as the *multistage-static formulation*), in which case far less is known. If these two formulations have a common optimal policy, i.e. the policy maker would be content with the given policy whether or not she has the power to recompute after each stage, we say that the policy is *time consistent*, and the problem is *weakly time consistent*. If every optimal policy for the multistage-static formulation is time consistent, i.e. it is impossible to devise a policy which is optimal at time zero yet suboptimal at a later time, we say that the problem is *strongly time consistent*. Such a property is desirable from a policy perspective, as it ensures that previously agreed upon policy decisions remain rational when the policy is actually implemented, possibly at a later time. We will also discuss an alternative notion of time consistency, in which we require that the optimal policy for the multistage-static formulation agrees with the sequence of decisions derived under a different notion of re-computation, in which the re-computed optimization problems are themselves (conditioned) versions of the multistage-static formulation (over the remaining time horizon).

Within the optimization and inventory control communities, much of the work on time consistency restricts its discussion of optimal policies to the setting in which the family of distributions from which nature can select satisfies a certain factorization property called *rectangularity*, which endows the associated minimax problem with a DP structure, and which we review in-depth in the supplemental appendix Section 7. Outside of this setting, most of the literature focuses on discussing and demonstrating hardness of the underlying optimization problems (cf. Iyengar (2005), Nilim et al. (2005), Wiesemann et al. (2013)). We note that this is in spite of the fact that previous literature has discussed the importance and relevance of such non-decomposable formulations from a modeling perspective (cf. Iyengar (2005)).

1.1. Our contributions

In this paper, we depart from much of the past literature by seeking both negative *and positive* results regarding time consistency when no such decomposition holds, i.e. the underlying family of distributions from which nature can select is non-rectangular. Our work is in the spirit of Grunwald et al. (2011), in which a definition of (weak) time consistency similar to ours was analyzed in the context of rectangularity and dynamic consistency (a concept defined in Epstein et al. (2003)), albeit in a substantially different context. Our work can also be viewed as providing a more in-depth and inventory-focused study of several notions of time-consistency studied in Shapiro (2009), Carpentier et al. (2012), Pflug et al. (2014), Homem-de-Mello et al. (2016), Shapiro et al. (2016a).

In contrast to many of these works in which all concepts are explained through the language of risk measures, here we explain all relevant concepts purely in the language of (robust) newsvendor models with moment constraints, a model popular in the operations management community, and hope that in doing so our work brings the concept of time-consistency to a broader audience.

We extend the work of Scarf (1958) (and followup work of Gallego (2001)) by considering the question of time consistency in multistage news vendor problems when the support and first two moments are known for the demand at each stage, and demand is stage-wise independent. In addition to refining multiple definitions related to time-consistency, we provide several illustrative examples showing that here the question of time consistency can be quite subtle. In particular: (i) the problem can fail to be weakly time consistent, (ii) the problem can be weakly but not strongly time consistent, and (iii) the problem can be strongly time consistent even if every associated optimal policy takes different values under the multistage-static and distributionally robust DP formulations. We also prove that, although the distributionally robust DP formulation always has an optimal policy of base-stock form, there may be no such optimal policy for the multistage-static formulation. We complement these observations by providing simple sufficient conditions for weak and strong time consistency.

Interestingly, in contrast to much of the related literature, our results show that time consistency may hold even when rectangularity does not. This stands in contrast to the analysis of Shapiro (2012) for the setting in which only the mean and support of the demand distribution are known, where the problem is always (weakly) time consistent, amenable to a simple DP solution, with both formulations having the same optimal value. Likewise, in the setting in which only the support is known, both formulations reduce to the so-called adjustable robust formulation described in Ben-Tal et al. (2004), where again (weak) time consistency always holds.

1.2. Outline of paper

The structure of the rest of the paper is organized as follows. In Section 2, we review the single-stage classical and distributionally robust formulations and their properties. In Section 3, we discuss the extension to the multi-stage setting, formally defining the multistage-static formulation, the relevant notions of time-consistency, and the distributionally robust DP formulation. In Section 4, we prove our sufficient conditions for weak and strong time consistency, and present several illustrative examples showing that here the question of time consistency can be quite subtle. In Section 5, we provide closing remarks and directions for future research. We also include a technical appendix in Section 6 containing proofs of several of our main results. Additional supplemental material and exposition, including a detailed discussion of rectangularity, and (for completeness) some additional results and proofs, are included in the supplemental appendix Section 7.

2. Single-stage formulation

In this section we review both the classical and distributionally robust single-stage formulation.

2.1. Classical formulation

Consider the following classical formulation of the news vendor problem: $\inf_{x\geq 0}\mathbb{E}[\Psi(x,D)]$, where $\Psi(x,d) := cx + b[d-x]_+ + h[x-d]_+$; and c,b,h are the ordering, backorder penalty, and holding costs, per unit, respectively. Unless stated otherwise we assume that $b > c > 0$ and $h \geq 0$. The expectation is taken with respect to the probability distribution of the demand D, which is modeled as a random variable having nonnegative support. It is well-known that here an optimal choice of x is the $\frac{b-c}{b+h}$ quantile of demand.

2.2. Distributionally robust formulation

Suppose now that the probability distribution of the demand D is not fully specified, but instead assumed to be a member of a family of distributions $\mathfrak{M}$. Then we consider the following distributionally robust formulation:

$$\inf_{x\geq 0}\psi(x), \tag{1}$$

where $\psi(x) := \sup_{Q\in\mathfrak{M}}\mathbb{E}_Q[\Psi(x,D)]$, and the notation $\mathbb{E}_Q$ emphasizes that the expectation is taken with respect to the distribution Q of the demand D.

We now introduce some additional notations to describe certain families of distributions. For a probability measure (distribution) Q, we let $\text{supp}(Q)$ denote the support of the measure, i.e. the smallest closed set $A \subseteq \mathbb{R}$ such that $Q(A) = 1$. With a slight abuse of notation, for a random variable Z, we also let $\text{supp}(Z)$ denote the support of the associated probability measure. For a given closed interval $\mathcal{I} \subseteq \mathbb{R}$, we let $\mathfrak{P}(\mathcal{I})$ denote the set of probability distributions Q such that $\text{supp}(Q) \subseteq \mathcal{I}$, and let δ_a denote the probability measure of mass one at $a \in \mathbb{R}$.

In this paper, we will study families of distributions satisfying moment constraints of the form $\mathfrak{M} := \left\{Q \in \mathfrak{P}(\mathcal{I}) : \mathbb{E}_Q[D] = \mu, \mathbb{E}_Q[D^2] = \mu^2 + \sigma^2\right\}$. Unless stated otherwise, it will be assumed that $\mathfrak{M}$ is indeed of this form, and is nonempty (i.e. $\mu \in [\alpha,\beta]$ and $\sigma^2 \leq (\beta-\mu)(\mu-\alpha)$ for some $\alpha \leq \beta$, where α,β may equal $\pm\infty$). We refer to Chapter 10 of Schmüdgen (2017) for a further discussion of classical results about the problem of moments. One can identify conditions under which $\mathfrak{M}$ is a singleton, as follows.

Observation 1 *If $-\infty < \alpha < \beta < +\infty$, $\mu \in [\alpha,\beta]$, and $\sigma^2 = (\beta-\mu)(\mu-\alpha)$, then $\mathfrak{M}$ consists of the one measure assigning α probability $\frac{\beta-\mu}{\beta-\alpha}$, and β probability $\frac{\mu-\alpha}{\beta-\alpha}$.*

We now rephrase $\psi(x)$ as the optimal value of a certain optimization problem. For use in later proofs, we define the following more general maximization problem, in terms of a general integrable objective function ζ:

$$\begin{array}{ll} \sup\limits_{Q\in\mathfrak{P}(\mathcal{I})} & \int \zeta(\tau)dQ(\tau) \\ \text{s.t.} & \int \tau dQ(\tau) = \mu,\ \int \tau^2 dQ(\tau) = \mu^2+\sigma^2. \end{array} \tag{2}$$

Our definitions imply that for all $x \in \mathbb{R}$, $\psi(x)$ equals the optimal value of Problem (2) for the special case that $\zeta(\tau) = \Psi(x,\tau)$. We note that Problem (2) is a classical problem of moments (see, e.g., Landau 1987), and by the the Richter-Rogosinski Theorem (e.g., Shapiro et al. 2009, Proposition 6.40), or results in Bertsimas et al. (2005), has an optimal solution with support on at most three points (if it posesses an optimal solution).

2.2.1. Scarf's solution. We note that the distributionally robust single-stage news vendor problem considered by Scarf (1958) is exactly Problem (1), when $\mathcal{I} = \mathbb{R}_+$. As it will be useful for later proofs, we briefly review Scarf's explicit solution here, and show some additional related results in the technical appendix. Define $\mathfrak{f}(z) := \big((z-\mu)^2 + \sigma^2\big)^{\frac{1}{2}}$ for all $z \in \mathbb{R}$.

Theorem 1 (Scarf (1958)) *Suppose that $b > c$, $c + h > 0$, $\mu > 0$, $\sigma > 0$, and $\mathcal{I} = \mathbb{R}_+$. Let $\kappa := \frac{b-h-2c}{b+h}$. Then for each $x \in \mathbb{R}$,*

$$\psi(x) = \begin{cases} c\mu + \frac{b+h}{2}\big((x-\mu)^2 + \sigma^2\big)^{\frac{1}{2}} - \frac{b-h-2c}{2}(x-\mu), & \text{if } x \geq \frac{\mu^2+\sigma^2}{2\mu}, \\ \frac{(h+c)\sigma^2 - (b-c)\mu^2}{\mu^2+\sigma^2}x + b\mu, & \text{if } x \in [0, \frac{\mu^2+\sigma^2}{2\mu}), \\ b\mu - (b-c)x, & \text{otherwise.} \end{cases} \tag{3}$$

As a consequence, a complete solution to the problem $\inf_{x\in\mathbb{R}} \psi(x)$ is as follows.
(i) *If $\frac{\sigma^2}{\mu^2} > \frac{b-c}{h+c}$, then the unique optimal solution is $x = 0$, and the optimal value is μb.*
(ii) *If $\frac{\sigma^2}{\mu^2} < \frac{b-c}{h+c}$, then the unique optimal solution is $x = \mu + \kappa\sigma(1-\kappa^2)^{-\frac{1}{2}}$, and the optimal value is $c\mu + \big((h+c)(b-c)\big)^{\frac{1}{2}}\sigma$.*
(iii) *If $\frac{\sigma^2}{\mu^2} = \frac{b-c}{h+c}$, then all $x \in [0, \mu + \kappa\sigma(1-\kappa^2)^{-\frac{1}{2}}]$ are optimal, and the optimal value is μb.*
In all cases $\arg\max_{Q\in\mathfrak{M}} \mathbb{E}_Q[\Psi(x,D)]$ is nonempty for every $x \in \mathbb{R}$, and the optimal solution set and value of the problem $\inf_{x\in\mathbb{R}} \psi(x)$ is identical to that of Problem (1).

3. Multistage formulation
3.1. Classical formulation

We begin by giving a quick review of the classical (i.e. non-robust) multistage news vendor problem. For a vector $z = (z_1, ..., z_n) \in \mathbb{R}^n$ and $1 \leq i \leq j \leq n$, denote $z_{[i,j]} := (z_i, ..., z_j)$. In particular for $i = 1$ we simply write $z_{[j]}$ for the vector consisting of the first j components of z, and set $z_{[0]} := \emptyset$.

We suppose that there is a finite time horizon T, and a (random) vector of demands $D = (D_1, \dots, D_T)$. By $d = (d_1, \dots, d_T)$ we usually denote a particular realization of the random vector D. We assume that the components of random vector D are *mutually independent*, and refer to this as the *stagewise independence* condition. We now define the family of admissible policies Π by introducing two families of functions, $\{y_t,\, t = 1, \dots, T\}$ and $\{x_t,\, t = 1, \dots, T\}$. Conceptually, y_t will correspond to the inventory level at the start of stage t, and x_t will correspond to the inventory level after having ordered in stage t, but before the demand in that stage is realized.

We will consider policies which are nonanticipative, i.e. decisions do not depend on realizations of future demand. We assume that y_1, the initial inventory level, is a given constant. We also require that one can only order a nonnegative amount of inventory at each stage. Thus the set of admissible policies Π should consist of those vectors of (measurable) functions $\pi = \{x_t(d_{[t-1]}), t = 1,\dots,T\}$, such that $x_t : \mathbb{R}_+^{t-1} \to \mathbb{R}$ satisfies $x_t(d_{[t-1]}) \geq y_t$, for all $d_{[t-1]} \in \mathbb{R}_+^{t-1}$ and $t = 1,...,T$, where $y_{t+1} = x_t(d_{[t-1]}) - d_t$, $t = 1,...,T-1$.

To prevent notational confusion, it will at times be helpful to parametrize the relevant optimization problems by the initial condition y_1, and we let $\Pi(y_1)$ denote the relevant set of feasible policies for a given y_1. Sometimes we will explicitly express x_t and y_t as a function of the associated policy π and demands $d_{[t]}$, with the notations $x_t^\pi(d_{[t-1]})$ and $y_t^\pi(d_{[t-1]})$, where we note that under such a notation the fixed initial condition y_1 is implicit. Combining the above, we can write the classical multistage news vendor problem (inventory problem) as follows:

$$\inf_{\pi\in\Pi(y_1)} \mathbb{E}\left\{\sum_{t=1}^{T} \rho^{t-1}\left[c_t\left(x_t^\pi(D_{[t-1]}) - y_t^\pi(D_{[t-1]})\right) + \Psi_t\left(x_t^\pi(D_{[t-1]}), D_t\right)\right]\right\}. \tag{4}$$

Here $\rho \in (0,1]$ is a discount factor, c_t, b_t, h_t are the ordering, backorder penalty and holding costs per unit in stage t, respectively, and $\Psi_t(x_t, d_t) := b_t[d_t - x_t]_+ + h_t[x_t - d_t]_+$. Unless stated otherwise, we assume that $b_t > c_t > 0$ and $h_t \geq 0$ for all $t = 1,...,T$.

Problem (4) can be viewed as an optimal control problem in discrete time with state variables y_t, control variables x_t and random parameters D_t. It is well known that Problem (4) can be solved using DP equations, which can be written as

$$V_t(y_t) = \inf_{x_t \geq y_t} \left\{c_t(x_t - y_t) + \mathbb{E}\left[\Psi_t(x_t, D_t) + \rho V_{t+1}(x_t - D_t)\right]\right\}, \tag{5}$$

$t = 1,...,T$, with $V_{T+1}(\cdot) \equiv 0$ (e.g., Zipkin (2000)). Note that the value functions $V_t(\cdot)$ are convex, and do not depend on the demand data because of the stagewise independence assumption. These equations naturally define a set of policies through the relation $x_t(y_t) \in \mathfrak{X}_t(y_t)$, where $\mathfrak{X}_t(y_t)$, $t = 1,...,T$, is the set of optimal solutions of the problem

$$\inf_{x_t \geq y_t} \left\{c_t(x_t - y_t) + \mathbb{E}[\Psi_t(x_t, D_t) + \rho V_{t+1}(x_t - D_t)]\right\}, \tag{6}$$

and the optimal value of Problem (4) is given by $V_1(y_1)$. Note that $x_t(y_t)$, $t = 1,\dots,T$, are functions of y_t, i.e., it suffices to consider policies (measurable functions) of the form $x_t = \pi_t(y_t)$; this fact is well known from optimal control theory (see, e.g., Bertsekas et al. (1978) for technical details). Of course, the assumption of stagewise independence is essential for this conclusion.

Under the specified conditions, the objective function of Problem (6) tends to $+\infty$ as $x_t \to \pm\infty$. It thus follows from convexity that this objective function possesses a (possibly non-unique) unconstrained minimizer x_t^* over $x \in \mathbb{R}$, and $\bar{x}_t := \max\{y_t, x_t^*\}$ is an optimal solution of Problem (6). In particular, the so-called *base-stock policy* is optimal for the inventory Problem (4), where we note that such a result is classical in the inventory literature. Recall that a policy $\pi \in \Pi$ is said to be a *base-stock* policy if there exist constants x_t^*, $t = 1, \ldots, T$, such that $x_t^\pi = \max\left\{y_t^\pi, x_t^*\right\}$, $t = 1, \ldots, T$. Without loss of generality, we suppose that in any base-stock policy $x_1^* \geq y_1$. Problem (4) can be solved using the DP formulation (5) and associated policy (6) in the following sense.

Lemma 3.1 *The optimal value of Problem* (4) *equals* $V_1(y_1)$. *Any policy* π *such that* $x_t^\pi(d_{[t-1]}) \in \mathfrak{X}_t\big(y_t^\pi(d_{[t-1]})\big)$ *are for all* $t = 1, ..., T$ *and* $d_{[t-1]} \in \mathbb{R}_+^{t-1}$, *is an optimal solution to Problem* (4). *Conversely, for any optimal policy* π *for Problem* (4), *and any* $t \in \{1, ..., T\}$, *there exists a set* $A \subseteq \mathbb{R}$ *such that* $\mathsf{Pr}\big(y_t^\pi(D_{[t-1]}) \in A\big) = 1$, *and* $x_t^\pi(D_{[t-1]}) \in \mathfrak{X}_t\big(y_t^\pi(D_{[t-1]})\big)$ *conditional on the event* $\{y_t^\pi(D_{[t-1]}) \in A\}$. *Furthermore, it follows from the convexity of the relevant cost-to-go functions* $V_t(y_t)$ *that any set of base-stock constants* $\{x_t^*, t = 1, \ldots, T\}$ *such that* $x_t^* \in \mathfrak{X}_t(0)$ *for all* $t \in [1, T]$ *will yield an optimal policy for Problem* (4).

3.2. Distributionally robust formulations

Suppose now that the distribution of the demand process is not known, and we only have at our disposal information about the support and first and second order moments. In this case, it is natural to use the framework previously developed for the single-stage problem (see Section 2) to handle the distributional uncertainty at each stage. However, in the multistage setting, there is a nontrivial question of how to model the associated uncertainty in the joint distribution of demand.

We will consider two formulations, one corresponding to the modeling choices of a policy maker who does not recompute her policy choices after each stage and one corresponding to a policy-maker who does. These two formulations are analogous to the two optimization models discussed in Iyengar (2005) and Nilim et al. (2005) in the framework of robust MDP, and can also be interpreted through the lens of (non)rectangularity of the associated families of priors (cf. Epstein et al. (2003), Iyengar (2005), Nilim et al. (2005)). We refer to these formulations as *multistage-static* and distributionally robust DP, respectively. Questions regarding the interplay between the sets of optimal policies of these two formulations are important from an implementability perspective, as a policy deemed optimal at time 0, but which does not remain optimal if the relevant decisions are re-examined at a later time, may not be implemented by the relevant stake-holders. We note that the particular definitions and formulations we introduce here are by no means the only way to define the relevant notions of time consistency, and again refer the reader to the survey by Etner et al. (2012), and other recent papers in the optimization community (cf. Iyengar (2005), Boda et al. (2006), Carpentier et al. (2012), Iancu et al. (2015), Homem-de-Mello et al. (2016), Shapiro et al. (2019)) for alternative perspectives.

We suppose that we have been given a sequence of closed (possibly unbounded) intervals $\mathcal{I}_t = [\alpha_t, \beta_t] \subset \mathbb{R}$, means μ_t, and variances σ_t^2, $t = 1, \ldots, T$.

3.2.1. Multistage-static formulation. We first consider the following formulation, referred to as *multistage-static*, in which the policy maker does not recompute her policy choices. Let

$$\mathfrak{M}_t := \left\{ Q_t \in \mathfrak{P}(\mathcal{I}_t) : \mathbb{E}_{Q_t}[D_t] = \mu_t,\ \mathbb{E}_{Q_t}[D_t^2] = \mu_t^2 + \sigma_t^2 \right\},\ t = 1, ..., T; \tag{7}$$

$$\mathfrak{M} := \{ Q = Q_1 \times \cdots \times Q_T : Q_t \in \mathfrak{M}_t,\ t = 1, ..., T \}. \tag{8}$$

That is, the set $\mathfrak{M}$ consists of probability measures given by direct products of probability measures $Q_t \in \mathfrak{M}_t$. This can be viewed as an extension of the stage-wise independence condition, employed in Section 3.1, to the considered distributionally robust case. In order for the sets $\mathfrak{M}_t$ to be nonempty we assume that $\mu_t \in [\alpha_t, \beta_t]$ and $\sigma_t^2 \le (\beta_t - \mu_t)(\mu_t - \alpha_t)$, $t = 1, ..., T$. According to (8), the associated minimax problem supposes that although the set of associated marginal distributions may be "worst-case", the joint distribution will always be a product measure (i.e. the demand will be independent across stages). The multistage-static formulation for the distributionally robust inventory problem can then be formulated as follows.

$$\inf_{\pi \in \Pi(y_1)} \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q\big[Z^\pi\big], \tag{9}$$

where $Z^\pi = Z^\pi(D_{[T]})$ is a function of $D_{[T]} = (D_1, ..., D_T)$ given by

$$Z^\pi(D_{[T]}) := \sum_{t=1}^{T} \rho^{t-1} \Big[c_t\big(x_t^\pi(D_{[t-1]}) - y_t^\pi(D_{[t-1]})\big) + \Psi_t\big(x_t^\pi(D_{[t-1]}), D_t\big) \Big], \tag{10}$$

and $\Pi(y_1)$ is the set of admissible policies defined previously in Section 3.1. We let $\mathrm{OPT}(y_1)$ denote the set of optimal policies for Problem (9). Of course, if the set $\mathfrak{M} = \{Q\}$ is a singleton, then formulation (9) coincides with formulation (4) taken with respect to the distribution $Q = Q_1 \times \cdots \times Q_T$ of the demand vector $D_{[T]}$. We note that the multistage-static formulation (9) is closely related to optimization with risk measures. Indeed, the functional $\sup_{Q \in \mathfrak{M}} \mathbb{E}_Q\big[Z\big]$ is a coherent risk measure (cf. Shapiro et al. (2009)). Very little is known about $\mathrm{OPT}(y_1)$, as this problem does not enjoy a DP formulation along the lines of (5).

3.2.2. Time consistency and distributionally robust DP equations. As informally referenced earlier, time inconsistency refers to the possibility that policy choices which seemed optimal from the perspective of time 0 no longer seem optimal if one re-performs one's minimax calculations at a later time. Although first addressed within the economics community, the issue of time (in)consistency has recently started to receive attention in the stochastic and robust optimization communities, in which closely related concepts such as Pareto robust optimality (Iancu et al. (2014)) have also been studied. We note that related issues were addressed even in the seminal work of Bellman (1957) on DP. Here we will formalize two particular definitions of time-consistency in our setting, each of which has appeared in various forms throughout the literature. These two definitions will stem from two different ways to formalize the precise manner in which the decision maker "re-optimizes" at later time periods. We begin with Definition I, which leads to a tractable DP formulation, and reason as follows. Suppose we wished to know whether a policy π which was optimal for the multi-stage static formulation had the property that, should one re-perform one's minimax calculation in the final period, one would make the same ordering decision. As she cannot change past decisions, the only policy decision she still has to make is the determination of the function x_T. However, she now has knowledge of $D_{[T-1]}$ and y_T, which she can incorporate into her minimax computations. We note that here we are faced with the modeling question of how to reconcile the use of $D_{[T-1]}$ and y_T's realized values in performing one's minimax computations with the previously assumed stagewise independence of demand. A natural approach, consistent with the economics literature on time consistency, is to reason as follows. As $D_{[T-1]}$ has already been realized, it is unreasonable to enforce independence of D_T on this realization, as it is no longer undetermined. Instead, the relevant minimax computation is carried out with this knowledge of the realization of $D_{[T-1]}$. At that time, such a policy-maker is thus led to the optimization $\inf_{x_T \geq y_T} \left\{ c_T(x_T - y_T) + \sup_{Q_T \in \mathfrak{M}_T} \mathbb{E}_{Q_T}[\Psi_T(x_T, D_T)] \right\}$, with $\mathfrak{Y}_T(y_T) := \arg\min_{x_T \geq y_T} \left\{ c_T(x_T - y_T) + \sup_{Q_T \in \mathfrak{M}_T} \mathbb{E}_{Q_T}[\Psi_T(x_T, D_T)] \right\}$ the corresponding set of optimal policy choices. Here we note that (for example) the inner maximization $\sup_{Q_T \in \mathfrak{M}_T} \mathbb{E}_{Q_T}[\Psi_T(x_T, D_T)]$ is implicitly a function of $D_{[T-1]}$, through the dependence on x_T. Thus time-consistency of an optimal policy π for the multi-stage static formulation should (at least as regards policy decisions in this final period) be equivalent to requiring that $x^{\pi}_T(D_{[T-1]}) \in \mathfrak{Y}_T\big(y^{\pi}_T(D_{[T-1]})\big)$. We note that there is a second subtlety here relating to in what precise sense (e.g. with probability 1 etc.) this inclusion must hold, which we explain by considering a special situation. Suppose (for discussion) that the moment and support constraints ensured that there is only one possible measure Q for demand, with finite support. Note that the optimality of a policy π is completely independent from π's behavior on demand trajectories with zero probability under Q. Thus there will be optimal policies for the static formulation which behave arbitrarily on these zero probability trajectories, independent of the solution to any dynamic program. If we deemed such a policy inconsistent, it would : 1. stand against the fundamental notion that time-consistency assesses whether disagreement between a static and dynamic optimizer is possible (as disagreement would only occur on impossible trajectories); and 2. in many cases trivialize the question of whether all optimal policies for the static problem are consistent (our notion of strong consistency). Such questions become even more subtle when the uncertainty set contains measures with different supports, as one cannot simply refer to something holding w.p.1 under a common reference measure (a way some past works implicitly deal with this issue). We note that many past works do not take this subtlety into account in their definitions, as the distinction was less important for the notions of time consistency studied there (typically not strong consistency). Here, we propose the natural and intuitive interpretation / definition that one should require the inclusion hold w.p.1 for every measure in $\mathfrak{M}$, as these are exactly those measures one believes possible.

Distributionally robust DP formulation. Before proceeding with our formal definition of time consistency, let us expand on the distributionally robust DP formulation, which we have defined only in the final period. Carrying out the same logic using backwards induction, we conclude that if a policy is to be deemed time-consistent when the policy-maker is (possibly) given the choice to recompute her minimax calculations in an arbitrary set of time periods, her choices should be consistent with the following distributionally robust DP equations.

$$V_t(y_t) = \inf_{x_t \geq y_t} \left\{ c_t(x_t - y_t) + \sup_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\Psi_t(x_t, D_t) + \rho V_{t+1}(x_t - D_t)] \right\}, \tag{11}$$

$t = 1, ..., T$, with $V_{T+1}(\cdot) \equiv 0$. The optimal value of DP formulation (11) is given by $V_1(y_1)$. Dynamic equations (11) naturally define a set of policies of the form $x_t = \pi_t(y_t)$, $t = 1, \ldots, T$, with $x_t = \pi_t(y_t)$ being measurable selections $x_t \in \mathfrak{Y}_t(y_t)$ from sets

$$\mathfrak{Y}_t(y_t) := \arg\min_{x_t \geq y_t} \left\{ c_t(x_t - y_t) + \sup_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\Psi_t(x_t, D_t) + \rho V_{t+1}(x_t - D_t)] \right\}, \ t = 1, ..., T. \tag{12}$$

We refer to (11) as the *distributionally robust DP formulation* and $V_1(y_1)$ as its optimal value.

We now observe that due to certain convexity properties, DP formulation (11) always possesses an optimal base-stock policy. We note that such results are generally well-known to hold in this setting (cf. Ahmed et al. (2007)).

Observation 2 *It follows from the convexity of the relevant cost-to-go functions $V_t(y_t)$ that DP formulation* (11) *possesses an optimal base-stock policy. Furthermore, any set of base-stock constants $\{x_t^*, t = 1, \ldots, T\}$ such that $x_t^* \in \mathfrak{Y}_t(0)$ for all $t \in [1, T]$ yields an optimal policy. Namely, for any such $\{x_t^*, t = 1, \ldots, T\}$, $\max\{y, x_t^*\} \in \mathfrak{Y}_t(y)$ for all $y \in \mathbb{R}$ and $t = 1, \ldots, T$.*

We close our discussion of the distributionally robust DP formulation with a final definition, formalizing our earlier discussion of for which measures one should require optimality (of decisions) under the distributionally robust DP formulation.

Definition 3.1 (Robust-w.p.1-optimal) *Let us say that a policy $\pi \in \Pi(y_1)$ is robust-w.p.1-optimal for the distributionally robust DP formulation if for all $Q \in \mathcal{M}$, w.p.1 $x_t^\pi(D_{[t-1]}) \in \mathfrak{Y}_t\big(y_t^\pi(D_{[t-1]})\big)$ for all $t \in [1, T]$.*

When $\mathfrak{M}_t$ is the set of all probability measures supported on the interval $[\alpha_t, \beta_t]$ (i.e., the pure robust formulation), Definition 3.1 requires that the optimality equations hold for all possible $d_{[t-1]}$ consistent with the given intervals, because each singleton measure $Q_t \in \mathfrak{M}_t$ determines a point $d_t \in [\alpha_t, \beta_t]$. This remains true if one is also given first moment information, and under certain regularity conditions remains true for more general moment constraints (e.g., Schmüdgen 2017).

First formal definition of time consistency. We now provide our first formal definition of time consistency, in light of our earlier discussion. We note that given the motivation behind time consistency, i.e. implementation of policies, a further subtlety must be considered. Clearly, it is desirable for there to exist at least one policy which is optimal both initially, and if reconsidered at later times. However, it is similarly undesirable for there to exist even one policy which could potentially be selected (i.e. optimal) initially, but deemed sub-optimal (i.e. non-implementable) at a later time. Although such a notion is of course a stringent requirement (as noted informally in passing in Homem-de-Mello et al. (2016)), we believe that its conceptual importance none-the-less makes it worthy of further study. This motivates the following definition(s) of time consistency, where we note that similar definitions were presented in Grunwald et al. (2011) in a different context motivated by considerations in decision theory and artificial intelligence. Then our first definition of time consistency is as follows.

Definition 3.2 (Time consistency: Definition I) *If a policy $\pi \in OPT(y_1)$ is robust-w.p.1-optimal for the distributionally robust DP formulation, we say that π is type-I* time consistent*. If there exists at least one policy $\pi \in OPT(y_1)$ which is type-I time consistent, we say that Problem* (9) *is type-I* weakly time consistent*. If all $\pi \in OPT(y_1)$ are type-I time consistent, we say that Problem* (9) *is type-I* strongly time consistent*.*

Of course the notion of strong time consistency makes sense only if $\text{OPT}(y_1)$ is non-empty. We note that our Definition I of time consistency is essentially identical to definitions given in Shapiro (2009), Shapiro et al. (2016a, 2019), and closely related to the notion of dynamic consistency given in Pflug et al. (2014).

Second formal definition of time consistency. We now formalize our second definition of time-consistency, and reason as follows. Suppose we wished to know whether a policy π which was optimal for the multi-stage static formulation had the property that, should one re-perform one's minimax calculation in the *second* period, one would make the same ordering decision. From the perspective of the second period, one could reason as follows. With the demand in period 1 already realized, and given the posited stage-wise independence of demands, at time two one faces the "remaining" static problem, which we now formally define after introducing some relevant additional notations. For $i \in [1, T-1]$, let $\mathfrak{M}_t^i := \mathfrak{M}_{t+i}$ (for $t \in [1, T-i]$); $\mathfrak{M}^i := \{Q = Q_1 \times \cdots \times Q_{T-i} : Q_t \in \mathfrak{M}_t^i,\ t = 1, ..., T-i\}$; $b_t^i := b_{t+i}, c_t^i := c_{t+i}, h_t^i := h_{t+i}$ (for $t \in [1, T-i]$). Here we note that for each $i \in [1, T-1]$, the cost parameters $\{b_t^i, h_t^i, c_t^i, t \in [1, T-i]\}$ and distributional uncertainty set $\mathfrak{M}^i$ define a $(T-i)$-period multi-stage static problem, completely analogous to (9), which has its own set of feasible policies $\Pi^i(z)$ and optimal policies $\mathrm{OPT}^i(z)$ for any given initial condition z. Then time-consistency of an optimal policy $\pi \in \mathrm{OPT}(y_1)$ for the multi-stage static formulation should (at least as regards policy decisions in this second period) be equivalent to requiring that (w.p.1, under any $Q \in \mathfrak{M}$) there exists $\pi^1 \in \mathrm{OPT}^1\big(y_2^\pi(D_{[1]})\big)$ s.t. $x_1^{\pi^1} = x_2^\pi(D_{[1]})$. Carrying out the same logic using forwards induction, we conclude that if a policy is to be deemed time-consistent when the policy-maker is (possibly) given the choice to recompute her minimax calculations in an arbitrary set of time periods, her choices should satisfy the following.

Definition 3.3 (Time consistency: Definition II) *If a policy $\pi \in OPT(y_1)$, and if for every $i \in [1, T-1]$ and $Q \in \mathfrak{M}$ w.p.1 there exists $\pi^i \in OPT^i\big(y_{i+1}^\pi(D_{[i]})\big)$ s.t. $x_1^{\pi^i} = x_{i+1}^\pi(D_{[i]})$, we say that π is type-II* time consistent*. If there exists at least one policy $\pi \in OPT(y_1)$ which is type-II time consistent, we say that Problem* (9) *is type-II* weakly time consistent*. If all $\pi \in OPT(y_1)$ are type-II time consistent, we say that Problem* (9) *is type-II strongly time consistent.*

Our Definition II of time consistency is related to definitions given in Carpentier et al. (2012), Osogami et al. (2012), Pflug et al. (2014), Homem-de-Mello et al. (2016), and is essentially a weakening of the notion of "inherited optimality" defined in Homem-de-Mello et al. (2016). Interestingly, our two definitions arise depending on whether one uses backwards inductive reasoning (in which the re-optimizer is implicitly using knowledge that she will later re-optimize), or forwards inductive reasoning (in which no such assumption is made). At this time no single consensus on a proper definition of time consistency (especially as it pertains to optimization) has been reached in the literature, although several recent works have begun to unify various definitions appearing in the literature (cf. De Lara et al. (2016), Shapiro et al. (2016a), Gérard et al. (2019)). Both our definitions can, in an appropriate sense, be viewed as special cases of the definition given in Shapiro et al. (2016a).

Comparison between type-I and type-II time consistency. Our two notions of time consistency each have pros and cons, and may be appropriate in different circumstances. One benefit of type-I time consistency is that it quantifies agreement with a policy which is *tractable* (by the associated DP equations), while the multi-stage static problems associated with type-II time consistency are generally intractable (cf. Lam et al. (2013)). However, we note that the formal problem of testing time consistency appears intractable under both definitions, as doing so seems to require characterizing the optimal policies of the static problem. We note that progress has been made on related questions in Iancu et al. (2015), and that formulating simple sufficient conditions for time consistency (as we do in Section 4) can in some cases sidestep this challenge. We leave a deeper exploration of the computational nature of time consistency, and e.g. what definitions might be appropriate for a policy-maker with bounded computational resources, as an open question.

A potential downside of type-I time consistency is that it may be more conservative. Indeed, it is straightforward that our definitions imply the following.

Observation 3 (Relationship of type-I and type-II time consistency) *Suppose $T = 2$. Then a policy $\pi \in OPT(y_1)$ is type-II time consistent iff for all $Q \in \mathcal{M}$, w.p.1 $x_2^\pi(D_{[1]}) \in \mathfrak{Y}_2\big(y_2^\pi(D_{[1]})\big)$. It follows that for $T = 2$, the type-I time consistency of a policy $\pi \in OPT(y_1)$ also implies the type-II time consistency of π. Under the same assumptions, type-I weak (strong) time consistency thus implies type-II weak (strong) time consistency.*

For $T \geq 3$ the question of formal comparison becomes more subtle, and we leave developing a complete understanding as an interesting direction for future research.

Our definitions of time consistency can, in a certain sense, be viewed as an extension of the definition typically used in the theory of risk measures to an optimization context. In Section 4.3.3, we show that it is possible for the multistage-static problem to have an optimal solution and to be strongly time consistent, but with a different optimal value than the distributionally robust DP formulation. That is, it is possible for the multistage-static problem to possess an optimal solution and to be strongly time consistent even when the rectangularity property does not hold. This stands in contrast to the definition of consistency typically used in the theory of risk measures, i.e. the notion of dynamic consistency coming from Epstein et al. (2003) and based on a certain stability of preferences over time, which may result in a problem being deemed inconsistent based on the values that a given optimal policy takes under the different formulations, and even the values taken by suboptimal policies (cf. Ruszczyński (2010), Grunwald et al. (2011)). In an optimization setting one may be primarily concerned only with the implementability of optimal policies, irregardless of their values and the values of suboptimal policies, and this is the approach we take here. We note that such optimization-oriented formulations have been considered in several recent works, e.g. Carpentier et al. (2012) and Homem-de-Mello et al. (2016), and our definitions are largely consistent with those works. We also note that several other recent works have started to bridge these optimization-oriented definitions with definitions coming from the theory of risk measures (cf. De Lara et al. (2016), Gérard et al. (2019)).

Before further exploring some of the subtle and interesting features of time (in)consistency for our model, we briefly review some previously known results for related models. Note that if the set $\mathfrak{M}$ is a singleton, then both the multistage-static and distributionally robust DP formulations collapse to the classical formulation. Hence both formulations have the same optimal value and type-I and type-II strong time consistency follows. If one only has information about the support $\mathcal{I}_t$, and hence takes $\mathfrak{M}_t$ to be the set of all probability measures supported on the interval $\mathcal{I}_t$, $t = 1, ..., T$, then both the multistage-static and distributionally robust DP formulations collapse to the so-called adjustable robust formulation (cf. Ben-Tal et al. (2004), Shapiro (2011)), which is purely deterministic. As a consequence, both formulations have the same optimal value and type-I and type-II weak time consistency follows. However, the recent work of Shapiro et al. (2019) (itself inspired in part by an earlier version of this paper, Xin et al. (2013)), shows that even in that setting strong time consistency need not hold, and studies related phenomena in several particular problems with general moment constraints. We note that all inventory problems considered in Shapiro et al. (2019) are type-I weakly time consistent with both formulations having the same optimal value due to the rectangularity of the underlying set of measures. Shapiro et al. (2019) also shows that there are interesting connections between the notions of weak and strong time-consistency and the concept of "strict monotonicity" for risk measures (e.g., Shapiro (2017)), and we leave further investigations of this connection as an interesting direction for future research. We also note that the existence of optimal time inconsistent policies was investigated earlier in several purely robust (i.e. deterministic) settings. In particular, Bertsimas et al. (2010) demonstrated the optimality of so-called affine policies in certain settings, and Delage et al. (2015) explicitly constructed optimal time inconsistent policies in an inventory control setting.

4. Time consistency: sufficient conditions and (counter) examples

4.1. Sufficient conditions for weak time consistency

In this section, we provide simple sufficient conditions for the type-I and type-II weak time consistency of Problem (9), which are a direct corollary of an earlier result of Jagannathan (1978). The condition is essentially equivalent to monotonicity of the associated base-stock constants. Intuitively, in this case the inventory manager can always order up to the optimal inventory level with which to enter the next time period, irregardless of previously realized demand. Thus any potential for the adversary to take advantage of previously realized demand information in the distributionally robust DP formulation is "masked" by the fact that the actual attained inventory level will always be this idealized level, under both formulations. We note that several previous works have identified monotonicity of base-stock levels as a condition which causes various inventory problems to become tractable, in a variety of settings (cf. Veinott (1965), Ignall et al. (1969), Jagannathan (1978), Zipkin (2000)). In particular, Jagannathan (1978) studied a very similar distributionally robust inventory model with moment constraints and identified monotonicity of base-stock levels as a sufficient condition for a myopic base-stock policy to be optimal. We recall these results of

Jagannathan (1978), yielding a sufficient condition for weak time consistency, and for completeness include a proof in the supplemental appendix Section 7.

We begin by providing a different (but equivalent) formulation for Problem (9), in which all relevant instances of y_t are rewritten in terms of the appropriate x_t functions, as this will clarify the precise structure of the relevant cost-to-go functions. As a notational convenience, let $c_{T+1} = 0$, in which case we define $\hat{\Psi}_t(x_t, d_t) := (c_t - \rho c_{t+1})x_t + b_t[d_t - x_t]_+ + h_t[x_t - d_t]_+,\ t = 1, ..., T$. Let us define the problem

$$\inf_{\pi\in\Pi} \sup_{Q\in\mathfrak{M}} \mathbb{E}_Q \left[\sum_{t=1}^{T} \rho^{t-1}\hat{\Psi}_t\big(x_t(y_t), D_t\big)\right] - c_1 y_1 + \sum_{t=1}^{T-1} \rho^t c_{t+1}\mu_t. \tag{13}$$

Then, using straightforward substitution we can make the following observation.

Observation 4 *Problem* (9) *and Problem* (13) *are equivalent, i.e. each policy* $\pi \in \Pi(y_1)$ *has the same value under both formulations.*

We now derive a lower bound for any policy by allowing the policy maker to reselect her inventory at the start of each stage, at no cost. This will both set up the notation for formally stating the sufficient condition implied by the results of Jagannathan (1978), and also prove useful in some of our later proofs. As it turns out, this bound is "attainable" when the set of base-stock levels is monotone increasing. For $x \in \mathbb{R}$, let $\eta_t(x) := \sup_{Q_t\in\mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(x, D_t)]; \Gamma_t^x := \arg\max_{Q_t\in\mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(x, D_t)]; \hat{\eta}_t := \inf_{x\in\mathbb{R}} \eta_t(x) = \inf_{x\in\mathbb{R}} \sup_{Q_t\in\mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(x, D_t)]; \hat{\Gamma}_t := \arg\min_{x\in\mathbb{R}} \eta_t(x) = \arg\min_{x\in\mathbb{R}} \sup_{Q_t\in\mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(x, D_t)]$. For $j \geq 1$, and probability measures $Q_1, \ldots, Q_j$, let us define $\otimes_{t=1}^{j} Q_t := Q_1 \times \ldots \times Q_j$, i.e. the associated product measure with the corresponding marginals. Then we have the following, whose proof follows from a straightforward induction (which we include for completeness in the supplemental appendix Section 7.5).

Lemma 4.1 *Suppose that the sets* $\Gamma_t^x, \hat{\Gamma}_t$ *are non-empty for all* $x \in \mathbb{R}$, $t = 1, ..., T$. *Let us fix any* $\pi = (x_1, \ldots, x_T) \in \Pi$, *and* $i \geq 0$. *Then for any given* $Q_1 \in \mathfrak{M}_1, \ldots, Q_i \in \mathfrak{M}_i$, *there exist* $Q_{i+1} \in \mathfrak{M}_{i+1}, \ldots, Q_T \in \mathfrak{M}_T$ *such that*

$$\mathbb{E}_{\otimes_{j=1}^{T} Q_j}\left[\hat{\Psi}_t\big(x_t(y_t), D_t\big)\right] \geq \hat{\eta}_t \text{ for all } t \geq i+1. \tag{14}$$

Furthermore, the optimal value of Problem (9) *is at least* $\sum_{t=1}^{T} \rho^{t-1}\hat{\eta}_t - c_1 y_1 + \sum_{t=1}^{T-1} \rho^t c_{t+1}\mu_t$.

We now show that the bound of Lemma 4.1 is "realizable" when the set of base-stock levels is monotone increasing, and that in this case the associated base-stock policy is optimal for both the multistage-static and distributionally robust DP formulations. In particular, in this setting, the associated base-stock policy is type-I time consistent. Furthermore, as this condition is inherited by all of the subproblems appearing in the definition of type-II time consistency, type-II time consistency follows as well. Although their results are not stated exactly in this manner, this sufficient condition is essentially equivalent to an earlier result of Jagannathan (1978). For completeness, we again include a proof in the supplemental appendix Section 7.6.

Theorem 2 (Jagannathan (1978)) *Suppose there exists a nondecreasing sequence x_t^*, $t = 1,...,T$, such that $y_1 \le x_1^*$, and $x_t^* \in \hat{\Gamma}_t$, $t = 1,...,T$. Also suppose $\mathcal{I}_t \subset \mathbb{R}_+$ for all $t = 1,...,T$. Then the base-stock policy π for which $x_t(y) = \max\{y, x_t^*\}$ for all $y \in \mathbb{R}$, is an optimal policy for the multistage-static formulation, and a robust-w.p.1-optimal policy for the distributionally robust DP formulations, and attains value $\sum_{t=1}^{T} \rho^{t-1}\hat{\eta}_t - c_1 y_1 + \sum_{t=1}^{T-1} \rho^t c_{t+1}\mu_t$ under both formulations. Consequently, this base-stock policy is type-I time consistent, and the multistage-static problem is type-I weakly time consistent.*

We note that Theorem 2 implies that if the parameters $\mu_t, \sigma_t, c_t, b_t, h_t$ and $\mathcal{I}_t$ are the same for all $t = 1,...,T$, and hence the sets $\mathfrak{M}_t$ are also the same for all t, then the multistage-static problem is type-I weakly time consistent, and the multistage-static and distributionally robust DP formulations have the same optimal value. For further related discussion, we refer the interested reader to Jagannathan (1978).

We now argue that the assumptions of Theorem 2 imply type-II time consistency as well. Indeed, note that in the definition of type-II time consistency, for the ith subproblem (on $T-i$ periods) one may define $\hat{\Gamma}_t^i, t = 1,\ldots,T-i$ for the corresponding $(T-i)$-period multi-stage static problem, in analogy with the definition of $\hat{\Gamma}_t, t = 1,\ldots,T$ given above for the original T-period multi-stage static problem. It then follows from definitions that $\hat{\Gamma}_t^i = \hat{\Gamma}_{t+i}$ for all $i \in [1, T-1]$ and $t \in [1, T-i]$. Thus as subsequences of monotonic sequences are themselves monotone, the assumption that there exists a nondecreasing sequence x_t^*, $t = 1,...,T$, such that $y_1 \le x_1^*$, and $x_t^* \in \hat{\Gamma}_t$, $t = 1,...,T$ implies that for all $i \in [1, T-1]$, the nondecreasing sequence $x_{t+i}^*, t = 1,\ldots,T-i$ satisfies $x_{t+i}^* \in \hat{\Gamma}_t^i, t = 1,\ldots,T-i$. In addition, assuming $y_1 \le x_1^*$, under the base-stock policy π (for the original multistage-static problem) for which $x_t(y) = \max\{y, x_t^*\}$ for all $y \in \mathbb{R}, t = 1,\ldots,T$, it holds (w.p.1, for all $Q \in \mathfrak{M}$) that for all $i \in [1, T-1]$, $y_{i+1}^{\pi}(D_{[i]}) = x_i^* - D_i \le x_{i+1}^*$. Consequently, Theorem 2 implies that the base-stock policy π^i for which $x_t(y) = \max\{y, x_{t+i}^*\}$ for all $y \in \mathbb{R}, t = 1,\ldots,T-i$, is optimal for the ith multistage-static subproblem in the definition of type-II time consistency, i.e. $\pi^i \in \mathrm{OPT}^i\big(y_{i+1}^{\pi}(D_{[i]})\big)$. It follows that (supposing $y_1 \le x_1^*$), for the base-stock policy π, for all $Q \in \mathfrak{M}$, w.p.1 there exists $\pi^i \in \mathrm{OPT}^i\big(y_{i+1}^{\pi}(D_{[i]})\big)$ s.t. $x_1^{\pi^i} = x_{i+1}^{\pi}(D_{[i]}) = x_{i+1}^*$. Applying the definition of type-II time consistency, we conclude the following.

Corollary 4.1 *Under the same assumptions as Theorem 2, the same base-stock policy described in Theorem 2 is also type-II time consistent, and the multistage-static problem is type-II weakly time consistent.*

We end this section by briefly discussing an additional sufficient condition for type-I weak time consistency, which may shed further insight onto the relationship between type-I and type-II time consistency. Intuitively, this condition formalizes the fact that the essential property of Theorem 2 is that the multistage-static problem has an optimal policy π under which (for all t) $x_t^{\pi}(D_{[t-1]})$ is a constant independent of $D_{[t-1]}$. Note that (barring degenerate situations such as the demand equalling zero w.p.1) this is equivalent to π being a base-stock policy with monotone increasing base-stock levels.

Lemma 4.2 *Suppose that there exists $\pi \in OPT(y_1)$ which is a base-stock policy with monotone increasing base-stock levels. Then the multistage-static problem is type-I weakly time consistent, and the multistage-static and distributionally robust DP formulations have the same optimal value.*

We defer the proof of Lemma 4.2, and some further related discussion, to the supplemental appendix Section 7.7. We note that such an independence from the realized value of demand, although sufficient for type-I weak time consistency, is not strictly necessary, as we will later demonstrate through an example presented in Section 4.3.3. Although by Observation 3, Lemma 4.2 extends to type-II time consistency for the case T = 2, for $T \geq 3$ we leave the connection between such an independence from $D_{[t-1]}$ and type-II time consistency as an interesting direction for future research.

4.2. Sufficient conditions for strong time consistency

In this section, we show that under additional assumptions, which ensure that the variance in each stage is sufficiently large, the multistage-static problem is strongly time consistent. As we will see, in this case there is a unique optimal base-stock policy, and in this policy all base-stock constants equal zero, the intuition being that when the variance is sufficiently large, it becomes undesirable to give nature any additional "wiggle room". We further note that such a base-stock policy has been widely adopted in practice and the resulting inventory system is a so-called Make-To-Order (MTO) or "Pull" system. In such a system, no inventory is carried and the replenishment is based on actual demands instead of forecasts (cf. Williams (1984), Arreola-Risa et al. (1998), Federgruen et al. (1999), Rajagopalan (2002), Kaminsky et al. (2009)). We will later see in Section 4.3.2 that deviating slightly from this setting may lead to a lack of strong time consistency. In particular, our results demonstrate that strong time consistency is a very fragile property. Our sufficient conditions are as follows.

Theorem 3 *Suppose that* $b'_t := b_t - c_t + \rho c_{t+1} > 0$, $h'_t := h_t + c_t - \rho c_{t+1} > 0$, $\sigma_t, \mu_t > 0$, $\mathcal{I}_t = \mathbb{R}_+, t = 1, \ldots, T$, $y_1 = 0$, *and* $\frac{\sigma_t^2}{\mu_t^2} > \frac{b'_t}{h'_t}$, $t = 1, ..., T$. *Then the set of optimal policies for the multistage-static problem is exactly the set of policies* $\Pi^0 := \{\pi = (x_1, \ldots, x_T) \in \Pi : x_1(y_1) = 0, x_t(z) = 0 \ \ \text{for all} \ \ z \leq 0 \ \ \text{and} \ \ t \in [1, T]\}$, *and the multistage-static problem is type-I strongly time consistent.*

We defer the proof to the technical appendix Section 6.2. By the same logic used to extend Theorem 2 to type-II time consistency and derive Corollary 4.1, we may analogously derive the following corollary.

Corollary 4.2 *Under the same assumptions as Theorem 3, the multistage-static problem is type-II strongly time consistent.*

We note that under certain rectangularity-related assumptions, necessary and sufficient conditions for the existence of time-inconsistent optimal policies in an inventory setting with moment constraints was very recently provided in Shapiro et al. (2019). However, those results are not applicable to the setting we consider, as the uncertainty sets we consider here are inherently non-rectangular, and thus (for example) our formulation allows for the non-existence of weak time-consistency, as well as the two formulations having different optimal values, and even the possibility that no policy of base-stock form is optimal for the multistage-static formulation (while the assumptions of Shapiro et al. (2019) do not allow for such behavior). Furthermore, Shapiro et al. (2019) considers only 2-period problems, while the sufficient conditions provided in this work hold in the general multi-period setting.

4.3. Further investigation of time (in)consistency

We now demonstrate that the question of time (in)consistency becomes quite delicate for inventory models with moment constraints, by considering a series of examples in which our model exhibits interesting (and sometimes counterintuitive) behavior. We view our results as a step towards understanding the subtleties which can arise when taking a policy-centric view of time consistency in an operations management setting. Throughout this section, we will let Π_s^{opt} denote the set of all optimal policies for the corresponding multistage-static problem, and Π_d^{opt} denote the set of all robust-w.p.1-optimal policies for the corresponding distributionally robust DP problem. We defer all proofs to the technical appendix, Sections 6.3 - 6.6.

4.3.1. Example: a multistage-static problem that is not weakly time consistent.

In this section, we explicitly provide an example for which the multistage-static problem is not weakly time consistent. Furthermore, for this example, the multistage-static and distributionally robust DP formulations have different optimal values. This stands in contradiction to the robust optimization setting (cf. Ben-Tal et al. (2004)) and the first-moment-constraint setting (cf. Shapiro (2012)) where weak time consistency always holds and the two formulations always have the same optimal value. Let

$$y_1 = 10 \quad , \quad \rho = 1 \quad , \quad \mathcal{I}_1 = [1,3], \quad \mu_1 = 2, \quad \sigma_1 = 1, \quad c_1 = 0, \quad b_1 = 2, \quad h_1 = 2,$$

$$\mathcal{I}_2 = \mathbb{R}_+, \quad \mu_2 = 8, \quad \sigma_2 = 2, \quad c_2 = 0, \quad b_2 = 1, \quad h_2 = 1.$$

Let $\tilde{\Pi}_s$ denote the set of policies $\tilde{\pi} = (\tilde{x}_1, \tilde{x}_2)$ such that $\tilde{x}_1(10) = 10$, $\tilde{x}_2(9) = 9$, $\tilde{x}_2(7) = 7$, and $\tilde{\Pi}_d$ denote the set of policies $\tilde{\pi} = (\tilde{x}_1, \tilde{x}_2)$ such that $\tilde{x}_1(10) = 10$, $\tilde{x}_2(9) = 9$, $\tilde{x}_2(7) = 8$. Note that here (and in later statements) we have defined a set of policies by specifying a required behavior at only a few values, and allow the behavior at all other values to be arbitrary (subject to the overall policy belonging to $\Pi(y_1)$).

Observation 5 $\Pi_s^{opt} = \tilde{\Pi}_s$*, and the optimal value of the multistage-static problem is 18. On the other hand,* $\Pi_d^{opt} \subseteq \tilde{\Pi}_d$*, and the optimal value of the distributionally robust DP problem is* $17 + \frac{\sqrt{5}}{2} > 18$*. Consequently, the multistage-static problem is not type-I or type-II weakly time consistent, and the multistage-static and distributionally robust DP problems have different optimal values.*

4.3.2. Example: a multistage-static problem that is weakly time consistent, but *not* strongly time consistent. In this section, we explicitly provide an example showing that it is possible for the multistage-static problem to be weakly time consistent, but not strongly time consistent. In particular, there is a base-stock policy π^*, with associated base-stock constants x_1^*, x_2^* satisfying the conditions of Theorem 2, which is optimal for the multi-stage static formulation and robust-w.p.1-optimal for the distributionally robust DP formulation, yet the multistage-static problem has other non-trivial optimal policies which are not robust-w.p.1-optimal for the distributionally robust DP formulation. The intuitive explanation is as follows. In the multistage-static formulation, one can leverage the randomness in the realization of D_1 to construct a policy π' such that with positive probability $x_2^{\pi'}(y_2)$ is slightly below x_2^*, and with the remaining probability is slightly above x_2^*. Since in the multistage-static formulation nature cannot observe the realized inventory in stage 2 before selecting a worst-case distribution, it turns out that such a policy incurs the same cost as π' under the multistage-static formulation. However, under the distributionally robust DP formulation, such a perturbation leads to sub-optimality. We note that such a lack of strong time consistency can also be interpreted as resulting from the fact that optimality of a policy for the static formulation does not require optimality for every possible measure which nature can select, analogous to the ideas explored (in the robust optimization setting) in Iancu et al. (2014). We note that in this example, even though the multistage-static problem is not strongly time consistent, both formulations have the same optimal value, as dictated by Theorem 2. Let

$$y_1 = 0 \quad , \quad \rho = 1 \quad , \quad \mathcal{I}_1 = [1,3], \quad \mu_1 = 2, \quad \sigma_1 = 1, \quad c_1 = 0, \quad b_1 = 1, \quad h_1 = 1,$$

$$\mathcal{I}_2 = \mathbb{R}_+, \quad \mu_2 = 10, \quad \sigma_2 = 1, \quad c_2 = 0, \quad b_2 = 1, \quad h_2 = 1.$$

Observation 6 *The multistage-static problem is type-I and type-II weakly time consistent, but not type-I or type-II strongly time consistent.*

4.3.3. Example: a multistage-static problem that is strongly time consistent, but the two formulations have a *different* optimal value. In this section, we explicitly provide an example showing that it is possible for the multistage-static problem to be strongly time consistent, yet for the two formulations to have different optimal values. We note that, although it is expected that there will be settings where the two formulations have different optimal values, it is somewhat surprising that this is possible even when the two formulations have the same set of optimal policies. As discussed previously, we note that this possibility stands in contrast to several related works which consider alternative, less policy-focused definitions of time consistency, e.g. those definitions appearing in the literature on risk measures, and shows that consistency may hold even when rectangularity does not. Let

$$y_1 = 0 \quad , \quad \rho = 1 \quad , \quad \mathcal{I}_1 = [1,3], \quad \mu_1 = 2, \quad \sigma_1 = 1, \quad c_1 = 0, \quad b_1 = 0, \quad h_1 = 0,$$

$$\mathcal{I}_2 = \mathbb{R}_+, \quad \mu_2 = 100, \quad \sigma_2 = 5, \quad c_2 = 2, \quad b_2 = 1, \quad h_2 = 1.$$

Let $\tilde{\Pi}$ denote the set of policies $\tilde{\pi} = (\tilde{x}_1, \tilde{x}_2)$ such that $\tilde{x}_1(0) = 102$, $\tilde{x}_2(101) = 101$, $\tilde{x}_2(99) = 99$.

Observation 7 $\Pi_s^{opt} = \tilde{\Pi}$, *and the multistage-static problem is type-I and type-II strongly time consistent. However, the optimal value of the multistage-static problem equals 5, while the optimal value of the distributionally robust DP problem equals* $\sqrt{26} > 5$.

4.3.4. Example: a multistage-static problem that has no optimal policy of base-stock form. In this section, we explicitly provide an example showing that it is possible for the multistage-static problem to have no optimal base-stock policy, where we note that in all our previous examples the associated multistage-static problem did indeed have an optimal base-stock policy (possibly different from that of the associated distributionally robust DP problem). Note that this stands in contrast to the distributionally robust DP formulation, which always has an optimal base-stock policy by Observation 2. It remains an interesting open question to develop a deeper understanding of the set of optimal policies for the multistage-static problem, where we again note that some preliminary investigations of such distributionally robust problems with independence constraints can be found in Lam et al. (2013). Both the results of Lam et al. (2013), and our own result, indicate that the structure of optimal policies for the multistage-static problem may be very complicated.

To prove the desired result, it will be useful to consider a family of problems parameterized by a parameter ϵ. In particular, let $\epsilon \in \big(0, \frac{1}{2}(\sqrt{6}-2)\big)$ be any sufficiently small strictly positive number. It may be easily verified that for any such ϵ, one has $\epsilon \in (0, \frac{1}{4})$,

$$\frac{1}{2} - 2\epsilon - \epsilon^2 > 0, \tag{15}$$

$$y_1 = 10 - \epsilon \quad , \quad \rho = 1 \quad , \quad \mathcal{I}_1 = [1-\epsilon,\ 3+\epsilon], \quad \mu_1 = 2, \quad \sigma_1 = 1, \quad c_1 = 0, \quad b_1 = 2, \quad h_1 = 2,$$

$$\mathcal{I}_2 = \mathbb{R}_+, \quad \mu_2 = 8, \quad \sigma_2 = 3, \quad c_2 = 0, \quad b_2 = 1, \quad h_2 = 1.$$

Observation 8 *Suppose ϵ satisfies* (15)*. Then any admissible policy $\tilde{\pi} = (\tilde{x}_1, \tilde{x}_2) \in \Pi$ satisfying $\tilde{x}_1(y_1) = y_1$, $\tilde{x}_2(D_1) = y_1 - D_1 + \epsilon$ belongs to Π_s^{opt}, and the corresponding optimal value equals $19 - 2\epsilon$. Moreover, no base-stock policy belongs to Π_s^{opt}.*

5. Conclusion

In this paper, we analyzed the notion of time consistency in the context of managing an inventory under distributional uncertainty. In particular, we studied the associated multistage distributionally robust optimization problem, when only the mean, variance and distribution support are known for the demand at each stage. Our contributions were three-fold. First, we gave two refined policy-centric definitions for time consistency in this setting, and put our definitions in the context of prior work. Second, we gave sufficient conditions for weak and strong time consistency. Third, we gave a series of examples of two-stage problems exhibiting interesting and counterintuitive time (in)consistency properties, showing that the question of time consistency can be quite subtle in this setting.

We departed from much of the past literature by demonstrating both negative *and positive* results regarding time consistency when the underlying family of distributions from which nature can select is non-rectangular, a setting in which most of the literature focuses on demonstrating hardness of the underlying optimization problems and other negative results. Furthermore, our example demonstrating that it is possible for the multistage-static problem to be strongly time consistent, but with a different optimal value than the distributionally robust DP formulation, stands in contrast to the definition of time consistency typically used in the theory of risk measures, i.e. the notion of dynamic consistency coming from Epstein et al. (2003), under which a problem may be deemed time inconsistent based on the values that a given optimal policy takes under the different formulations, and even the values taken by suboptimal policies. Indeed, our definitions are motivated by the fact that in an optimization setting, one may be primarily concerned only with the implementability of optimal policies, irregardless of their values and the values of suboptimal policies, building on the more optimization-oriented definitions provided in Carpentier et al. (2012) and the recent work Homem-de-Mello et al. (2016).

Our work leaves many interesting directions for future research. The general question of time consistency, and how its many definitions appearing in the literature relate to one-another, remains poorly understood. It is also an intriguing question to understand the accuracy of time-consistent approximations for time-inconsistent problems, e.g. how much our two formulations can differ in optimal value and policy, even when time inconsistency occurs, along the lines of Huang et al. (2011), Asamov et al. (2015), and Iancu et al. (2015). On a related note, it is largely open to develop a broader understanding of the optimal solution to the multistage-static problem, or even approximately optimal solutions, as well as related algorithms, where we note that preliminary investigations along these lines were carried out in Lam et al. (2013).

References

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

Arreola-Risa, A., G.A. DeCroix. 1998. Make-to-order versus make-to-stock in a production-inventory system with general production times. *IIE Transactions.* **30**(8) 705-713.

Artzner, P., F. Delbaen, J.M. Eber, D. Heath, H. Ku. 2007. Coherent multiperiod risk adjusted values and Bellmans principle. *Annals of Operations Research.* **152** 5–22.

Asamov, T., A. Ruszczyński. 2015. Time-consistent approximations of risk-averse multistage stochastic optimization problems. *Mathematical Programming.* **153**(2) 459-493.

Bellman, R.E. 1957. *Dynamic Programming.* Princeton University Press, Princeton, NJ.

Ben-Tal, A., B. Golany, A. Nemirovski, J.P. Vial. 2005. Retailer-supplier flexible commitments contracts: a robust optimization approach. *Manufacturing & Service Operations Management* **7** 248–271.

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

Ben-Tal, A., G. Boaz, S. Shimrit. 2009. Robust multi-echelon multi-period inventory control. *European Journal of Operational Research.* **199**(3) 922–935.

Bertsekas, D.P., S.E. Shreve. 1978. *Stochastic Optimal Control: The Discrete Time Case.* Academic Press, NY.

Bertsimas, D., D. A. Iancu, P.A. Parrilo. 2010. Optimality of affine policies in multistage robust optimization. *Mathematics of Operations Research.* **35**(2) 363-394.

Bertsimas, D., I. Popescu. 2005. Optimal inequalities in probability theory: A convex optimization approach. *SIAM Journal on Optimization.* **15**(3) 780-804.

Bertsimas, D., A. Thiele. 2006. A robust optimization approach to inventory theory. *Operations Research.* **54**(1) 150-168.

Bielecki, T. R., I. Cialenco, M. Pitera. 2018. A unified approach to time consistency of dynamic risk measures and dynamic performance measures in discrete time. *Mathematics of Operations Research.* **43**(1) 204-221.

Boda, K., J.A. Filar. 2006. Time consistent dynamic risk measures. *Mathematical Methods in Operations Research.* **63** 169–186.

Bonnans, J. F., A. Shapiro. 2000. *Perturbation Analysis of Optimization Problems.* Springer-Verlag, NY.

Carpentier, P., J.P. Chancelier, G. Cohen, M. De Lara, P. Girardeau. 2012. Dynamic consistency for stochastic optimal control problems. *Annals of Operations Research.* **200** 247–263.

Carrizosaa, E., A. Olivares-Nadal, P. Ramirez-Cobob. 2016. Robust newsvendor problem with autoregressive demand. *Computers & Operations Research.* **68** 123-133.

Chen, Z., G. Li, J. Guo. 2013. Optimal investment policy in the time consistent mean-variance formulation. *Insurance: Mathematics and Economics.* **52** 145–156.

Chen, X., M. Sim, D. Simchi-Levi, P. Sun. 2007. Risk aversion in inventory management. *Operations Research.* **55**(5) 828-842.

Chen, W., M. Sim. 2009. Goal-driven optimization. *Operations Research.* **57**(2) 342-357.

Chen, X., P. Sun. 2012. Optimal structural policies for ambiguity and risk averse inventory and pricing models. *SIAM Journal on Control and Optimization.* **50**(1) 133-146.

Cheridito, P., M. Kupper. 2009. Recursiveness of indifference prices and translationinvariant preferences. *Mathematics and Financial Economics.* **2** 173–188.

Choi, S., A. Ruszczynski. 2008. A risk-averse newsvendor with law invariant coherent measures of risk. *Operations Research Letters.* **36**(1) 77–82.

Choi, S., A. Ruszczyński, Y. Zhao. 2011. A multiproduct risk-averse newswendor with law-invariant coherent measures of risk. *Operations Research.* **59** 346–354.

De Lara, M., V. Leclère. 2016. Building up time-consistency for risk measures and dynamic optimization. *European Journal of Operational Research.* **249**(1) 177-187.

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

Edgeworth, F. 1888. The mathematical theory of banking. *Royal Statistical Society.* **51** 113–127.

Epstein, L. G., M. Schneider. 2003. Recursive multiple-priors. *Journal of Economic Theory.* **113**(1) 1-31.

Etner, J., M. Jeleva, J.-M. Tallon. 2012. Decision theory under ambiguity. *Journal of Economic Surveys.* **26**(2) 234-270.

Federgruen, A., Z. Katalan. 1999. The impact of adding a maketo-order item to a make-to-stock production system. *Management Science.* **45**(7) 980-994.

Gabrel, V., C. Murat, A. Thiele. 2014. Recent advances in robust optimization: an overview. *European Journal of Operational Research.* **235**(3) 471-483.

Gallego, G. 1998. New bounds and heuristics for (Q, r) policies. *Management Science.* **44**(2) 219–233.

Gallego, G. 2001. Minimax analysis for finite-horizon inventory models. *IIE Transactions.* **33**(10) 861–874.

Gallego, G., K. Katircioglu, B. Ramachandran. 2007. Inventory management under highly uncertain demand. *Operations Research Letters.* **35**(3) 281-289.

Gallego, G., I. Moon. 1993. The distribution free newsboy problem: review and extensions. *Journal of the Operational Research Society.* **44**(8) 825–834.

Gallego, G., I. Moon. 1994. Distribution free procedures for some inventory models. *Journal of the Operational research Society.* 651–658.

Gérard, H., M. De Lara, Jean-Philippe Chancelier. 2019. Equivalence between time consistency and nested formula. *Annals of Operations Research.* 1-21.

Grunwald, P. D., J. Y. Halpern. 2011. Making decisions using sets of probabilities: updating, time consistency, and calibration. *Journal of Artificial Intelligence Research* **42**(1) 393-426.

Hanasusanto, G., A. Grani, D. Kuhn, W. Wallace, S. Zymler. 2015. Distributionally robust multi-item newsvendor problems with multimodal demand distributions. *Mathematical Programming.* **152**(1) 1-32.

Hansen, L. P., T. J. Sargent. 2001. Robust control and model uncertainty. *American Economic Review.* **91**(2) 60-66.

Homem-de-Mello, T., B. Pagnoncelli. 2016. Risk aversion in multistage stochastic programming: A modeling and algorithmic perspective. *European Journal of Operational Research.* **249**(1) 188-199.

Huang, P., D. Iancu, M. Petrik, D. Subramanian. 2011. The Price of Dynamic Inconsistency for Distortion Risk Measures. Technical Report.

Iancu, D.A., M. Petrik, D. Subramanian. 2015. Tight Approximations of Dynamic Risk Measures. *Mathematics of Operations Research.* **40**(3) 655-682.

Ignall, E., A. Veinott. 1969. Optimality of myopic inventory policies for several substitute products. *Management Science.* **15**(5) 284-304.

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

Isii, K., 1962. On sharpness of Tchebycheff-type inequalities. *Ann. Inst. Stat. Math.*, **14** 185–197.

Iyengar, G.N.. 2005. Robust Dynamic Programming. *Mathematics of Operations Research*. **30** 257–280.

Jagannathan, R. 1978. A minimax ordering policy for the infinite stage dynamic inventory problem. *Management Science*. **24** 1138-1149.

Kaminsky, P., O. Kaya. 2009. Combined make-to-order/make-to-stock supply chains. *IIE Transactions*. **41**(2) 103-119.

Kasugai, H., T. Kasegai. 1961. Note on minimax regret ordering policy-static and dynamic solutions and a comparison to maximin policy. *Journal of the Operations Research Society of Japan*. **3**(4) 155–169.

Klabjan, D., D. Simchi-Levi, M. Song. 2013. Robust stochastic lot-sizing by means of histograms. *Production and Operations Management*. **22** 691-710.

Lam, H., S. Ghosh. 2013. Iterative methods for robust estimation under bivariate distributional uncertainty. *Proceedings of the Winter Simulation Conference*.

Moments in mathematics, H.J. Landau (ed.), *Proc. Sympos. Appl. Math.*, 37, Amer. Math. Soc., Providence, RI, 1987.

Levi, R., G. Perakis, J. Uichanco. 2015. The data-driven newsvendor problem: new bounds and insights. *Operations Research*. **63**(6) 1294-1306.

Lovejoy, W. 1992. Stopped myopic policies for some inventory models with uncertain demand distributions. *Management Science*. **38**(5) 688-707.

Natarajan, K., L. Zhou. 2007. A mean-variance bound for a three-piece linear function. *Probability in the Engineering and Informational Sciences*. **21**(4) 611-621.

Nilim, A., L. El Ghaoui. 2005. Robust Control of Markov Decision Processes with Uncertain Transition Matrices. *Operations Research*. **53** 780–798.

Osogami, T., T. Morimura. 2012. Time-consistency of optimization problems. *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*.

Perakis, G., G. Roels. 2008. Regret in the newsvendor model with partial information. *Operations Research*. **56**(1) 188-203.

Pflug, G. C., A. Pichler. 2014. Time Consistency, pp. 175-208, Chapter 5 in *Multistage Stochastic Optimization*, Springer.

Popescu, I., 2005. A Semidefinite Programming Approach to Optimal Moment Bounds for Convex Classes of Distributions. *Mathematics of Operations Research*. **50**(3) 632-657.

Rajagopalan, S. 2002. Make-to-order or make-to-stock: model and application. *Management Science*. **48**(2) 241-256.

Riedel, F. 2004. Dynamic coherent risk measures. *Stochastic Processes and their Applications*. **112** 185–200.

Roorda, B., J.M. Schumacher. 2007. Time consistency conditions for acceptability measures, with an applications to Tail Value at Risk. *Insurance: Mathematics and Economics*. **40** 209–230.

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

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

Scarf, H. 1958. A min-max solution of an inventory problem. K. Arrow, ed *Studies in the Mathematical Theory of Inventory and Production.* Stanford University Press, Stanford, CA, 201–209.

Scarf, H. 1959. Bayes solution of the statistical inventory problem. *The Annals of Mathematical Statistics.* 490–508.

Scarf, H. 1960. Some remarks on Bayes solutions to the inventory problem. *Naval Research Logistics Quarterly.* **7**(4) 591–596.

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

See, C. T., M. Sim. 2010. Robust approximation to multiperiod inventory management. *Operations Research.* **58**(3) 583-594.

Shapiro, A., D. Dentcheva, A. Ruszczyński. 2009. *Lectures on Stochastic Programming: Modeling and Theory.* SIAM, Philadelphia.

Shapiro, A. 2009. On a time consistency concept in risk averse multistage stochastic programming. *Operations Research Letters.* **37** 143–147.

Shapiro, A. 2011. A dynamic programming approach to adjustable robust optimization. *Operations Research Letters.* **39**(2) 83-87.

Shapiro, A. 2012. Minimax and risk averse multistage stochastic programming. *European Journal of Operational Research.* **219** 719–726.

Shapiro, A. 2016. Rectangular sets of probability measures. *Operations Research.* **64**(2) 528-541.

Shapiro, A. 2017. Interchangeability principle and dynamic equations in risk averse stochastic programming. *Operations Research Letters.* **45**(4) 377-381.

Shapiro, A., A. Pichler. 2016a. Time and dynamic consistency of risk averse stochastic programs. Published electronically in Optimization Online.

Shapiro, A., K. Ugurlu. 2016b. Decomposability and time consistency of risk averse multistage programs. *Operations Research Letters.* **44**(5) 663-665.

Shapiro, A., L. Xin. 2019. Time inconsistency of optimal policies of distributionally robust inventory models. *Operations Research.* Forthcoming.

Strotz, R. H.. 1955. Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies.* **23**(3) 165-180.

Veinott, Jr., A. F.. 1965. Optimal policy for a multi-product, dynamic, nonstationary inventory problem. *Management Science.* **12**(3) 206-222.

Wang, T. 1999. A class of dynamic risk measure. *Working Paper*. University of British Columbia.

Wiesemann, W., D. Kuhn, B. Rustem. 2013. Robust Markov decision processes. *Mathematics of Operations Research*. **38**(1) 153-183.

Williams, T.M. 1984. Special products and uncertainty in production/inventory systems. *European Journal of Operations Research*. **15** 46-54.

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

Yang, J. 2013. Inventory and Price Control under Time-consistent Coherent and Markov Risk Measure. Preprint.

Yue, J., B. Chen, M. Wang. 2006. Expected value of distribution information for the newsvendor problem. *Operations Research*. **54**(6) 1128–1136.

Zhu, Z., J. Zhang, Y. Ye. 2013. Newsvendor optimization with limited distribution information. *Optimization methods and software*. **28**(3) 640-667.

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

6. Technical Appendix

6.1. Extensions of Scarf (1958)

For use in later proofs, we now state two extensions of the results of Scarf (1958). First, consider a generalization of Problem (1) in which one can select any distribution D_1 for x. Specifically, let us consider the following minimax problem:

$$\inf_{Q_1 \in \mathfrak{P}(\mathcal{I})} \phi(Q_1), \tag{16}$$

where $\phi(Q_1) := \sup_{Q_2 \in \mathfrak{M}} \mathbb{E}_{Q_1 \times Q_2}\big[\Psi(D_1, D_2)\big]$, and the notation $\mathbb{E}_{Q_1 \times Q_2}$ indicates that for any choices for the marginal distributions Q_1, Q_2 of D_1 and D_2, the expectation is taken with respect to the associated product measure, under which D_1 and D_2 are independent. In this case, we have the following result, whose proof follows from Theorem 1 and a straightforward probabilistic argument, and which we include (for completeness) in the supplemental appendix Section 7.3.

Proposition 6.1 *Suppose that $b > c$, $c + h > 0$, $\mu > 0$, $\sigma > 0$, $\frac{\sigma^2}{\mu^2} > \frac{b-c}{h+c}$, and $\mathcal{I} = \mathbb{R}$. Then Problem (16) has the unique optimal solution $\bar{Q}_1 = \delta_0$.*

Next, we state a generalization presented in Natarajan et al. (2007) to certain piecewise affine functions with three pieces. For completeness, we provide a proof in the supplemental appendix Section 7.4.

Theorem 4 (Natarajan et al. (2007)) *Suppose that there exist $c_1, c_2 > 0$ such that $c_1 < c_2$, and $\zeta(d) = \max\{-d + c_1, 0, d - c_2\}$ for all $d \in \mathbb{R}$. Let $\eta := \frac{1}{2}(c_1 + c_2)$, and recall that $\mathfrak{f}(z) := \big((z-\mu)^2 + \sigma^2\big)^{\frac{1}{2}}$. Further suppose that $\sigma > 0$, $\mathcal{I} = \mathbb{R}_+$, $\frac{1}{4}(2\mu - 3c_1 + c_2)(3c_2 - c_1 - 2\mu) \le \sigma^2$, and $\eta - \mathfrak{f}(\eta) \ge 0$. Then the unique optimal solution to the primal Problem (2) is the probability measure Q having support at two points $h_1 = \eta - \mathfrak{f}(\eta)$ and $h_2 = \eta + \mathfrak{f}(\eta)$, with $Q(h_1) = \sigma^2 \left(\sigma^2 + \big(\eta - \mathfrak{f}(\eta) - \mu\big)^2\right)^{-1}$, and $Q(h_2) = 1 - Q(h_1)$.*

6.2. Proof of Theorem 3

Proof of Theorem 3 : Let Π^{opt} denote the set of optimal policies for the multistage-static problem. It follows from Theorem 1.(i) and Theorem 2 that $\Pi^0 \subseteq \Pi^{opt}$, and every policy $\pi \in \Pi^0$ is time consistent. Thus to prove the theorem, it suffices to demonstrate that $\Pi^0 = \Pi^{opt}$, and we begin by showing that $\bar{\pi} = (\bar{x}_1, \dots, \bar{x}_T) \in \Pi^{opt}$ implies $\bar{x}_1(y_1) = 0$. Indeed, it follows from Lemma 4.1 that $\bar{\pi} \in \Pi^{opt}$ implies $\sup_{Q \in \mathfrak{M}_1} \mathbb{E}_Q\big[\hat{\Psi}_1\big(\bar{x}_1(y_1), D_1\big)\big] = \hat{\eta}_1 = b_1\mu_1$. That $\bar{x}_1(y_1)$ must equal 0 then follows from Theorem 1.

We now show that $\bar{\pi} \in \Pi^{opt}$ implies $\bar{x}_2(z) = 0$ for all $z \le 0$. We proceed by contradiction. Suppose that there exists $z' \le 0$ such that $\bar{x}_2(z') \ne 0$. It is easily verified that there exists $Q_1 \in \mathfrak{M}_1$ such that $Q_1(-z') > 0$, and consequently for this choice of Q_1, $\bar{x}_2(y_2)$ is not a.s. equal to 0. We conclude from Proposition 6.1 that there exists $Q_2 \in \mathfrak{M}_2$ such that $\mathbb{E}_{Q_1 \times Q_2}\big[\hat{\Psi}_2\big(\bar{x}_2(y_2), D_2\big)\big] > \hat{\eta}_2 = b_2\mu_2$. As we have already demonstrated that $\bar{x}_1(y_1) = 0$, and $Q_1 \in \mathfrak{M}_1$, we conclude that $\mathbb{E}_{Q_1}\big[\hat{\Psi}_1\big(\bar{x}_1(y_1, D_1\big)\big] = \hat{\eta}_1 = b_1\mu_1$. Combining with Lemma 4.1 then yields a contradiction. The proof that $\bar{x}_t(z) = 0$ for all $z \le 0$ and $t \ge 3$ follows from a nearly identical argument, and we omit the details. □

6.3. Proof of Observation 5

We first characterize the set of optimal policies for the multistage-static problem.

Lemma 6.1 $\Pi_s^{opt} = \tilde{\Pi}_s$*, and the multistage-static problem has optimal value 18.*

Proof : It follows from Observation 1 that $\mathfrak{M}_1$ consists of the single probability measure Q_1 such that $Q_1(1) = Q_1(3) = \frac{1}{2}$. Let D_1 denote a random variable distributed as Q_1. Note that for any policy $\pi = (x_1, x_2) \in \Pi$, one has that $x_1(y_1) = x_1(10) \ge 10$. Consequently, $\mathsf{Pr}(x_1(y_1) \ge D_1) = 1$, and $|x_1(y_1) - D_1| = x_1(y_1) - D_1$ w.p.1. It then follows from a straightforward calculation that the cost of any policy $\pi = (x_1, x_2) \in \Pi$ under the multistage-static formulation equals

$$2x_1(10) - 4 + \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\Big[\tfrac{1}{2}\Big(\big|x_2\big(x_1(10) - 1\big) - D_2\big| + \big|x_2\big(x_1(10) - 3\big) - D_2\big|\Big)\Big]. \tag{17}$$

Let $\bar{\pi} = (\bar{x}_1, \bar{x}_2)$ denote any optimal policy for the multistage-static problem, i.e. $\bar{\pi} \in \Pi_s^{opt}$. Then it follows from (17) and a straightforward contradiction argument that

$$\bar{x}_1(10) = 10 \quad , \quad \big(\bar{x}_2(9), \bar{x}_2(7)\big) \in \operatorname*{arg\,min}_{(x,y): x \ge 9, y \ge 7} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\big[\tfrac{1}{2}\big(|x - D_2| + |y - D_2|\big)\big]. \tag{18}$$

Furthermore, it follows from Lemma 4.1 and Theorem 1 that

$$\inf_{(x,y):x\geq 9,y\geq 7}\sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}\left[\tfrac{1}{2}\left(|x-D_2|+|y-D_2|\right)\right] \quad\geq\quad \sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}\left[|8-D_2|\right] \quad=\quad 2. \tag{19}$$

Noting that $\frac{1}{2}\left(|9-D_2|+|7-D_2|\right)=1+\max(-D_2+7,0,D_2-9)$, it then follows from a straightforward calculation and Theorem 4 that

$$\sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}\left[\tfrac{1}{2}\left(|9-D_2|+|7-D_2|\right)\right]=2. \tag{20}$$

Combining the above, we conclude that $\tilde{\Pi}_s\subseteq\Pi_s^{opt}$. Also, it then follows from a straightforward calculation that the multistage-static problem has optimal value 18.

We now prove that $\tilde{\Pi}_s=\Pi_s^{opt}$. Indeed, suppose for contradiction that there exists some optimal policy $\hat{\pi}=(\hat{x}_1,\hat{x}_2)\notin\tilde{\Pi}_s$. In that case, it follows from (18) that $\frac{1}{2}\left(\hat{x}_2(9)+\hat{x}_2(7)\right)>8$. However, it then follows from Jensen's inequality, Theorem 1, and (19) that

$$\sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}\left[\tfrac{1}{2}\left(|\hat{x}_2(9)-D_2|+|\hat{x}_2(7)-D_2|\right)\right] \quad\geq\quad \sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}\left[\left|\tfrac{1}{2}\left(\hat{x}_2(9)+\hat{x}_2(7)\right)-D_2\right|\right] \quad>\quad 2.$$

Combining with (19) and (20) yields a contradiction, completing the proof $\square$

We now (partially) characterize the set of robust-w.p.1-optimal policies for the distributionally robust DP problem.

Lemma 6.2 $\Pi_d^{opt}\subseteq\tilde{\Pi}_d$, *and the distributionally robust DP problem has optimal value* $17+\frac{\sqrt{5}}{2}$.

Proof : Let $\bar{\pi}=(\bar{x}_1,\bar{x}_2)$ denote any robust-w.p.1-optimal policy for the distributionally robust DP problem, i.e. $\bar{\pi}\in\Pi_d^{opt}$. Then it again follows from a straightforward contradiction argument that $\bar{x}_1(10)=10$. It then follows from (12) that $\bar{x}_2(9)\in\arg\min_{x\geq 9}\sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}[|x-D_2|]$, and $\bar{x}_2(7)\in\arg\min_{x\geq 7}\sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}[|x-D_2|]$. The lemma then follows from Theorem 1 and a straightforward calculation. $\square$

Combining Lemmas 6.1 and 6.2 completes the proof of Observation 5.

6.4. Proof of Observation 6

We first prove that the multistage-static problem is type-I and type-II weakly time consistent.

Lemma 6.3 *The multistage-static problem is type-I and type-II weakly time consistent, and both the multistage-static and distributionally robust DP problems have optimal value 2.*

Proof : Note that $\hat{\Psi}_1(x_1,d_1)=|x_1-d_1|$, and $\hat{\Psi}_2(x_2,d_2)=|x_2-d_2|$. It follows from Observation 1 that $\mathfrak{M}_1$ consists of the single probability measure Q_1 such that $Q_1(1)=Q_1(3)=\frac{1}{2}$. It follows from Theorem 1 and a straightforward calculation that $\hat{\Gamma}_1=[1,3]$, $\hat{\Gamma}_2=10$, $\hat{\eta}_2=1$. Combining the above with Theorem 2, we conclude that the base-stock policy π such that $x_1(y)=\max\{3,y\}$, and $x_2(y) = \max\{10, y\}$ for all $y \in \mathbb{R}$, is optimal for both the multistage-static formulation and robust-w.p.1-optimal for the distributionally robust DP formulation, which have common optimal value 2. $\square$

We now prove that the multistage-static problem is not type-I or type-II strongly time consistent.

Lemma 6.4 *The multistage-static problem is not type-I or type-II strongly time consistent.*

Proof : Consider the policy $\pi' = (x_1', x_2')$ such that

$$x_1'(y) = \max\{3, y\}, \quad \text{and} \quad x_2'(y) = \begin{cases} 9.9, & \text{if } y \leq 0, \\ \max\{10.1, y\}, & \text{otherwise.} \end{cases} \tag{21}$$

We first show that $\pi' \in \Pi_s^{opt}$. It follows from a straightforward calculation that the cost of π' under the multistage-static formulation equals

$$\mathbb{E}_{Q_1}|3 - D_1| + 0.1 + \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \max\left\{9.9 - D_2, 0, D_2 - 10.1\right\}. \tag{22}$$

It is easily verified that the conditions of Theorem 4 are met, and we may apply Theorem 4 to conclude that $\arg\max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \max\left\{9.9 - D_2, 0, D_2 - 10.1\right\}$ is the probability measure Q_2 such that $Q_2(9) = \frac{1}{2}$, $Q_2(11) = \frac{1}{2}$. It follows that the value of (22) equals 2, and we conclude that $\pi' \in \Pi_s^{opt}$, completing the proof.

However, it follows from a straightforward calculation (and considering the measure $Q_1 \in \mathfrak{M}_1$ such that $Q_1(1) = Q_1(3) = \frac{1}{2}$) that type-II time consistency of π' would imply that $9.9 \in \arg\min_{x \geq 0} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}[|x - D_2|]$. But Theorem 1 implies that ths arg min must be the singleton $\{10\}$, reaching a contradiction and completing the proof. $\square$

Combining Lemmas 6.3 and 6.4 completes the proof of Observation 6.

6.5. Proof of Observation 7

We first characterize the set of optimal policies for the multistage-static problem.

Lemma 6.5 $\Pi_s^{opt} = \tilde{\Pi}$*, and the multistage-static problem has optimal value 5.*

Proof : It follows from Observation 1 that $\mathfrak{M}_1$ consists of the single probability measure Q_1 such that $Q_1(1) = Q_1(3) = \frac{1}{2}$. In this case, the cost of any policy $\pi = (x_1, x_2) \in \Pi$ under the multistage-static formulation equals

$$\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[\mathbb{E}_{Q_1}\left[2\Big(x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big)\Big) + \left|x_2\big(x_1(0) - D_1\big) - D_2\right|\right]\right]. \tag{23}$$

We now prove that for any policy $\bar{\pi} = (\bar{x}_1, \bar{x}_2) \in \Pi_s^{opt}$, one has that

$$\bar{x}_2\big(\bar{x}_1(0) - 1\big) = \bar{x}_1(0) - 1 \quad \text{and} \quad \bar{x}_2\big(\bar{x}_1(0) - 3\big) = \bar{x}_1(0) - 3. \tag{24}$$

Indeed, note that w.p.1, it follows from the triangle inequality that

$$\begin{aligned}
&2\Big(x_2\big(x_1(0)-D_1\big)-\big(x_1(0)-D_1\big)\Big)+\big|x_2\big(x_1(0)-D_1\big)-D_2\big| \\
&= 2\Big(x_2\big(x_1(0)-D_1\big)-\big(x_1(0)-D_1\big)\Big)+\big|x_2\big(x_1(0)-D_1\big)-\big(x_1(0)-D_1\big)+\big(x_1(0)-D_1\big)-D_2\big| \\
&\geq 2\Big(x_2\big(x_1(0)-D_1\big)-\big(x_1(0)-D_1\big)\Big)+\big|\big(x_1(0)-D_1\big)-D_2\big|-\big|x_2\big(x_1(0)-D_1\big)-\big(x_1(0)-D_1\big)\big| \\
&= x_2\big(x_1(0)-D_1\big)-\big(x_1(0)-D_1\big)+\big|x_1(0)-D_1-D_2\big|. \qquad (25)
\end{aligned}$$

Now, suppose for contradiction that (24) does not hold. It follows that $\mathbb{E}_{Q_1}\big[x_2\big(x_1(0)-D_1\big)-\big(x_1(0)-D_1\big)\big]>0$, and combining with (25), we conclude that (23) is strictly greater than

$$\sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}\Big[\mathbb{E}_{Q_1}\Big[\big|x_1(0)-D_1-D_2\big|\Big]\Big]. \qquad (26)$$

Noting that (26) is the cost incurred by some policy satisfying (24) completes the proof.

We now complete the proof of the lemma. It suffices from the above to prove that

$$\arg\min_{x_1\in\mathbb{R}_+}\sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}\left[\tfrac{1}{2}\big(|x_1-1-D_2|+|x_1-3-D_2|\big)\right]=\{102\}. \qquad (27)$$

It follows from a straightforward calculation that as long as $x_1\geq 3$, $(x_1-100)(104-x_1)\leq 25$ and $x_1-2-\big((x_1-2-100)^2+25\big)^{\frac{1}{2}}\geq 0$, which holds for all $x_1\in[100,104]$, the conditions of Theorem 4 are met. We may thus apply Theorem 4 to conclude that for all $x_1\in[100,104]$,

$$\sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}\left[\tfrac{1}{2}\big(|x_1-1-D_2|+|x_1-3-D_2|\big)\right] \qquad (28)$$

has the unique optimal solution $\hat{Q}_2$ such that

$$\hat{Q}_2\Big(x_1-2-\big((x_1-2-100)^2+25\big)^{\frac{1}{2}}\Big)=25\Big(25+\big(x_1-2-\big((x_1-2-100)^2+25\big)^{\frac{1}{2}}-100\big)^2\Big)^{-1},$$

and

$$\hat{Q}_2\Big(x_1-2+\big((x_1-2-100)^2+25\big)^{\frac{1}{2}}\Big)=1-25\Big(25+\big(x_1-2-\big((x_1-2-100)^2+25\big)^{\frac{1}{2}}-100\big)^2\Big)^{-1}.$$

It then follows from a straightforward calculation that for $x_1\in[100,104]$, (28) has the value $g(x_1):=\big(x_1^2-204x_1+10429\big)^{\frac{1}{2}}$. It is easily verified that g is a strictly convex function on $[100,104]$, g has its unique minimum on that interval at the point 102, and $g(102)=5$. The desired result then follows from the fact that (28) is a convex function of x_1 on $\mathbb{R}$. $\square$

We now prove that the multistage-static problem is type-I and type-II strongly time consistent.

Lemma 6.6 *The multistage-static problem is type-I and type-II strongly time consistent, and the optimal value of the distributionally robust DP problem equals* $\sqrt{26}$.

Proof : First, we note that as in the multistage-static setting, any policy $\bar{\pi} = (\bar{x}_1, \bar{x}_2) \in \Pi_d^{opt}$ also satisfies (24). The proof is very similar to that used for the multistage-static case, and we omit the details. To prove the lemma, it thus suffices to prove that

$$\arg\min_{x_1 \in \mathbb{R}^+} \left(\tfrac{1}{2} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|x_1 - 1 - D_2|\right] + \tfrac{1}{2} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|x_1 - 3 - D_2|\right] \right) = \{102\}. \tag{29}$$

It is easily verified that for all $x_1 \in [100, 104]$, we may apply Theorem 1 to conclude that $\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|x_1 - 1 - D_2|\right] = \left((x_1 - 101)^2 + 25\right)^{\frac{1}{2}}$; $\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|x_1 - 3 - D_2|\right] = \left((x_1 - 103)^2 + 25\right)^{\frac{1}{2}}$. We conclude that for all $x_1 \in [100, 104]$,

$$\tfrac{1}{2} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|x_1 - 1 - D_2|\right] + \tfrac{1}{2} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|x_1 - 3 - D_2|\right] \tag{30}$$

equals

$$g(x_1) := \tfrac{1}{2} \left(\left((x_1 - 101)^2 + 25\right)^{\frac{1}{2}} + \left((x_1 - 103)^2 + 25\right)^{\frac{1}{2}} \right). \tag{31}$$

It is easily verified that $g(x)$ is a strictly convex function of x on $[100, 104]$, g has its unique minimum on that interval at the point 102, and $g(102) = \sqrt{26}$. The desired result then follows from the fact that (30) is a convex function of x_1 on $\mathbb{R}$. $\square$

Combining Lemmas 6.5 and 6.6 completes the proof of Observation 7.

6.6. Proof of Observation 8

Let $\tilde{Q}_2$ denote the probability measure such that $\tilde{Q}_2(5) = \tilde{Q}_2(11) = \frac{1}{2}$. It may be easily verified that $\tilde{Q}_2 \in \mathfrak{M}_2$. We begin by proving the following auxiliary lemma.

Lemma 6.7

$$\sup_{Q_1 \in \mathfrak{M}_1,\ Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_1 \times Q_2}\Big[|10 - D_1 - D_2|\Big] = 3.$$

Proof : Note that $\mathbb{E}_{Q_1 \times Q_2}\Big[|10 - D_1 - D_2|\Big] = \mathbb{E}_{Q_2}\Big[\mathbb{E}_{Q_1}\Big[|10 - D_1 - D_2| \Big| D_2\Big]\Big]$. Let $\phi_{Q_1}(d) \triangleq \mathbb{E}_{Q_1}\Big[|10 - D_1 - D_2| \Big| \{D_2 = d\}\Big]$, and $q(d) \triangleq \frac{1}{6}(d-8)^2 + \frac{3}{2} = \frac{73}{6} - \frac{8}{3}d + \frac{1}{6}d^2$. As $\tilde{Q}_2 \in \mathfrak{M}_2$, to prove the lemma, it follows from the well-known duality theory for Problem (2) that it suffices to demonstrate that for all $Q_1 \in \mathfrak{M}_1$, $q(5) = \phi_{Q_1}(5), q(11) = \phi_{Q_1}(11)$, and $q(d) \geq \phi_{Q_1}(d)$ for all $d \in \mathbb{R}$, as in this case for any $Q_1 \in \mathfrak{M}_1$, $\sup_{Q_2 \in \mathfrak{M}_2} E_{Q_2}[\phi_{Q_1}(D_2)] = E_{\tilde{Q}_2}[q(D_2)] = 3$. We now prove that $q(d) \geq \phi_{Q_1}(d)$ for all $d \in \mathbb{R}$. For any $Q_1 \in \mathfrak{M}_1$, since $10 - D_1 \in [7 - \epsilon,\ 9 + \epsilon]$ w.p.1, it follows that $\phi_{Q_1}(d) = 10 - \mu_1 - d = 8 - d$ if $d \in [0, 7 - \epsilon]$, and $\phi_{Q_1}(d) = d + \mu_1 - 10 = d - 8$ if $d \in [9 + \epsilon, \infty)$. It is easily verified that $q(d) - (8 - d) \geq 0$, and $q(d) - (d - 8) \geq 0$, for all $d \in \mathbb{R}$. It follows that $q(d) \geq \phi_{Q_1}(d)$ for all $d \in (-\infty, 7 - \epsilon] \bigcup [9 + \epsilon, \infty)$. Noting that $\phi_{Q_1}(d)$ is a convex function of d on $(-\infty, \infty)$, we conclude that $\phi_{Q_1}(d) \leq \max\left(\phi_{Q_1}(7 - \epsilon), \phi_{Q_1}(9 + \epsilon)\right)$ for all $d \in [7 - \epsilon, 9 + \epsilon]$. As it is easily verified that $\inf_{d \in \mathbb{R}} q(d) = \frac{3}{2}$, to prove that $q(d) \geq \phi_{Q_1}(d)$ for $d \in [7 - \epsilon, 9 + \epsilon]$, it suffices to show that $\max\left(\phi_{Q_1}(7 - \epsilon), \phi_{Q_1}(9 + \epsilon)\right) \leq \frac{3}{2}$. As $\phi_{Q_1}(7 - \epsilon) = 8 - (7 - \epsilon) = 1 + \epsilon < \frac{3}{2}$, and

$\phi_{Q_1}(9+\epsilon) = (9+\epsilon) - 8 = 1+\epsilon < \frac{3}{2}$, combining the above we conclude that $q(d) \geq \phi(d)$ for all $d \in \mathbb{R}$. As it is easily verified that $q(5) = \phi_{Q_1}(5) = 3$ and $q(11) = \phi_{Q_1}(11) = 3$, combining the above completes the proof. $\square$

Proof of Observation 8 : Note that the cost under any policy $\pi = (x_1, x_2) \in \Pi$ under the multistage-static formulation equals $\sup_{Q_1 \in \mathfrak{M}_1,\ Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_1 \times Q_2}\Big[2|x_1(y_1) - D_1| + |x_2(D_1) - D_2|\Big]$. As $D_1 \leq 3+\epsilon \leq 10-\epsilon$ w.p.1, and $x_1(y_1) \geq y_1 = 10-\epsilon$, we conclude that w.p.1 $|x_1(y_1) - D_1| = x_1(y_1) - D_1 \geq 10 - \epsilon - D_1$. Combining with the fact that $\mu_1 = 2$, we conclude that $\mathbb{E}_{Q_1 \times Q_2}\Big[2|x_1(y_1) - D_1|\Big] \geq 2\big(10-\epsilon-2\big) = 2(8-\epsilon)$. As $\frac{\sigma_2^2}{\mu_2^2} = \frac{9}{64} < \frac{b_2}{h_2} = 1$, and $\big(h_2 b_2\big)^{\frac{1}{2}} \sigma_2 = 3$, it follows from Lemma 4.1 and Theorem 1 that $\mathbb{E}_{Q_1 \times Q_2}\Big[|x_2(D_1) - D_2|\Big] \geq 3$. Combining the above, we conclude that the cost incurred under any policy π is at least $19 - 2\epsilon$.

We now show that the cost incurred under any such policy $\tilde{\pi}$ achieves this bound, and is thus optimal. In particular, $\sup_{Q_1 \in \mathfrak{M}_1,\ Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_1 \times Q_2}\Big[2|\tilde{x}_1(y_1) - D_1| + |\tilde{x}_2(D_1) - D_2|\Big]$ equals

$$
\begin{aligned}
&\sup_{Q_1 \in \mathfrak{M}_1,\ Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_1 \times Q_2}\Big[2|10 - \epsilon - D_1| + |10 - D_1 - D_2|\Big] \\
&= \sup_{Q_1 \in \mathfrak{M}_1,\ Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_1 \times Q_2}\Big[2(10 - \epsilon - D_1) + |10 - D_1 - D_2|\Big] \\
&= 2(10 - \epsilon - \mu_1) + \sup_{Q_1 \in \mathfrak{M}_1,\ Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_1 \times Q_2}\Big[|10 - D_1 - D_2|\Big] \quad = \quad 19 - 2\epsilon,
\end{aligned}
$$

where the final equality follows from Lemma 6.7.

Next we show that there is no optimal base-stock policy, i.e. no base-stock policy belongs to Π_s^{opt}. Indeed, let us suppose for contradiction that $\hat{\pi}$ is a base-stock policy with constants $\hat{x}_1$, $\hat{x}_2$. The cost incurred under such a policy $\hat{\pi}$ equals

$$
\sup_{Q_1 \in \mathfrak{M}_1,\ Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_1 \times Q_2}\Big[2|\max(\hat{x}_1, y_1) - D_1| + \big|\max\big(\max(\hat{x}_1, y_1) - D_1, \hat{x}_2\big) - D_2\big|\Big].
$$

It follows from the fact that $D_1 \leq 3+\epsilon < 10-\epsilon$ w.p.1 for all $Q_1 \in \mathfrak{M}_1$, and a straightforward contradiction argument (the details of which we omit), that $\hat{\pi}$ cannot be optimal unless $\hat{x}_1 \leq 10-\epsilon$, in which case repeating our earlier arguments, we conclude that $\max(\hat{x}_1, y_1) = 10-\epsilon$, and for any $Q_1 \in \mathfrak{M}_1, Q_2 \in \mathfrak{M}_2$, $\mathbb{E}_{Q_1 \times Q_2}\big[2|\max(\hat{x}_1, y_1) - D_1|\big] = 2(8-\epsilon)$. Thus to prove the desired claim, it suffices to demonstrate that

$$
\inf_{\hat{x}_2 \in \mathbb{R}} \sup_{Q_1 \in \mathfrak{M}_1, Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_1 \times Q_2}\Big[\Big|\max\{10 - \epsilon - D_1, \hat{x}_2\} - D_2\Big|\Big] > 3. \tag{32}
$$

We treat two different cases: $\hat{x}_2 \in (-\infty, 7+\frac{1}{2}\epsilon]$ and $\hat{x}_2 \in [7+\frac{1}{2}\epsilon, \infty)$. If $\hat{x}_2 \leq 7+\frac{1}{2}\epsilon$, let the probability measure $\tilde{Q}_1$ be such that $\tilde{Q}_1(1) = \tilde{Q}_1(3) = \frac{1}{2}$, where it is easily verified that $\tilde{Q}_1 \in \mathfrak{M}_1$. In this case,

$$
\sup_{Q_1 \in \mathfrak{M}_1, Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_1 \times Q_2}\Big[\Big|\max\{10 - \epsilon - D_1, \hat{x}_2\} - D_2\Big|\Big] \tag{33}
$$

is at least

$$\begin{aligned}
&\sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{\tilde{Q}_1\times Q_2}\Big[\Big|\max\{10-\epsilon-D_1,\hat{x}_2\}-D_2\Big|\Big] \\
&\quad= \sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_2}\Big[\frac{1}{2}\left|\max\{7-\epsilon,\hat{x}_2\}-D_2\right| + \frac{1}{2}\left|9-\epsilon-D_2\right|\Big], \qquad (34)
\end{aligned}$$

where the final equality follows from the fact that $\hat{x}_2 \le 7+\frac{1}{2}\epsilon$ implies $\max\{9-\epsilon,\hat{x}_2\} = 9-\epsilon$. It follows from convexity of the absolute value function that (34) is at least

$$\sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_2}\Big[\Big|\frac{1}{2}\max\{7-\epsilon,\hat{x}_2\} + \frac{1}{2}(9-\epsilon) - D_2\Big|\Big]. \qquad (35)$$

Note that

$$\frac{1}{2}\max\{7-\epsilon,\hat{x}_2\} + \frac{1}{2}(9-\epsilon) \quad\ge\quad \frac{1}{2}(7-\epsilon) + \frac{1}{2}(9-\epsilon) \quad=\quad 8-\epsilon. \qquad (36)$$

Letting $z \triangleq \frac{1}{2}\max\{7-\epsilon,\hat{x}_2\} + \frac{1}{2}(9-\epsilon)$, note that (35) equals $\sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}\Big[(z-D_2)^+ + (D_2-z)^+\Big]$. Applying Theorem 1 with $c=0, b=h=1$, and noting that $\frac{\mu_2^2+\sigma_2^2}{2\mu_2} = \frac{73}{16} < 8-\epsilon = z$, we conclude that (35) equals

$$\Big(\big(\frac{1}{2}\max\{7-\epsilon,\hat{x}_2\} + \frac{1}{2}(9-\epsilon) - 8\big)^2 + 9\Big)^{\frac{1}{2}}. \qquad (37)$$

Combining (36) with the fact that $\frac{1}{2}\max\{7-\epsilon,\hat{x}_2\} + \frac{1}{2}(9-\epsilon) \;\le\; \frac{1}{2}(7+\frac{1}{2}\epsilon) + \frac{1}{2}(9-\epsilon) \;=\; 8-\frac{1}{4}\epsilon$, we conclude that (37) is strictly greater than 3, completing the proof of (32) for the case $\hat{x}_2 \le 7+\frac{1}{2}\epsilon$.

Alternatively, if $\hat{x}_2 \ge 7+\frac{1}{2}\epsilon$, let the probability measure $\tilde{Q}_1$ be such that $\tilde{Q}_1(\frac{1+2\epsilon}{1+\epsilon}) = \frac{(1+\epsilon)^2}{(1+\epsilon)^2+1}$ and $\tilde{Q}_1(3+\epsilon) = \frac{1}{(1+\epsilon)^2+1}$. Again, it is easily verified that $\tilde{Q}_1 \in \mathfrak{M}_1$. In this case, (33) is at least

$$\sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_2}\left[\frac{1}{(1+\epsilon)^2+1}\left|\hat{x}_2-D_2\right| + \frac{(1+\epsilon)^2}{(1+\epsilon)^2+1}\left|\max\left\{10-\epsilon-\frac{1+2\epsilon}{1+\epsilon},\hat{x}_2\right\}-D_2\right|\right]. \qquad (38)$$

It follows from convexity of the absolute value function that (38) is at least

$$\sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_2}\left[\left|\frac{1}{(1+\epsilon)^2+1}\hat{x}_2 + \frac{(1+\epsilon)^2}{(1+\epsilon)^2+1}\max\left\{10-\epsilon-\frac{1+2\epsilon}{1+\epsilon},\hat{x}_2\right\}-D_2\right|\right]. \qquad (39)$$

Letting $z \triangleq \frac{1}{(1+\epsilon)^2+1}\hat{x}_2 + \frac{(1+\epsilon)^2}{(1+\epsilon)^2+1}\max\left\{10-\epsilon-\frac{1+2\epsilon}{1+\epsilon},\hat{x}_2\right\}$, note that (39) equals $\sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}\Big[(z-D_2)^+ + (D_2-z)^+\Big]$. Furthermore,

$$\begin{aligned}
&\frac{1}{(1+\epsilon)^2+1}\hat{x}_2 + \frac{(1+\epsilon)^2}{(1+\epsilon)^2+1}\max\left\{10-\epsilon-\frac{1+2\epsilon}{1+\epsilon},\hat{x}_2\right\} \\
&\quad\ge \frac{1}{(1+\epsilon)^2+1}\left(7+\frac{1}{2}\epsilon\right) + \frac{(1+\epsilon)^2}{(1+\epsilon)^2+1}\left(10-\epsilon-\frac{1+2\epsilon}{1+\epsilon}\right) \quad=\quad 8+\frac{\frac{1}{2}-2\epsilon-\epsilon^2}{(1+\epsilon)^2+1}\epsilon. \qquad (40)
\end{aligned}$$

Applying Theorem 1 with $c=0, b=h=1$, and noting that $\frac{\mu_2^2+\sigma_2^2}{2\mu_2} = \frac{73}{16} < 8+\frac{\frac{1}{2}-2\epsilon-\epsilon^2}{(1+\epsilon)^2+1}\epsilon = z$ (having applied (15)), we conclude that (39) equals

$$\left(\Big(\frac{1}{(1+\epsilon)^2+1}\hat{x}_2 + \frac{(1+\epsilon)^2}{(1+\epsilon)^2+1}\max\left\{10-\epsilon-\frac{1+2\epsilon}{1+\epsilon},\hat{x}_2\right\} - 8\Big)^2 + 9\right)^{\frac{1}{2}}. \qquad (41)$$

Combining with (40) and (15), we conclude that (41) is strictly greater than 3, completing the proof of (32) for the case $\hat{x}_2 \le 7+\frac{1}{2}\epsilon$, which completes the proof. $\square$

7. Supplemental Appendix

7.1. Review of rectangularity and connection to our definitions.

To contextualize our definitions within the broader literature, we here briefly review the relevant notion of rectangularity. Our definitions will closely follow those given in Shapiro (2016), although we note that many closely related definitions have appeared previously throughout the literature (see e.g. Iyengar (2005)). Consider the cost $Z^\pi = Z^\pi(D_{[T]})$ of a policy π, defined in (10). Let $\widehat{\mathfrak{M}}$ be a set of probability distributions for the demand vector $D_{[T]}$, and let $Q \in \widehat{\mathfrak{M}}$. At the moment we do not assume that Q is of the product form $Q = Q_1 \times \cdots \times Q_T$. We can write

$$\mathbb{E}_Q[Z^\pi] = \mathbb{E}_Q\left[\mathbb{E}_{Q|D_1}\left[\cdots \mathbb{E}_{Q|D_{[T-2]}}\left[\mathbb{E}_{Q|D_{[T-1]}}[Z^\pi]\right]\right]\right], \tag{42}$$

where $\mathbb{E}_{Q|D_{[t]}}[Z^\pi]$ is the conditional expectation, given $D_{[t]}$, with respect to the distribution Q of $D_{[T]}$. Of course, this conditional expectation is a function of $D_{[t]}$. Consequently,

$$\sup_{Q\in\widehat{\mathfrak{M}}} \mathbb{E}_Q[Z^\pi] \le \sup_{Q\in\widehat{\mathfrak{M}}} \mathbb{E}_Q\left[\sup_{Q\in\widehat{\mathfrak{M}}} \mathbb{E}_{Q|D_1}\left[\cdots \sup_{Q\in\widehat{\mathfrak{M}}} \mathbb{E}_{Q|D_{[T-1]}}[Z^\pi]\right]\right]. \tag{43}$$

The right hand side of (43) leads to the nested formulation

$$\inf_{\pi\in\Pi}\left\{\sup_{Q\in\widehat{\mathfrak{M}}} \mathbb{E}_Q\left[\sup_{Q\in\widehat{\mathfrak{M}}} \mathbb{E}_{Q|D_1}\left[\cdots \sup_{Q\in\widehat{\mathfrak{M}}} \mathbb{E}_{Q|D_{[T-1]}}[Z^\pi]\right]\right]\right\}. \tag{44}$$

In particular, if the set $\widehat{\mathfrak{M}}$ is defined in the form (8), i.e., consists of products of probability measures (with the t-th measure drawn from $\mathfrak{M}_t$), then formulation (44) simplifies to

$$\inf_{\pi\in\Pi}\left\{\sup_{Q_1\in\mathfrak{M}_1} \mathbb{E}_{Q_1}\left[\sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_2|D_1}\left[\cdots \sup_{Q_T\in\mathfrak{M}_T} \mathbb{E}_{Q_T|D_{[T-1]}}[Z^\pi]\right]\right]\right\}. \tag{45}$$

It follows from (43) that the optimal value of (44) is greater than or equal to the optimal value of the multistage-static Problem (9). Moreover, the optimal value of (44) can be strictly greater than the optimal value of (9). Let us demonstrate this through the following simple example.

Example 1 *Let $T = 2$ and the set $\widehat{\mathfrak{M}}$ be of the product form (8). Suppose further that $\mathcal{I}_1 = [0,1]$, $\mu_1 = 1/2$ and $\sigma_1^2 = 1/4$. Then $\mathfrak{M}_1 = \{Q_1\}$ is a singleton with $Q_1 = p_1\delta_0 + p_2\delta_1$, $p_1 = p_2 = 1/2$, i.e., with probability $1/2$ the demand D_1 can be either zero or one. Let us fix some policy $\pi \in \Pi$ and let $Z^\pi = Z^\pi(D_1, D_2)$ be the corresponding objective function. Then the associated cost under formulation (9) equals*

$$\sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_1\times Q_2}[Z^\pi(D_1,D_2)] = \sup_{Q_2\in\mathfrak{M}_2}\Big(p_1\mathbb{E}_{Q_2}[Z^\pi(0,D_2)] + p_2\mathbb{E}_{Q_2}[Z^\pi(1,D_2)]\Big), \tag{46}$$

while the associated cost under formulation (44) equals

$$\mathbb{E}_{Q_1}\left[\sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_2|D_1}[Z^\pi(D_1,D_2)]\right] = p_1 \sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_2}[Z^\pi(0,D_2)] + p_2 \sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_2}[Z^\pi(1,D_2)]. \tag{47}$$

Note that the worst case distribution $Q_2 \in \mathfrak{M}_2$ in (46) has to be the same for all possible realizations of the demand D_1. In contrast, the worst case distribution $Q_2 \in \mathfrak{M}_2$ in (47) is allowed to depend on realized D_1. Hence the right hand side of (47) can be strictly greater than the right hand side of (46).

In line with the definition given in Shapiro (2016), we say that the set $\widehat{\mathfrak{M}}$ of probability measures is rectangular if such strict inequality does not occur for any r.v. Z^π, i.e. the two formulations are equivalent in their optimal values. More formally, we make the following definition.

Definition 7.1 *Consistent with the definition given in Shapiro (2016), we say that the set $\widehat{\mathfrak{M}}$ of probability measures is rectangular if for every measurable and non-negative function f,*

$$\sup_{Q\in\widehat{\mathfrak{M}}} \mathbb{E}_Q[f(D_{[T]})] = \sup_{Q\in\widehat{\mathfrak{M}}} \mathbb{E}_Q\left[\sup_{Q\in\widehat{\mathfrak{M}}} \mathbb{E}_{Q|D_1}\Big[\cdots \sup_{Q\in\widehat{\mathfrak{M}}} \mathbb{E}_{Q|D_{[T-1]}}[f(D_{[T]})]\Big]\right]. \tag{48}$$

We note that under additional compactness assumptions on the support of $D_{[T]}$, Shapiro (2016) formally explores several related concepts and subtleties of this definition, e.g. proves that one can associate a rectangular set of probability measures to any given (possibly non-rectangular) set of probability measures, but as such a compactness condition does not hold in our setting we do not explore that further here.

For a rectangular set $\widehat{\mathfrak{M}}$ the static formulation

$$\inf_{\pi\in\Pi} \sup_{Q\in\widehat{\mathfrak{M}}} \mathbb{E}_Q\big[Z^\pi\big], \tag{49}$$

is equivalent (in an appropriate sense, see e.g. Shapiro (2016, 2017)) to the formulation (44). Furthermore, the natural generalization of the distributionally robust DP equations (11) can be applied to (49), with both formulations having a common optimal policy and the same optimal value.

We note that the concept of rectangularity has been central to the past literature on time consistency (cf. Epstein et al. (2003), Grunwald et al. (2011), Iancu et al. (2015)), especially as it relates to optimization (cf. Iyengar (2005), Nilim et al. (2005), Wiesemann et al. (2013)). In several of these works, connections were made between tractability of the associated robust MDP and various notions of rectangularity (e.g. (s,a)-rectangularity, s-rectangularity). We refer the interested reader to Wiesemann et al. (2013) and the references therein for details. Our definition of rectangularity is aimed directly at the decomposability property of the static formulation ensuring its equivalence to the corresponding dynamic formulation (see Shapiro (2016) for details).

In general the set of product measures $\mathfrak{M}$ we consider in this work is not rectangular, as certified by the possible lack of weak time-consistency shown in our examples. We note that a rectangular analogue of the set of measures $\mathfrak{M}$ defined in (7)-(8) would be the set of all joint distributions Q for $D_{[T]}$ such that

$$D_t \in \mathfrak{P}(\mathcal{I}_t) \quad , \quad \mathbb{E}_Q[D_t|D_{[t-1]}] = \mu_t \quad , \quad \mathbb{E}_Q[D_t^2|D_{[t-1]}] = \mu_t^2 + \sigma_t^2, \quad t = 1, \ldots, T. \tag{50}$$

Non-rectangular (and intractable) formulations for robust MDP are described in both Iyengar (2005) and Nilim et al. (2005). In Iyengar (2005), it is referred to as the static formulation, while in Nilim et al. (2005), it is referred to as the stationary formulation. In both of these settings, these non-rectangular formulations essentially equate to requiring nature to select the same transition kernel every time a given state (and action, depending on the formulation) is encountered, as opposed to being able to select a different kernel every time a given state is visited in the robust MDP, and we refer the reader to Iyengar (2005), Nilim et al. (2005), and Wiesemann et al. (2013) for details. Although our multistage-static formulation could similarly be phrased in terms of a particular kind of dependency between the choices of nature in a robust MDP framework, and would be significantly different from either of the aforementioned non-rectangular formulations, we do not pursue such an investigation here, and leave the formalization of such connections as a direction for future research.

7.2. Supplemental Proof of Theorem 1

Proof of Theorem 1 : We first compute the value of $\psi(x)$ for all $x \in \mathbb{R}$, and proceed by a case analysis. First, suppose $x < 0$. In this case, $\mathbb{E}_Q[\Psi(x, D)] = cx + b(\mu - x)$ for all $Q \in \mathfrak{M}$, and thus

$$\psi(x) = cx + b(\mu - x). \tag{51}$$

Now, suppose $x \geq 0$. Then it is easily verified that

$$\psi(x) = cx + \frac{(h-b)(x-\mu)}{2} + \frac{b+h}{2} \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q\left[|x - D|\right]. \tag{52}$$

Hence to compute $\psi(x)$, it suffices to solve $\sup_{Q\in\mathfrak{M}} \mathbb{E}_Q\left[|x - D|\right]$, and we proceed by a case analysis. Recall that $\mathfrak{f}(z) := \left((z-\mu)^2 + \sigma^2\right)^{\frac{1}{2}}$ for all $z \in \mathbb{R}$, and $\mathfrak{f}^{-1}(z)$ denotes the reciprocal of $\mathfrak{f}(z)$.

First, suppose $x \geq \frac{\mu^2+\sigma^2}{2\mu}$. Let us define $\bar{\lambda} = (\bar{\lambda}_0, \bar{\lambda}_1, \bar{\lambda}_2)$ such that

$$\bar{\lambda}_0 := \frac{1}{2}\left(x^2 \mathfrak{f}^{-1}(x) + \mathfrak{f}(x)\right), \quad \bar{\lambda}_1 := -x\mathfrak{f}^{-1}(x), \quad \bar{\lambda}_2 := \frac{1}{2}\mathfrak{f}^{-1}(x),$$

and let $\bar{g}(d) := \bar{\lambda}_0 + \bar{\lambda}_1 d + \bar{\lambda}_2 d^2$ for all $d \in \mathbb{R}$. Then it follows from a straightforward calculation that $\bar{g}(d)$ and $|x - d|$ are tangent at $\bar{d}_1 := x - \mathfrak{f}(x)$ and $\bar{d}_2 := x + \mathfrak{f}(x)$, and consequently $\bar{g}(d) \geq |x - d|$ for all $d \in \mathbb{R}_+$. Hence $\bar{\lambda}$ is feasible for the dual of Problem (2). Also, as $x \geq \frac{\mu^2+\sigma^2}{2\mu}$ implies $\bar{d}_1 \geq 0$, it is easily verified that the probability measure $\bar{Q}$ such that

$$\bar{Q}(\bar{d}_1) = \sigma^2\Big(\sigma^2 + \big(x - \mathfrak{f}(x) - \mu\big)^2\Big)^{-1}, \quad \bar{Q}(\bar{d}_2) = 1 - \sigma^2\Big(\sigma^2 + \big(x - \mathfrak{f}(x) - \mu\big)^2\Big)^{-1}$$

is feasible for the primal Problem (2). It follows from standard duality results that $\bar{Q}$ is an optimal primal solution. Combining the above and simplifying the relevant algebra, we conclude that in this case

$$\psi(x) = \psi_1(x) := c\mu + \frac{b+h}{2}\mathfrak{f}(x) - \frac{b-h-2c}{2}(x-\mu). \tag{53}$$

Alternatively, suppose $x \in [0, \frac{\mu^2+\sigma^2}{2\mu})$. Let us define $\hat{\lambda} = (\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2)$ such that

$$\hat{\lambda}_0 := x \quad , \quad \hat{\lambda}_1 := 1 - 4x\mu(\mu^2+\sigma^2)^{-1} \quad , \quad \hat{\lambda}_2 := 2x\big(\mu(\mu^2+\sigma^2)^{-1}\big)^2,$$

and let $\hat{g}(d) := \hat{\lambda}_0 + \hat{\lambda}_1 d + \hat{\lambda}_2 d^2$ for all $d \in \mathbb{R}$. Then it follows from a straightforward calculation that $\hat{g}(d)$ and $|x-d|$ are tangent at $\hat{d}_1 := \mu^{-1}(\mu^2+\sigma^2)$, and intersect at $\hat{d}_2 := 0$, with $\hat{g}'(0) \geq -1$. It follows that $\hat{g}(d) \geq |x-d|$ for all $d \in \mathbb{R}_+$. Hence $\hat{\lambda}$ is feasible for the dual of Problem (2). Also, it is easily verified that the probability measure $\hat{Q}$ such that

$$\hat{Q}(\hat{d}_1) = \mu^2(\mu^2+\sigma^2)^{-1}, \quad \hat{Q}(\hat{d}_2) = 1 - \mu^2(\mu^2+\sigma^2)^{-1}$$

is feasible for the primal Problem (2). It follows from standard duality results that $\hat{Q}$ is an optimal primal solution. Combining the above and simplifying the relevant algebra, we conclude that in this case

$$\psi(x) = \psi_2(x) := \frac{(h+c)\sigma^2 - (b-c)\mu^2}{\mu^2+\sigma^2}x + b\mu. \tag{54}$$

We now use the above to complete the proof of the theorem. Note that since by assumption $b > c$, it follows from (51) that $\arg\min_{x\in\mathbb{R}} \psi(x) \subseteq \mathbb{R}_+$. Recall that $\kappa = \frac{b-h-2c}{b+h}$. Furthermore, our assumptions, i.e. $b > c, h+c > 0$, imply that $|\kappa| < 1$. Let $\chi := \mu + \kappa\sigma(1-\kappa^2)^{-\frac{1}{2}}$. It follows from a straightforward calculation that ψ_1 is a strictly convex function on $\mathbb{R}$, and $\psi_1(\chi) = 0$, i.e. ψ_1 is strictly decreasing on $(-\infty, \chi)$, and strictly increasing on (χ, ∞). Furthermore, it follows from a similar calculation that

$$\frac{\sigma^2}{\mu^2} - \frac{b-c}{h+c} \quad \text{is the same sign as} \quad \frac{\mu^2+\sigma^2}{2\mu} - \chi. \tag{55}$$

We now proceed by a case analysis. First, suppose $\frac{\sigma^2}{\mu^2} > \frac{b-c}{h+c}$. In this case, ψ_2 is a linear function with strictly positive slope, and thus $\arg\min_{x\in[0,\frac{\mu^2+\sigma^2}{2\mu}]} \psi(x) = \{0\}$. Furthermore, it follows from (55) that $\chi < \frac{\mu^2+\sigma^2}{2\mu}$, which implies that ψ_1 is strictly increasing on $[\frac{\mu^2+\sigma^2}{2\mu}, \infty)$. It follows from the continuity of ψ that $\arg\min_{x\geq\frac{\mu^2+\sigma^2}{2\mu}} \psi(x) = \{\frac{\mu^2+\sigma^2}{2\mu}\}$. Combining the above, we conclude that

$\arg\min_{x\in\mathbb{R}} \psi(x) = \{0\}$.

Next, suppose $\frac{\sigma^2}{\mu^2} < \frac{b-c}{h+c}$. In this case, ψ_2 is a linear function with strictly negative slope, and thus $\arg\min_{x\in[0,\frac{\mu^2+\sigma^2}{2\mu}]} \psi(x) = \{\frac{\mu^2+\sigma^2}{2\mu}\}$. Furthermore, it follows from (55) that $\chi > \frac{\mu^2+\sigma^2}{2\mu}$, which implies that $\arg\min_{x\geq\frac{\mu^2+\sigma^2}{2\mu}} \psi(x) = \{\chi\}$. Combining the above, we conclude that $\arg\min_{x\in\mathbb{R}} \psi(x) = \{\chi\}$.

Finally, suppose that $\frac{\sigma^2}{\mu^2} = \frac{b-c}{h+c}$. In this case, ψ_2 is a constant function, and thus $\arg\min_{x\in[0,\frac{\mu^2+\sigma^2}{2\mu}]} \psi(x) = [0, \frac{\mu^2+\sigma^2}{2\mu}]$. Furthermore, it follows from (55) that $\chi = \frac{\mu^2+\sigma^2}{2\mu}$, which implies that $\arg\min_{x\geq\frac{\mu^2+\sigma^2}{2\mu}} \psi(x) = \{\frac{\mu^2+\sigma^2}{2\mu}\}$. Combining the above, we conclude that $\arg\min_{x\in\mathbb{R}} \psi(x) = [0, \frac{\mu^2+\sigma^2}{2\mu}]$.

Combining all of the above with another straightforward calculation completes the proof of the theorem. □

7.3. Supplemental Proof of Proposition 6.1

Proof of Proposition 6.1 : Let $\delta := \frac{\sigma^2}{\mu^2+\sigma^2}, \tau := \frac{\mu^2+\sigma^2}{\mu}$. Let Q_2^* be the probability measure such that

$$Q_2^*(0) = \delta, \quad Q_2^*(\tau) = 1-\delta.$$

Recall that $b-c>0$, and $(h+c)\sigma^2 > (b-c)\mu^2$, which we denote by assumption A1. Note that the value of any feasible solution Q_1 to Problem (16) is at least $\mathbb{E}_{Q_1\times Q_2^*}\Big[\Psi(D_1,D_2)\Big]$, which itself equals the sum of $c\mu$ and

$$\mathbb{E}_{Q_1}\Big[\Big(\delta\big((b-c)[0-D_1]_+ + (h+c)[D_1-0]_+\big) + (1-\delta)\big((b-c)[\tau-D_1]_+ + (h+c)[D_1-\tau]_+\big)\Big) I(D_1>0)\Big] \tag{56}$$

$$+\,\mathbb{E}_{Q_1}\Big[\Big(\delta\big((b-c)[0-D_1]_+ + (h+c)[D_1-0]_+\big) + (1-\delta)\big((b-c)[\tau-D_1]_+ + (h+c)[D_1-\tau]_+\big)\Big) I(D_1<0)\Big] \tag{57}$$

$$+\,\mathbb{E}_{Q_1}\Big[\Big(\delta\big((b-c)[0-D_1]_+ + (h+c)[D_1-0]_+\big) + (1-\delta)\big((b-c)[\tau-D_1]_+ + (h+c)[D_1-\tau]_+\big)\Big) I(D_1=0)\Big] \tag{58}$$

Note that if $P(D_1>0)>0$, then (56) is at least

$$\begin{aligned}
&\mathbb{E}\Big[\frac{\sigma^2}{\mu^2+\sigma^2}(h+c)D_1 + \frac{\mu^2}{\mu^2+\sigma^2}(b-c)(\frac{\mu^2+\sigma^2}{\mu} - D_1)\big|D_1>0\Big]P(D_1>0) \\
> \quad &\mathbb{E}\Big[\frac{\mu^2}{\mu^2+\sigma^2}(b-c)D_1 + \frac{\mu^2}{\mu^2+\sigma^2}(b-c)(\frac{\mu^2+\sigma^2}{\mu} - D_1)\big|D_1>0\Big]P(D_1>0) \quad \text{by A1} \\
= \quad &(b-c)\mu P(D_1>0).
\end{aligned} \tag{59}$$

Similarly, if $P(D_1<0)>0$, then (57) is at least

$$\begin{aligned}&\mathbb{E}\Big[-\frac{\sigma^2}{\mu^2+\sigma^2}(b-c)D_1+\frac{\mu^2}{\mu^2+\sigma^2}(b-c)(\frac{\mu^2+\sigma^2}{\mu}-D_1)\big|D_1<0\Big]P(D_1<0)\\ =\quad&\mathbb{E}\Big[(b-c)(\mu-D_1)\big|D_1<0\Big]P(D_1<0)\quad>\quad(b-c)\mu P(D_1<0).\end{aligned}\tag{60}$$

Furthermore, if $P(D_1=0)>0$, then (58) equals $(b-c)\mu P(D_1=0)$. Combining with (59), (60), and the fact that the measure δ_0 attains value $b\mu$ (by Theorem 1), completes the proof. □

7.4. Supplemental Proof of Theorem 4

Proof of Theorem 4 : Recall that $\eta:=\frac{1}{2}(c_1+c_2)$, and $\mathfrak{f}(z):=\big((z-\mu)^2+\sigma^2\big)^{\frac{1}{2}}$ for all $z\in\mathbb{R}$. Also, letting $h_1(d):=-d+c_1, h_2(d):=d-c_2$ for all $d\in\mathbb{R}$, we have that $\zeta(d)=\max\{h_1(d),0,h_2(d)\}$ for all $d\in\mathbb{R}$. Let λ denote the vector in $\mathcal{R}^3$ s.t.

$$\lambda_0=\tfrac{1}{2}\big(\eta^2+(\eta-\mu)^2+\sigma^2\big)\mathfrak{f}^{-1}(\eta)+\frac{c_1-c_2}{2},\quad \lambda_1=-\eta\mathfrak{f}^{-1}(\eta),\quad \lambda_2=\tfrac{1}{2}\mathfrak{f}^{-1}(\eta).$$

Let $g(d):=\lambda_0+\lambda_1 d+\lambda_2 d^2$. We now prove that $g(d)\geq\zeta(d)$ for all $d\in\mathbb{R}$. It follows from a straightforward calculation that $g(d)$ is tangent to $h_1(d)$ at $d_1:=\eta-\mathfrak{f}(\eta)$, and $g(d)$ is tangent to $h_2(d)$ at $d_2:=\eta+\mathfrak{f}(\eta)$. Thus $g(d)\geq\max\big(h_1(d),h_2(d)\big)$ for all $d\in\mathbb{R}$, and to prove the desired claim it suffices to demonstrate that $g(d)\geq 0$ for all $d\geq 0$. It is easily verified that for all $d\in\mathbb{R}$,

$$g(d)=\tfrac{1}{2}\mathfrak{f}^{-1}(\eta)(d-\eta)^2+\tfrac{1}{2}\big(\mathfrak{f}(\eta)+c_1-c_2\big).\tag{61}$$

Recall that

$$\frac{1}{4}(2\mu-3c_1+c_2)(3c_2-c_1-2\mu)\leq\sigma^2,$$

which we denote by assumption A2. It follows from another straightforward calculation that assumption A2 is equivalent to requiring that $\frac{1}{2}\big(\mathfrak{f}(\eta)+c_1-c_2\big)\geq 0$. Combining with (61), we conclude that A2 implies $g(d)\geq 0$ for all $d\in\mathbb{R}$, completing the proof that $g(d)\geq\zeta(d)$ for all $d\in\mathbb{R}$. Hence λ is feasible for the dual of Problem (2). Also, it is easily verified that Q is feasible for the primal Problem (2). It follows from standard duality results that Q is an optimal primal solution, and λ is an optimal dual solution. That these optimal solutions are unique again follows from standard duality results and a straightforward contradiction argument. Combining the above and simplifying the relevant algebra completes the proof. □

7.5. Supplemental Proof of Lemma 4.1

Proof of Lemma 4.1 : Suppose $i\in\{0,...,T\}$ and $Q_1,\dots,Q_i$ are fixed. As a notational convenience, for $k\in[1,T]$, let $\mathbb{E}_k[\cdot]$ denote $\mathbb{E}_{\otimes_{j=1}^k Q_j}[\cdot]$. We now prove that (14) holds for all $t\geq i+1$, and proceed by induction. Our particular induction hypothesis will be that there exist $Q_{i+1},\dots,Q_{i+n}$ such that

$$\mathbb{E}_{i+n}\big[\hat{\Psi}_t\big(x_t(y_t),D_t\big)\big]\geq\hat{\eta}_t \text{ for all } t\in[i+1,i+n].\tag{62}$$

We first treat the base case $n = 1$. It follows from Jensen's inequality, and the independence structure of the measures in $\mathfrak{M}$, that for any $Q_{i+1} \in \mathfrak{M}_{i+1}$,

$$\mathbb{E}_{i+1}\big[\hat{\Psi}_{i+1}\big(x_{i+1}(y_{i+1}), D_{i+1}\big)\big] \geq \mathbb{E}_{Q_{i+1}}\big[\hat{\Psi}_{i+1}\big(\mathbb{E}_i[x_{i+1}(y_{i+1})], D_{i+1}\big)\big].$$

Taking Q_{i+1} to be any element of $\Gamma_{i+1}^{\mathbb{E}_i[x_{i+1}(y_{i+1})]}$ ($\Gamma_1^{x_1(y_1)}$ if $i = 0$) completes the proof for $n = 1$.

Now, suppose the induction holds for some n. It again follows from Jensen's inequality, and the independence structure of the measures in $\mathfrak{M}$, that for any $Q_{i+n+1} \in \mathfrak{M}_{i+n+1}$,

$$\mathbb{E}_{i+n+1}\big[\hat{\Psi}_{i+n+1}\big(x_{i+n+1}(y_{i+n+1}), D_{i+n+1}\big)\big] \geq \mathbb{E}_{Q_{i+n+1}}\big[\hat{\Psi}_{i+n+1}\big(\mathbb{E}_{i+n}[x_{i+n+1}(y_{i+n+1})], D_{i+n+1}\big)\big].$$

Taking Q_{i+n+1} to be any element of $\Gamma_{i+n+1}^{\mathbb{E}_{i+n}[x_{i+n+1}(y_{i+n+1})]}$ completes the induction, and the proof, where the second part of the lemma follows by letting $i = 0$. □

7.6. Supplemental Proof of Theorem 2

Proof of Theorem 2 : Note that under these assumptions, for any measure $Q \in \mathcal{M}$ (and in fact any non-negative joint distribution for demand), for any such base-stock policy π, w.p.1 $x_t^\pi(y_t) = x_t^*$ for all $t = 1, ..., T$. It then follows from a straightforward induction that π is a robust-w.p.1-optimal policy for the distributionally robust DP formulation, and furthermore for all $t = 1, ..., T$ and $y \leq x_t^*$,

$$V_t(y) = \hat{\eta}_t - c_t x_{t-1}^* + c_t D_{t-1} + \sum_{s=t+1}^{T} \rho^{s-t}(\hat{\eta}_s + c_s \mu_{s-1}),$$

and

$$V_1(y) = \sum_{t=1}^{T} \rho^{t-1}\hat{\eta}_t - c_1 y + \sum_{t=1}^{T-1} \rho^t c_{t+1}\mu_t.$$

Combining with Lemma 4.1 and Observation 2 completes the proof. □

7.7. Supplemental Proof of Lemma 4.2

Before proving Lemma 4.2, it will be useful to review some facts surrounding the distributionally robust DP equations, and their connection to a certain minimax optimization problem over a rectangular set of probability measures. In particular, let us define $\widehat{\mathfrak{M}}$ to be the set of all joint distributions Q for $D_{[T]}$ such that

$$D_t \in \mathfrak{P}(\mathcal{I}_t) \quad , \quad \mathbb{E}_Q[D_t|D_{[t-1]}] = \mu_t \quad , \quad \mathbb{E}_Q[D_t^2|D_{[t-1]}] = \mu_t^2 + \sigma_t^2, \quad t = 1, \ldots, T.$$

Then as discussed in Section 7.1, and as is generally well-known from the literature on rectangularity (see e.g. (Shapiro et al. 2009, section 6.7.3)) and Shapiro (2016), the connection between the minimax problem

$$\inf_{\pi \in \Pi(y_1)} \sup_{Q \in \widehat{\mathfrak{M}}} \mathbb{E}_Q\big[Z^\pi\big], \tag{63}$$

and the distributionally robust DP formulation, is as follows.

Observation 9 *The optimal value of Problem* (45) *equals* $V_1(y_1)$*. Any policy* π *such that* $x_t^\pi(d_{[t-1]}) \in \mathfrak{Y}_t\big(y_t^\pi(d_{[t-1]})\big)$ *for all* $t = 1, ..., T$ *and* $d_{[t-1]} \in \mathbb{R}_+^{t-1}$*, is an optimal solution to Problem* (63)*. Conversely, for any optimal policy* π *for Problem* (63)*, and any measure* $Q \in \arg\max_{Q \in \widehat{\mathfrak{M}}} \mathbb{E}_Q[Z^\pi]$*, it holds w.p.1 that* $x_t^\pi(D_{[t-1]}) \in \mathfrak{Y}_t\big(y_t^\pi(D_{[t-1]})\big)$ *for all* $t = 1, ..., T$*.*

We now complete the proof of Lemma 4.2.

Proof : Let $\pi \in \mathrm{OPT}(y_1)$ denote a base-stock policy with monotone increasing base-stock levels (with initial level at least y_1) $\{c_t, t \in [1, T]\}$. For $t \in [1, T]$, let Q_t^* denote any fixed measure in $\Gamma_t^{c_t} = \arg\max_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(c_t, D_t)]$, where non-emptiness follows from Theorem 1, and $Q^* := \otimes_{j=1}^T Q_j^*$. Then it follows from a straightforward proof by contradiction, the details of which we omit and the logic of which follows similarly to the proof of Theorem 2, that Q^* belongs to both $\arg\max_{Q \in \mathfrak{M}} \mathbb{E}_Q\big[Z^\pi\big]$ and $\arg\max_{Q \in \widehat{\mathfrak{M}}} \mathbb{E}_Q\big[Z^\pi\big]$. Combined with the facts that $\pi \in \mathrm{OPT}(y_1)$ and $\mathfrak{M} \subseteq \widehat{\mathfrak{M}}$, we conclude that π is also an optimal solution to Problem (63), and that both Problems (9) and (63) have the same value. Combining with Observation 9 completes the proof.

7.7.1. Additional discussion. It is natural to ask whether positing that the distributionally robust DP (as opposed to the static formulation) has an optimal policy with monotone increasing base-stock levels (but which themselves do not satisfy the conditions of Theorem 2) has any such implications for time consistency. We leave this as an interesting open question, as the approach we take to proving Lemma 4.2 seems unable to generalize in this way. Indeed, the key logic used in our proof of Lemma 4.2 is that: 1. the distributionally robust DP can be framed as an appropriate minimax optimization; 2. any monotone base-stock policy takes the same value under this minimax problem and the static formulation; 3. the value of the static formulation is upper-bounded by the value of this minimax problem; and thus 4. optimality for the static formulation implies optimality for this minimax problem (corresponding to the distributionally robust DP). If instead one posits that the monotone base-stock policy is optimal for the distributionally robust DP, the chain of inequalities breaks, and it seems that a fundamentally different approach is needed.

7.8. Acknowledgments

The authors sincerely thank Alex Shapiro for initiating this research, and his continued support and insights throughout this project. Research of David A. Goldberg was partly supported by NSF award CMMI 1757394.