NEW APPROACHES TO INVENTORY CONTROL: ALGORITHMS, ASYMPTOTICS AND ROBUSTNESS

A Thesis
Presented to
The Academic Faculty

by

Linwei Xin

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Industrial and Systems Engineering

Georgia Institute of Technology
August 2015

NEW APPROACHES TO INVENTORY CONTROL: ALGORITHMS, ASYMPTOTICS AND ROBUSTNESS

Approved by:

Professor David A. Goldberg, Advisor
School of Industrial and Systems
Engineering
Georgia Institute of Technology

Professor Alexander Shapiro, Advisor
School of Industrial and Systems
Engineering
Georgia Institute of Technology

Professor Shabbir Ahmed
School of Industrial and Systems
Engineering
Georgia Institute of Technology

Professor Jim Dai
School of Operations Research and
Information Engineering
Cornell University

Professor Alan Scheller-Wolf
Tepper School of Business
Carnegie Mellon University

Date Approved: April 14, 2015

To Huijun, my wife

Guihuan and Xiaoyan, my parents

ACKNOWLEDGEMENTS

First of all, I want to express my deepest gratitude to my advisors Professors David A. Goldberg and Alex Shapiro, for all their help, support, and guidance over the last five years. I have been very lucky and honored to work with them on exciting problems. Thinking about hard problems with David (day and night) is one of most rewarding and enjoyable things during my Ph.D.. His true passion about research has always inspired me to be a better researcher. I am extremely grateful to his dedication and patience, and have learned tremendously from both his mentoring and personality. All my knowledge on stochastic programming comes from Alex (and his great books). He introduced me to the broad field of optimization under uncertainty and led our first work on stochastic inventory control. His knowledge and wisdom have been constantly shaped my point of view.

I am much indebted to all my committee members, Professors Shabbir Ahmed, Jim Dai, and Alan Scheller-Wolf. Shabbir has been a tremendous source of support and help. I sincerely appreciate his continuous encouragement and insightful career advice. Jim has provided me with invaluable guidance during my studies. I am very grateful for his constructive feedback and generous help during my job searching. I benefited a lot from Alan's visit to ISyE in April 2014. A chapter of this thesis was indeed inspired by our discussion, and my work is much enriched with his knowledge and insights.

Besides my thesis committee, I want to thank many outstanding faculty members at Georgia Tech, who taught, guided, and helped me in various ways. In particular, I want to thank Professors Sigrun Andradottir, Hayriye Ayhan, Renato Monteiro, George Nemhauser, and Arkadi Nemirovski, for lecturing me unforgettable courses

TABLE OF CONTENTS

LIST OF TABLES

SUMMARY

The fundamental problem of managing an inventory over time in the presence of stochastic demand is one of the core problems of operations research. This thesis is motivated by the need (in both government and industry) to understand such complex inventory systems used to model many of society's most important problems. In particular, we investigate *simple*, *efficient* and *robust* inventory policies for several fundamental models commonly used in the study of stochastic inventory systems. Some of these policies have been already implemented in practice and we provide strong theoretic support for their practical utilization in industry. Furthermore, the results on the performance of these policies often yield a rule-of-thumb that is applicable in a variety of settings.

There are five chapters in this thesis. In the first chapter, we provide the overall motivation and problem formulations, and summarize our main contributions. Chapter 2 - 5 constitute the main body of the thesis. In the second chapter, we study lost sales inventory model with positive lead time. We significantly improve the bound on the rate of convergence of constant-order policies in the lead time. In particular, we prove that a simple constant-order policy actually converges exponentially fast to optimality as the lead time grows. In addition, our bound is simple and explicit, demonstrating good performance of constant-order policies for realistic lead time values. Our results provide theoretical justification for the good performance of such simple policies, and open the window to making the results and methodology practical.

In the third chapter, we investigate dual-sourcing inventory systems. These systems are notoriously difficult to optimize due to the complex structure of the optimal solution and the curse of dimensionality. Recently, so-called Tailored Base-Surge (TBS) policies have been proposed as a heuristic for the dual-sourcing problem. Although numerical experiments by several authors have suggested that such policies perform well as the lead time difference between the two sources grows large, providing a theoretical foundation for this phenomenon has remained a major open problem. We provide such a theoretical foundation by proving that a simple TBS policy is indeed asymptotically optimal as the lead time of the regular source grows large, with the lead time of the express source held fixed. Since many companies are already implementing such TBS policies, our results provide strong theoretical support for the widespread use of TBS policies in practice.

In the fourth chapter, we explore the concept of time consistency in the context of distributionally robust inventory models with second moment constraints. Recently, several communities have observed that a subtle phenomena known as time inconsistency, which never happens in the classic (non-robust) setting, can arise in the framework of distributionally robust optimization. In particular, there have been two fundamentally different formulations (i.e., the multistage-dynamic and multistage-static formulations) proposed in the literature, depending on whether the underlying optimization model is static or dynamic in nature. We provide several illustrative examples showing that here the question of time consistency can be quite subtle and complement these observations by providing simple sufficient conditions for time consistency. We also prove that, although the multistage-dynamic formulation always has an optimal policy of base-stock form, there may be no such optimal policy for the multistage-static formulation. Interestingly, our results show that time consistency may hold even when rectangularity does not.

In the fifth chapter, we study distributionally robust inventory control with martingale demand. Although distributionally robust inventory models have been analyzed previously, the cost and policy implications of positing different dependency structures remains poorly understood. We combine the framework of distributionally robust optimization with the theory of martingales, and study a novel distributionally robust model in which the sequence of future demands is assumed to belong to a family of martingales. We explicitly compute the optimal policy and shed light on the interplay between the optimal policy and worst-case distribution. We also compare to the analogous setting in which demand is independent across periods. Our results shed light on several intriguing phenomena regarding the impact of correlations on distributionally robust models, and provide a first step towards developing a conditional-expectation based theory of dynamic distributionally robust forecasting.

CHAPTER I

INTRODUCTION

1.1 Motivation

The fundamental problem of managing an inventory over time in the presence of stochastic demand is one of the core problems of operations research. This thesis is motivated by the need (in both government and industry) to understand such complex inventory systems used to model many of society's most important problems. Such models, which arise in applications as diverse as military and sustainable operations (cf. [80]), the provisioning of renewable energy resources (cf. [141]), health care operations (cf. [144]), the management of global supply chains (cf. [134]), and cloud computing (cf. [107]), are often characterized by their high-dimensionality, uncertainty, and complicated dependency structure. Unfortunately, a perfect understanding of such systems seems beyond our reach - many associated problems have been proven to be computationally intractable, or their solutions have complicated structure. Furthermore, many of the approximations developed in the literature to understand real-world inventory systems are themselves quite complicated. How can we expect policy-makers to gain insight from, not to mention implement such results?

The approach we have taken to answer the above question in this thesis is to investigate *simple*, *efficient* and *robust* inventory policies for several fundamental models commonly used in the study of stochastic inventory systems, including lost-sales inventory systems, dual-sourcing inventory systems, and distributionally robust inventory systems. Some of these policies have been already implemented in practice and we provide strong theoretic support for their practical utilization in industry.

Furthermore, the results on the performance of these policies often yield a rule-of-thumb that is applicable in a variety of settings.

1.2 Model formulation

In this section, we describe the fundamental models studied in this thesis.

1.2.1 Single-sourcing inventory model

Let D_t represent the demand distribution in period t, $t \geq 1$. Let T be the time horizon, representing the total number of time periods we consider. Let L be the deterministic lead time, i.e., a multi-period delay between when an order for more inventory is placed and when that order is received. Let c_t, h_t, b_t be the unit ordering cost, holding cost and under-stocking penalty in period t respectively. In addition, let I_t denote the on-hand inventory, and $\mathbf{x}_t = (x_{1,t}, \ldots, x_{L,t})$ denote the pipeline vector of orders placed but not yet delivered, at the beginning of time period t, where $x_{i,t}$ is the order to be received in period $i + t - 1$. The ordered sequence of events in period t is then as follows.

- A new amount of inventory $x_{1,t}$ is delivered and added to the on-hand inventory;

- A new order is placed;

- The demand D_t is realized;

- Costs for period t are incurred, and the on-hand inventory and pipeline vector are updated.

Note that the on-hand inventory is updated according to

$$
I_{t+1} = \begin{cases} I_t + x_{1,t} - D_t & \text{if unmet demand can be backlogged;} \\ \max(0, I_t + x_{1,t} - D_t) & \text{if unmet demand is lost forever.} \end{cases}
$$

We call the former model a backlog inventory model and the latter one a lost-sales inventory model. The pipeline vector is updated such that $x_{1,t}$ is removed, $x_{i,t+1}$ is

set equal to $x_{i+1,t}$ for $i \in [1, L-1]$, and $x_{L,t+1}$ is set equal to the new order placed. We require that this new order $x_{L,t+1}$ be a function of realized demands, but cannot depend on future demands. We call the corresponding family of policies *admissible*, and denote this family by Π. Define C_t to be the sum of the ordering cost, holding cost and under-stocking penalty in time period t:

$$C_t \triangleq c_t x_{L,t+1} + h_t \left(I_t + x_{1,t} - D_t\right)^+ + b_t \left(I_t + x_{1,t} - D_t\right)^-,$$

where $x^+ \triangleq \max(x, 0)$, $x^- \triangleq \max(-x, 0)$. Then the problem is to find an admissible inventory control policy that minimizes the total discounted cost over a finite time horizon T, i.e.,

$$\min_{\pi \in \Pi} \sum_{t=1}^{T} \rho^{t-1} \mathbb{E}\left[C_t^\pi\right],$$

where $\rho \in (0, 1]$ is a discount factor and C_t^π is the cost incurred in period t under policy π, or minimizes the long-run average cost, i.e.,

$$\min_{\pi \in \Pi} \limsup_{T \to \infty} \frac{\sum_{t=1}^{T} \mathbb{E}\left[C_t^\pi\right]}{T}.$$

1.2.2 Dual-sourcing inventory model

In a dual-sourcing inventory model with backlogging, the decision-maker has the choice to order from two different sources, the regular source (R) with longer lead time but lower per-unit ordering cost, and the express source (E) with shorter lead time but higher per-unit ordering cost. Let $L \geq 1$ be the deterministic lead time of the regular source (R), and $L_0 \geq 0$ the deterministic lead time of the express source (E), where $L \geq L_0 + 1$. Let c_R, c_E be the unit purchase costs of the regular and express sources, and h, b be the unit holding and backorder costs respectively, with $c \triangleq c_E - c_R > 0$. In addition, let I_t denote the on-hand inventory at the start of period t (before any orders or demands are received), and $q_t^R (q_t^E)$ denote the order placed from R(E) at the beginning of period t. Note that due to the leadtimes, the order received from R(E) in period t is $q_{t-L}^R (q_{t-L_0}^E)$. As a notational convenience, we define

3

$q_k^R = 0, k = -(L-1), \ldots, 0$; and $q_k^E = 0, k = -(L_0 - 1), \ldots, 0$. For $t = 1, \ldots, T$, the events in period t are ordered as follows.

- Ordering decisions from R and E are made (i.e. q_t^E, q_t^R are chosen);

- New inventory $q_{t-L}^R + q_{t-L_0}^E$ is delivered and added to the on-hand inventory;

- The demand D_t is realized, costs for period t are incurred, and the inventory is updated.

Note that the on-hand inventory is updated according to $I_{t+1} = I_t + q_{t-L}^R + q_{t-L_0}^E - D_t$, and may be negative since backorder is allowed. We require that the new orders q_t^R and q_t^E are non-negative measurable (and thus deterministic) functions of the realized demands. We call the corresponding family of policies *admissible*, and denote this family by Π. Let $G(y)$ be the sum of the holding and backorder costs when the inventory level equals y in the end of a time period, i.e. $G(y) \overset{\Delta}{=} hy^+ + by^-$, where $x^+ \overset{\Delta}{=} \max(x, 0)$, $x^- \overset{\Delta}{=} \max(-x, 0)$. Let C_t be the sum of the ordering, holding and backorder costs incurred in time period t, i.e.

$$C_t \overset{\Delta}{=} c_R q_t^R + c_E q_t^E + G(I_t + q_{t-L}^R + q_{t-L_0}^E - D_t).$$

Then the problem is to find an admissible inventory control policy that minimizes the long-run average cost.

1.2.3 Distributionally robust inventory model

1.2.3.1 Single-period

The models in previous sections require a complete specification of the underlying demand distribution $\{D_t\}_{t\geq 1}$. 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

4

the distributionally robust optimization paradigm, one assumes that the joint distribution (over time) of the sequence of future demands belongs to some set of joint distributions, and solves the min-max problem of computing the control policy which is optimal against a worst-case distribution belonging to this set.

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 the single-period distributionally robust newsvendor problem can be formulated as below:

$$\min_{x \geq 0} \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q \left[cx + b[d-x]_+ + h[x-d]_+ \right],$$

where c, b, h are the per unit ordering, backorder penalty, and holding costs respectively, and the notation $\mathbb{E}_Q$ emphasizes that the expectation is taken with respect to the distribution Q of the demand D.

1.2.3.2   Multi-period

In the distributionally robust setting, we assume backlogging and zero ordering lead time, i.e., orders are delivered immediately. Let c_t, b_t, h_t be the per unit ordering, backorder penalty, and holding costs in period t respectively. Let D_t be the demand in period t, and x_t, y_t be the inventory level at period t before and after placing an order respectively, $t = 1, \ldots, T$, where we note that the order must be placed in period t before D_t is known. We will consider policies which are nonanticipative, i.e. decisions do not depend on realizations of future demand. Let Π denote the family of all such policies. We assume that x_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, i.e., $y_t \geq x_t$. The inventory update is according to

$$x_{t+1} = y_t - D_t, \ t = 1, ..., T-1,$$

and the cost incurred in period t equals

$$C_t \overset{\Delta}{=} c_t(y_t - x_t) + h_t[y_t - D_t]_+ + b_t[D_t - y_t]_+.$$

5

Let $\mathfrak{M}$ be a collection of T-dimensional demand distributions. Then the multi-period distributionally robust inventory problem can be formulated as below:

$$\min_{\pi \in \Pi} \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q[\sum_{t=1}^{T} C_t^\pi]. \tag{1}$$

In particular, if the set $\mathfrak{M}$ is a singleton, then it is reduced to a classic (non-robust) multi-period problem.

1.3 Problem formulation and literature review

In this section, we formally describe the main problems that we will address in this thesis, and review the relevant literature.

1.3.1 Lost sales

It is a classical result that a so-called base-stock (i.e. order-up-to) policy, based only on the total inventory position (i.e. sum of the current inventory and all orders in the pipeline vector), is optimal in backlog models (cf. [170], [99], [193]). However, it is known that such simple policies are no longer optimal for models with lost sales and positive lead times (cf. [113]). Although the model has been studied now for over fifty years, the optimal policy remains poorly understood (cf. [74], [137], [206], [138], [139], [153], [140], [49], [110], [109], [109]), and we refer to [24] for a comprehensive review. Furthermore, *for over fifty years, inventory models with lost sales and positive lead times were generally considered intractable, as the primary solution method (dynamic programming) suffered from the curse of dimensionality as the lead time grew.* As noted in [24], this has led to many researchers using models with backlogging as approximations for settings in which a lost sales assumption is more appropriate, which may lead to very suboptimal solutions. Due to the difficulty of computing the optimal policy, there has been considerable focus on understanding structural properties of an optimal policy, and analyzing heuristics.

A very simple and natural policy, which will be the subject of our own investigations, is the so-called *constant-order policy*, which places the same order in every period, independent of the state of the system. Perhaps surprisingly, [158] proved that for lost sales inventory models with positive lead times, sometimes the best constant-order policy outperforms the more sophisticated base-stock policy, and performed a detailed analysis under a certain asymptotic scaling. This phenomena was further illuminated by the computational study of [214], which confirmed that in several scenarios the constant-order policy performed favorably. In all of their experiments, the constant-order policy always incurred an expected cost at most twice that incurred by the optimal policy; in 62.5% of the cases, it incurred a cost at most 1.33 times that incurred by the optimal policy; and in 38% of the cases, it incurred a cost at most 1.12 times that incurred by the optimal policy.

These observations were recently given a solid theoretical foundation by [75], who proved that for lost sales inventory models with positive lead times, as the lead time grows with all other parameters remaining fixed, the best constant-order policy is in fact *asymptotically optimal*. This is quite surprising, as the policy is so simple, and performs nearly optimally exactly in the setting which had stumped researchers for over fifty years. However, the bounds proven there are impractical, requiring the lead time to be very large before the constant-order policy becomes nearly optimal, e.g. requiring a lead time which is $\Omega(\epsilon^{-2})$ to ensure a $(1 + \epsilon)$-approximation guarantee, and involving a massive prefactor. The authors note that the numerical experiments of [214] suggest that the constant-order policy performs quite well even for small lead times, and pose closing this gap (thus making the results practical) as an open problem. The authors also point out that if one could prove that the constant-order policy performs well even for small to moderate lead times, this would open the door for the creation of practical hybrid algorithms, which solve large dynamic programs when the lead time is small, and gradually transition to more naive algorithms for

larger lead times.

1.3.2 Dual-sourcing

Although dual-sourcing strategy is attractive and very relevant to practice, optimizing a dual-sourcing inventory system is notoriously challenging. Such inventory systems have been studied now for over forty years and there is a vast literature investigating periodic review dual-sourcing inventory models as well as their variants, but the structure of the optimal policy remains poorly understood, with the exception of when the system is consecutive, i.e., the lead time difference between the two sources is exactly one (cf. [9], [45], [143], [67], [199]). Furthermore, it is well known that a dual-sourcing inventory system can be regarded as a generalization of a lost-sales inventory system (cf. [183]). Indeed, the intractability of both the dual-sourcing and lost-sales inventory models has a common source - as the lead time grows, the state-space of the natural dynamic programming (DP) formulation grows exponentially, rendering such techniques impractical.

As an exact solution seems out of reach, the operations research and management communities have instead investigated certain structural properties of the optimal policy (cf. [91]), and exerted considerable effort towards constructing various heuristic policies (cf. [191], [173], [183], [29]). A simple and natural policy that is implemented in practice, which will be the subject of our own investigations, is the so-called *Tailored Base-Surge* (TBS) policy. It was first proposed and analyzed in [3], where we note that closely related standing order policies had been studied previously (cf. [162, 106]). Under such a TBS policy, a constant order is placed at the regular source in each period to meet a *base* level of demand, while the orders placed at the express source follow an order-up-to rule to manage demand *surges*. We refer to Mini-Case 6 in [131] for more about the motivation and background of TBS policies. Note that dual-sourcing inventory systems in which a constant-order policy is implemented for

the regular source are essentially equivalent to single-sourcing inventory systems with constant returns, which have been investigated in the literature (cf. [64], [46]).

[3] analyzed TBS policies in a continuous review model, and their focus was to find the best TBS policy. Numerical results in [118], [163] showed that TBS policies are comparable to DI policies, and outperform DI policies for some problem instances. [3] conjectured that this policy performs more effectively as the lead time difference between the two sources grows. [105] analyzed a periodic review model and studied the performance of TBS policy. They provided an explicit bound on the performance of TBS policies compared to the optimal one when the demand had a specific structure, and provided numerical experiments suggesting that the performance of the TBS policy improves as the lead time difference grows large. However, to date there is no theoretical justification for the good behavior of TBS policies as the lead time difference grows large, and giving a solid theoretical foundation to this observed phenomena remains a major open question.

1.3.3 Time consistency

In the classical inventory control setting, the problem is stated as a minimization of the expected value of the relevant ordering, backorder, and holding costs. Such a formulation requires a complete specification of the probability distribution of the underlying demand process. 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 [168] 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. [71, 72, 69, 73, 151, 208, 70, 148, 37, 174, 85, 212]).

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 (see, e.g., [213] and the references therein). Recently, distributionally robust variants of such multistage problems have begun to receive attention in the literature (cf. [73, 2, 43, 174, 180, 117]). 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 [187] - for a more recent overview we refer the reader to the recent survey by [57], and the references therein. Questions of time consistency have also attracted attention in the mathematical finance literature, in the context of assessing the risk and value of investments over time, and have played an important role in the associated theory of coherent risk measures (cf. [196, 6, 161, 42, 165]). These concepts have also been studied from the perspective of robust control across various academic communities (cf. [86, 103, 145, 79, 32, 200]), and have also begun to receive attention in the setting of inventory control (cf. [38, 39, 205]).

Recently, [180] proved that for the setting in which only the mean and support of the demand are known, such distributionally robust inventory control problem is always time consistent. However, it is still an open challenge to understand the question of time consistency in other distributionally robust settings.

1.3.4 Martingale demand

In many practical settings of interest, demands are correlated over time (cf. [98, 169, 171]). As a result, there is a vast literature investigating inventory models with correlated demand, including: studies of the so-called bull-whip effect (cf. [35, 122, 167]); models with Markov-modulated demand (cf. [60, 98, 112]); and models with

forecasting, including models in which demand follows an auto-regressive/moving average (ARMA) or exponentially smoothed process (cf. [8, 26, 73, 111, 127, 132, 150]); and models obeying the Martingale Model of Forecast Evolution (MMFE) and its many generalizations (cf. [56, 77, 87, 101, 128, 189]). Although several of these works offer insights into the qualitative impact of correlations on the optimal policy (and associated costs) when managing an inventory over time, these results are typically proven under very particular distributional assumptions, which assume perfect knowledge of all relevant distributions.

As we discussed in the previous section, recently there has been a growing interest in developing inventory control policies which are robust to model misspecification. There are several works which formulate dynamic programming approaches to distributionally robust/risk averse inventory models (cf. [2, 73, 180, 204]). More generally, such dynamic problems can typically be formulated as so-called robust Markov decision processes (MDP) (cf. [103, 145, 200]). However, to our knowledge, none of these works consider applications to correlated demand or forecasting models, with the exception of the very general Bayesian model considered recently in the excellent work of [117]. Furthermore, there seems to have been no systematic study of the *qualitative* impact of positing different joint dependency structures in such multi-stage distributionally robust inventory control problems, i.e. seeing which insights previously derived under specific distributional assumptions extend to the distributionally robust setting, and furthermore what new insights manifest only in the distributionally robust setting. The quest to develop such an understanding in the broader context of stochastic optimization (not specifically inventory control) was recently initiated in [1], where the authors define the so-called *price of correlations* as the ratio between the optimal minimax value when all associated random variables (r.v.) are independent, and the setting where they may take any joint distribution belonging to the allowed family. Although the authors do not look specifically at any inventory

problems, they stress the general importance of understanding how positing different joint distribution uncertainty impacts the underlying stochastic optimization.

Combining the above, we are led to the following questions:

1. Can we construct effective dynamic distributionally robust variants of the time series and forecasting models used in Operations Research?

2. Can we develop a theory of how positing different correlation structures qualitatively impacts the optimal policy for such models?

1.4   Main contributions of this thesis

1.4.1   Lost sales

In chapter 2, we make significant progress towards resolving the open problem posed in section 1.3.1. In particular, for the infinite-horizon variant of the finite-horizon problem considered by [75], we prove that the optimality gap of the same constant-order policy actually converges *exponentially fast* to zero, i.e. we prove that a lead time which is $O\big(\log(\epsilon^{-1})\big)$ suffices to ensure a $(1 + \epsilon)$-approximation guarantee. We demonstrate that the corresponding rate of exponential decay is at least as fast as the exponential rate of convergence of the expected waiting time in a related single-server queue to its steady-state value, which we prove to be monotone in the ratio of the lost-sales penalty to the holding cost. We also derive simple and explicit bounds for the optimality gap. For the special case of exponentially distributed demand, we further compute all expressions appearing in our bound in closed form, and numerically evaluate them, demonstrating good performance for a wide range of parameter values. Our main proof technique combines convexity arguments with ideas from queueing theory, and is simpler than the coupling argument of [75].

1.4.2 Dual-sourcing

In chapter 3, we resolve the open question posed in section 1.3.2, by proving that, when the lead time of the express source is held fixed, a simple TBS policy is asymptotically optimal as the lead time of the regular source grows large. Our results provide a solid theoretical foundation for the conjectures and numerical experiments of [3] and [105]. Interestingly, the simple TBS policy performs nearly optimally exactly when standard DP-based methodologies become intractable due to the aforementioned "curse of dimensionality". Furthermore, as the "best" TBS policy can be computed by solving a convex program that does not depend on the lead time of the regular source (cf. [105]), our results lead directly to very efficient algorithms (with complexity independent of the lead time of the regular source) with asymptotically optimal performance guarantees. Perhaps most importantly, since many companies are already implementing such TBS policies (cf. [3]), our results provide strong theoretical support for the widespread use of TBS policies in practice. Our main proof technique combines a steady-state approach, novel convexity and lower-bounding arguments, a certain interchange of limits result, and ideas from the theory of random walks and queues, significantly extending the methodology and applicability of a novel framework for analyzing inventory models with large lead times recently introduced in [75] and [202] in the context of lost-sales models with positive lead times.

1.4.3 Time consistency

In chapter 4, 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 [79], 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 [54]), albeit in a substantially different context motivated by

questions in decision theory and artificial intelligence.

We extend the work of [168] (and followup work of [73]) 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. 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 dynamic formulations. We also prove that, although the multistage-dynamic 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 [180] for the setting in which only the mean and support of the demand distribution are known, where the problem is always time consistent, amenable to a simple dynamic programming 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 [15], where again time consistency always holds. Here our model is rich enough to exhibit a variety of interesting behaviors, including both time inconsistency, as well as strong time consistency even when no dynamic programming formulation exists, and the two formulations take different values and hence the rectangularity property does not hold.

1.4.4 Martingale demand

In chapter 5, we take a first step towards answering the questions raised in section 1.3.4, by analyzing in-depth a distributionally robust multi-stage inventory optimization problem which naturally generalizes and unifies a non-trivial family of inventory models with demand governed by simple time-series and forecasting. In particular, perhaps the simplest family of time-series and forecasting models are those in which demand is assumed to evolve as a one-dimensional (additive or multiplicative) random walk, with increments that form a martingale difference sequence. Such a forecasting model represents a special case of the MMFE framework, in which all future demand forecasts are updated by the same (random) amount in each period. We note that special cases of this model, e.g. that in which demand evolve as a Gaussian random walk with independent increments, have been studied previously (cf. [133]). Furthermore, related minimax optimization problems have been considered more broadly within the economics, finance, and robust control communities (cf. [10, 41, 58, 86, 146, 160, 190]). Intuitively, the setting we consider provides a unified, distribution-free approach to models with the property that the expected demand in period $t + 1$ equals the realized demand in period t, for all t.

More formally, we consider a distributionally robust multi-period inventory control problem (with backlogged demand) in which the sequence of future demands is assumed to belong to the set of all martingales with given mean and support (assumed the same in every period). Our contributions are three-fold. First, we explicitly compute the optimal policy in closed form, which is of state-dependent base-stock form. Our main proof technique involves a non-trivial induction, combining ideas from convex analysis and probability. Second, we shed light on the interplay between the optimal policy and worst-case distribution. In particular, we show that at optimality, in each period the adversary always puts some probability at 0, and some probability on a different quantity. Combined with the martingale property, this implies that at

15

optimality, the worst-case demand distribution corresponds to the setting in which demand may become obsolete at a random time, a scenario of practical interest which has been studied previously in the literature (cf. [31, 34, 108, 149, 184]). We also compute the limiting dynamics as the time horizon diverges, by proving convergence to an appropriate weak limit. Third, in the spirit of [1], we compare to the analogous setting in which demand is independent across periods (analyzed previously in [180]), and identify qualitative differences between these two models. In particular, we show that the cost incurred by an optimal policy in the martingale-demand model is always no greater than the corresponding cost in the independent-demand model, and their limiting ratio is exactly $\frac{1}{2}$ in the perfectly symmetric case.

1.5   Conclusion

In this thesis, we studied simple, efficient and robust inventory policies for several fundamental models commonly used in the study of stochastic inventory systems. We proved several asymptotic results on the performance of simple policies in lost-sales and dual-sourcing inventory systems. We analyzed the phenomenon of time (in)consistency in the context of distributionally robust inventory and provided several illustrative examples showing that here the question of time consistency can be quite subtle. We also proposed a novel multi-period inventory model by combining the framework of distributionally robust optimization with the theory of martingales and shed light on the interplay between the optimal policy and worst-case distribution, which served as the first step towards establishing a conditional-expectation based theory of dynamic distributional robust forecasting.

This thesis leaves several interesting directions for future research. In chapters 2 and 3, our methodology lays the foundations for a completely new approach to analyzing inventory models with large lead times. So far, this approach has been successful in yielding key insights and efficient algorithms for two settings previously

believed intractable: lost-sales models with large lead times, and dual-sourcing models with large lead time gap. We believe that our techniques have the potential to make similar progress on many other difficult supply chain optimization problems of practical relevance in which there is a lag between when policy decisions are made and when those decisions are implemented.

The general question of time consistency remains poorly understood. In addition, our work in chapter 4 has shown that this question can be quite subtle. For the particular model we consider here, it would be interesting to develop a better understanding of precisely when time consistency holds. Of course, it is still an open challenge to understand the question of time consistency more broadly, how precisely the various definitions of time consistency presented throughout the literature relate to one-another, and more generally to understand the relationship between different ways to model multistage optimization under uncertainty.

We believe that the framework built in chapter 5 towards establishing a conditional-expectation based theory of dynamic distributional robust forecasting. It would be interesting to consider more general conditional moment constraints. Furthermore, although it is often clear how to specify the marginal distribution in each time period, understanding the effects of positing various joint distributions over time remains an interesting challenge. It would be interesting to develop a deeper understanding of such price of correlations in robust stochastic optimization problems.

CHAPTER II

OPTIMALITY GAP OF CONSTANT-ORDER POLICIES DECAYS EXPONENTIALLY IN THE LEAD TIME FOR LOST SALES MODELS

This chapter is based on [202].

2.1 Introduction and literature review

It is well-known that there is a fundamental dichotomy in the theory of inventory models, depending on the fate of unmet demand. If unmet demand remains in the system and can be met at a later time, we say the system exhibits *backlogged demand*; if unmet demand is lost to the system, we say the system exhibits *lost sales*. Which of these assumptions is appropriate depends heavily on the application of interest. For example, in many retail applications one must manage an inventory in a competitive environment, i.e. demand can in principle be met by a competing supplier, making lost sales a more appropriate assumption. Indeed, as pointed out in [24], recent studies have shown that retailers across many sectors lose over 85% of the potential demand which they cannot satisfy immediately, and we refer the interested reader to [78], and [194] for further details.

A second important feature of many inventory models, intimately related to the above dichotomy, is that of positive lead times, i.e. settings in which there is a multi-period delay between when an order for more inventory is placed and when that order is received. In principle, this feature leads to an enlarged state-space (growing linearly with the lead time), to track all orders already placed but not yet received, i.e. the pipeline vector. It is a classical result, indeed one of the foundational results of the

field, that models with backlogged demand remain tractable even in the presence of positive lead times. Namely, it can be proven that a so-called base-stock (i.e. order-up-to) policy, based only on the total inventory position (i.e. sum of the current inventory and all orders in the pipeline vector), is optimal in this setting (cf. [170], [99], [193]). Intuitively, this follows from the fact that when demand is backlogged, inventory is a linear function of orders placed and past demands, along with certain convexity arguments. However, it is known that such simple policies are no longer optimal for models with lost sales and positive lead times (cf. [113]). *For over fifty years, inventory models with lost sales and positive lead times were generally considered intractable, as the primary solution method (dynamic programming) suffered from the curse of dimensionality as the lead time grew.* As noted in [24], this has led to many researchers using models with backlogging as approximations for settings in which a lost sales assumption is more appropriate, which may lead to very suboptimal solutions.

Although the optimal policy for lost-sales models with positive lead times remains poorly understood, the model has been studied now for over fifty years (cf. [74], [137], [206], [138], [139], [153], [140], [49], [110], [109], [109]), and we refer to [24] for a comprehensive review. Due to the difficulty of computing the optimal policy, there has been considerable focus on understanding structural properties of an optimal policy, and analyzing heuristics. In particular, convexity results were obtained in [113], [138], and [215], and used to bound the optimal ordering quantity. [104] compared the optimal costs between the backlogged and lost sales systems with identical problem parameters, and showed that the lost sales system always had a lower cost. [93] further proved that the base-stock policy was asymptotically optimal as the lost-sales penalty became large compared to the holding cost, and similar results were derived in [129]. In a breakthrough work, [123] proposed the family of so called dual-balancing policies, motivated by previous work on other relevant models (cf. [124], [126]), and

proved that the cost incurred by such a policy was always within a factor of 2 of optimal. Another recent line of research, based on carefully truncating and rounding the relevant dynamic programs, yields efficient approximation algorithms *for any fixed lead time*, with run-time polynomial in the inverse of the desired approximation error but possibly growing exponentially in the other problem inputs (e.g. the lead time) (cf. [83, 84, 36]). *Despite this progress, the aforementioned work leaves open the problem of deriving efficient algorithms with arbitrarily small error, when the lead time is large.*

A very simple and natural policy, which will be the subject of our own investigations, is the so-called *constant-order policy*, which places the same order in every period, independent of the state of the system. Perhaps surprisingly, [158] proved that for lost sales inventory models with positive lead times, sometimes the best constant-order policy outperforms the more sophisticated base-stock policy, and performed a detailed analysis under a certain asymptotic scaling. This phenomena was further illuminated by the computational study of [214], which confirmed that in several scenarios the constant-order policy performed favorably. In all of their experiments, the constant-order policy always incurred an expected cost at most twice that incurred by the optimal policy; in 62.5% of the cases, it incurred a cost at most 1.33 times that incurred by the optimal policy; and in 38% of the cases, it incurred a cost at most 1.12 times that incurred by the optimal policy.

These observations were recently given a solid theoretical foundation by [75], who proved that for lost sales inventory models with positive lead times, as the lead time grows with all other parameters remaining fixed, the best constant-order policy is in fact *asymptotically optimal*. This is quite surprising, as the policy is so simple, and performs nearly optimally exactly in the setting which had stumped researchers for over fifty years. However, the bounds proven there are impractical, requiring the lead time to be very large before the constant-order policy becomes nearly optimal, e.g.

requiring a lead time which is $\Omega(\epsilon^{-2})$ to ensure a $(1+\epsilon)$-approximation guarantee, and involving a massive prefactor. The authors note that the numerical experiments of [214] suggest that the constant-order policy performs quite well even for small lead times, and pose closing this gap (thus making the results practical) as an open problem. The authors also point out that if one could prove that the constant-order policy performs well even for small to moderate lead times, this would open the door for the creation of practical hybrid algorithms, which solve large dynamic programs when the lead time is small, and gradually transition to more naive algorithms for larger lead times.

2.1.1   Outline of chapter

The rest of the chapter is organized as follows. We formulate our problem, and introduce several elementary properties of the stationary inventory process, in Section 2.2.1. We describe the constant-order policy in Section 2.2.2, and review the results of [75] in Section 2.2.3. We state our main results in Section 2.2.4. We provide a more detailed analysis (both analytical and numerical) for the special case of exponentially distributed demand in Section 2.2.4.1, and discuss the monotonicity of our bounds in the ratio of the lost-sales penalty to the holding cost in Section 2.2.4.2. The proof of our main results are given in Section 2.3. Finally, we summarize our main results and propose directions for future research in Section 2.4. A technical appendix is provided in Section 2.5.

2.2   Main results

2.2.1   Model description, problem statement, and assumptions

In this section, we formally define our lost-sales inventory optimization problem. Note that the general framework of lost sales inventory model is already introduced in Section 1.2.1. Through this whole chapter, we assume that the demand process and all costs are stationary, and there is no ordering costs. Namely, $\{D_t, t \geq 1\}$ is

a sequence of independent and identically distributed (i.i.d.) demand realizations, distributed as the non-negative random variable (r.v.) D with distribution $\mathcal{D}$, which we assume to have finite mean, and (to rule out certain degenerate cases) to have strictly positive variance. In addition, $c_t = 0$, $h_t = h$ and $b_t = b$ for all t. Let us recall other notations here. Let T be the time horizon and L be the deterministic lead time. Let I_t denote the on-hand inventory, and $\mathbf{x}_t = (x_{1,t}, \ldots, x_{L,t})$ denote the pipeline vector of orders placed but not yet delivered, at the beginning of time period t, where $x_{i,t}$ is the order to be received in period $i + t - 1$. Recall that the ordered sequence of events in period t is then as follows.

- A new amount of inventory $x_{1,t}$ is delivered and added to the on-hand inventory;

- A new order is placed;

- The demand D_t is realized;

- Costs for period t are incurred, and the on-hand inventory and pipeline vector are updated.

Note that the on-hand inventory is updated according to $I_{t+1} = \max(0, I_t + x_{1,t} - D_t)$, and the pipeline vector is updated such that $x_{1,t}$ is removed, $x_{i,t+1}$ is set equal to $x_{i+1,t}$ for $i \in [1, L-1]$, and $x_{L,t+1}$ is set equal to the new order placed. We require that this new order $x_{L,t+1}$ be a (possibly random) function of realized demands, inventory levels, ordering quantities, and pipeline vectors, as well as the problem primitives $h, c, T, L, \mathcal{D}$ and current time t, but cannot depend on future demands. We call the corresponding family of policies *admissible*, and denote this family by Π. Define C_t to be the sum of the holding cost and lost-sales penalty in time period t:

$$C_t \triangleq h\left(I_t + x_{1,t} - D_t\right)^+ + b\left(I_t + x_{1,t} - D_t\right)^-,$$

where $x^+ \triangleq \max(x, 0)$, $x^- \triangleq \max(-x, 0)$. For simplicity, we suppose that the problem initial conditions are to start with the all zeros pipeline vector, i.e. $\mathbf{x}_1 = \mathbf{0}$, and zero

inventory, i.e. $I_1 = 0$. We note that our problem will differ from that considered in [75] in a single important way: we will consider the corresponding *infinite-horizon problem*, while [75] considered the finite-horizon problem. Namely, for a policy π, let $C(\pi)$ denote the long-run average cost incurred:

$$C(\pi) \triangleq \limsup_{T \to \infty} \frac{\sum_{t=1}^{T} \mathbb{E}[C_t]}{T}.$$

The corresponding infinite-horizon (i.e. long-run average cost) lost-sales inventory optimization problem is given by

$$\text{OPT}(L) \triangleq \inf_{\pi \in \Pi} C(\pi). \tag{2}$$

Recall that a stationary policy is one that places orders only based on the current state information (i.e., the on-hand inventory and pipeline vector), as well as the problem primitives $h, c, L, \mathcal{D}$, but *not* the current time period t or time horizon T. Under a stationary policy, the evolution of the on-hand inventory and pipeline vector evolves as a discrete time, finite-dimensional Markov Chain. It follows from the results of [95] that: an optimal policy for Problem 2 exists (i.e. is not simply approached), and furthermore that there always exists at least one such optimal policy which is stationary, so restricting oneself to the family of stationary policies is without loss of generality (w.l.o.g.). We will further assume that of these stationary optimal policies, there exists at least one such policy π^* whose corresponding induced Markov chain converges in distribution to a unique stationary measure when initialized with $\mathbf{x}_1 = \mathbf{0}$ and $I_1 = 0$. We will also assume that under this policy π^* (with the given initialization), $\mathbb{E}[\mathbf{x}_t], \mathbb{E}[I_t], \mathbb{E}[C_t]$ are finite for all t, and converge to the corresponding expected values under the given stationary measure (which we also assume to be finite), i.e. L^1 convergence. We refer to the set of such policies as convergent. Such a convergence is to be expected from the basic theory of Markov chains, and we refer the interested reader to [7] and [130] for further details. We note that this is especially so, in light of the results of [215], which demonstrate that there exists

an optimal stationary policy which, under the given initialization, with probability (w.p.) 1 belongs to a fixed compact set for all time (i.e. the ordering quantities and inventory levels are uniformly bounded over time as functions of $h, c, L, \mathcal{D}$ only). For any such stationary and convergent policy π, let $(\mathcal{I}^\pi, \chi^\pi)$ denote a vector distributed as the stationary measure of the corresponding Markov chain, with $\mathcal{I}^\pi$ corresponding to the stationary inventory level, and χ^π corresponding to the stationary pipeline vector.

2.2.2 Constant-order policy

In this section, we formally define the constant-order policy, and characterize the best constant-order policy. As a notational convenience, let us define all empty sums to equal zero, let $\mathbf{1}$ denote the vector with all entries equal to unity, e denote Euler's number, $\log(x)$ denote the natural logarithm of x, $\frac{1}{\infty}$ denote 0, $\frac{1}{0}$ denote ∞, $\log(\infty)$ denote ∞, and $\mathbb{I}(A)$ denote the indicator of the event A. For any $r \in [0, \mathbb{E}[D])$, the constant-order policy π_r is the policy that places the constant order r in every period. It is well-known (cf. [75]) that the corresponding steady-state on-hand inventory level, which we denote by I_∞^r, has the same distribution as the steady-state waiting time in the corresponding $GI/GI/1$ queue with interarrival distribution $\mathcal{D}$ and processing time distribution the constant r. For two r.v.s X, Y, let $X \sim Y$ denote equivalence in distribution between X and Y. In that case, it is well-known (cf. [7]) that

$$I_\infty^r \sim \sup_{j \geq 0} \left(jr - \sum_{i=1}^{j} D_i \right). \tag{3}$$

We note that in [75], the authors considered a slightly modified constant-order policy which ordered $I_\infty^r + r$ in the first period and r in all subsequent periods, to make the corresponding sequence of inventory levels stationary. As both policies have the same steady-state distribution, for our purposes this distinction is irrelevant.

We now formalize the notion of the best constant-order policy, and begin by briefly reviewing several well-known properties of the stationary inventory level under

any stationary and convergent policy π (not just the constant-order policy). As the inventory update dynamics imply that $\mathcal{I}^\pi \sim (\mathcal{I}^\pi + \chi_1^\pi - D)^+$, with D independent of $\mathcal{I}^\pi$ and χ_1^π, it follows that $\mathbb{E}[\mathcal{I}^\pi] = \mathbb{E}[(\mathcal{I}^\pi + \chi_1^\pi - D)^+]$. A straightforward algebraic manipulation further demonstrates that

$$\mathbb{E}\left[(\mathcal{I}^\pi + \chi_1^\pi - D)^-\right] = \mathbb{E}[D] - \mathbb{E}[\chi_1^\pi]. \tag{4}$$

Combining the above, we conclude that

$$\mathbb{E}[\chi_1^\pi] \le \mathbb{E}[D], \tag{5}$$

and

$$C(\pi) = h\mathbb{E}[\mathcal{I}^\pi] + b\mathbb{E}[D] - b\mathbb{E}[\chi_1^\pi]. \tag{6}$$

Customizing (3) - (6) to the constant-order policy, we conclude that for any $r \in [0, \mathbb{E}[D])$,

$$C(\pi_r) = h\mathbb{E}\left[\sup_{j \ge 0}\left(jr - \sum_{i=1}^{j} D_i\right)\right] + b\mathbb{E}[D] - br, \tag{7}$$

and note that $\mathbb{E}[\mathcal{I}^{\pi_r}] < \infty$ for all $r \in [0, \mathbb{E}[D])$ (cf. [7]). As it is easily verified that $C(\pi_r)$ is thus a convex function of r on $[0, \mathbb{E}[D])$, to find the best possible constant-order policy, it suffices to select the r minimizing this one-dimensional convex function over the compact set $[0, \mathbb{E}[D]]$. We note that the existence of at least one such optimal r follows from the well-known properties of convex optimization over a compact set, and that the set of all such optimal solutions must be bounded away from $\mathbb{E}[D]$, since by assumption $h > 0$, and $\lim_{r \uparrow \mathbb{E}[D]} \mathbb{E}[I_\infty^r] = \infty$ (since D has strictly positive variance, cf. [7]). Let $r_\infty \in \arg\min_{0 \le r \le \mathbb{E}[D]} C(\pi_r)$ denote the infimum of this set of optimal ordering quantities, in which case the best constant-order policy will refer to π_{r_∞}.

2.2.3 Review of results of [75]

In this section, we formally review the results of [75], and begin by introducing some additional notations. Let Q denote the $\frac{b}{b+h}$ quantile of the demand distribution, i.e.

$Q \triangleq \inf\{s \in \mathbb{R}^+ : \mathbb{P}(D > s) \leq \frac{h}{b+h}\}$. We note that Q is the optimal inventory level for the corresponding single-stage newsvendor problem, i.e. for any policy π and any time t, $\mathbb{E}[C_t^\pi] \geq g \triangleq h\mathbb{E}[(Q-D)^+] + b\mathbb{E}[(D-Q)^+]$ (cf. [213]). Equivalently, $g = \text{OPT}(0)$, the long-run optimal cost when there is zero leadtime, as in this case one can always order-up to this optimal level Q. That $g \in (0, \infty)$ follows from the assumption that D has finite mean and is not deterministic. Also, let σ denote the standard deviation of D, $\zeta \triangleq \mathbb{E}\big[|D - \mathbb{E}[D]|^3\big]\sigma^{-3}$ denote the so-called skewness of D,

$$m \triangleq \left\lceil \left(26\big(3\zeta + b(h\sigma)^{-1}\mathbb{E}[D] + 1\big) \right)^2 \right\rceil,$$

and $y(\epsilon) \triangleq$

$$\max\left(2^{14}h(Q + 2^{\frac{3}{2}}\mathbb{E}[D])(\mathbb{E}^2[D] + \mathbb{E}[D^2])^3\sigma^{-6}m^3g^{-1}\epsilon^{-1}, \left(12bg^{-1}\big((2bh^{-1})^{\frac{1}{2}} + 3\big) \right)^2 \epsilon^{-2} \right).$$

Then the main result of [75] is as follows. We only state the implications of those results for the infinite-horizon problem, as that is our focus in this chapter and to do otherwise would require several additional definitions and notations, but do note that the results of [75] also apply to the finite-horizon setting.

Theorem 1 *Suppose $\mathbb{E}[D^3] < \infty$. Then for all $\epsilon \in (0,1)$ and $L \geq y(\epsilon)$,*

$$\frac{C(\pi_{r_\infty})}{OPT(L)} \leq 1 + \epsilon.$$

Theorem 1 represented significant progress in our understanding of lost sales models with large lead times, as it proved that the simple constant-order policy performs well exactly when the problem becomes challenging to solve by dynamic programming and other means, i.e. when L becomes large. However, as discussed in [75], the explicit bounds of Theorem 1 require L to be so large as to make the results impractical. In addition to the massive prefactor, they require L to be $\Omega(\epsilon^{-2})$ to achieve a $(1+\epsilon)$-approximation, which requires e.g. a lead time on the order of 400 to be within 5% of optimal. As pointed out in [75], this massive prefactor and unfavorable scaling with

ϵ leave much to be desired, and are a far cry from the good numerical performance of the constant-order policy even for small lead times demonstrated in [214]. [75] pose tightening these bounds and closing this gap as a significant open question, as doing so would represent a large step towards making the bounds practical, e.g. proving that when L is small one can solve a large dynamic program, and when L becomes even moderately large one can use simple policies such as the constant-order policy.

2.2.4 Main results

In this section, we present our main results, demonstrating that the optimality gap of the best constant-order policy decays exponentially in the lead time. For $\theta \geq 0$, let us define

$$\phi(\theta) \stackrel{\Delta}{=} \exp(\theta r_\infty)\mathbb{E}[\exp(-\theta D)] \quad , \quad \gamma \stackrel{\Delta}{=} \inf_{\theta \geq 0} \phi(\theta),$$

and $\vartheta \in \arg\min_{\theta \geq 0} \phi(\theta)$ denote the supremum of the set of minimizers of $\phi(\theta)$, where we define ϑ to equal ∞ if the above infimum is not actually attained. Note that $\phi(\theta)$ is a continuous and convex function of θ on $(0, \infty)$, and right-continuous function of θ at 0. In addition, it follows from [65] Theorem 2.27 that $\phi(\theta)$ is right-differentiable at zero, with derivative equal to $r_\infty - \mathbb{E}[D]$, assuming only that $\mathbb{E}[D] < \infty$ (along with our default assumption of non-negativity). As $r_\infty < \mathbb{E}[D]$, we conclude from the definition of derivative and a straightforward contradiction argument that $\vartheta > 0$ (i.e. ϑ is strictly positive), and $\gamma \in [0, 1)$ (i.e. γ is strictly less than 1). It follows from the celebrated Cramér's Theorem, and more generally the theory of large deviations, that up to exponential order (and under appropriate technical assumptions), $\mathbb{P}(kr_\infty \geq \sum_{i=1}^{k} D_i)$ decays like γ^k as $k \to \infty$ (cf. [48]). Furthermore, as we will explore in detail later in the proof of our main result, γ corresponds (again up to exponential order, under appropriate assumptions) to the rate at which the expected waiting time in an initially empty single-server queue, with inter-arrival distribution D and processing time distribution (the constant) r_∞, converges to its steady-state value (cf. [116]).

Then our main result is as follows.

Theorem 2 *For all $L \geq 1$,*

$$\frac{C(\pi_{r_\infty})}{OPT(L)} \leq 1 + h\big((1-\gamma)g\big)^{-1}\bigg(\mathbb{E}[D] - r_\infty + \big(e\vartheta(L+1)\big)^{-1}\bigg)\gamma^{L+1}. \tag{8}$$

Our results prove that for the corresponding infinite-horizon problem, the optimality gap of the constant-order policy converges exponentially fast to zero. In particular, a lead time which is $O\big(\log(\epsilon^{-1})\big)$ suffices to ensure a $(1+\epsilon)$-approximation guarantee. This contrasts with the bounds of [75], which had an inverse polynomial dependence on ϵ. Furthermore, our explicit bounds are much tighter than those of [75], and we only require a finite first moment, i.e. our results also hold for heavy-tailed distributions. This takes a large step towards answering several open questions posed in [75] with regards to deriving bounds tight enough to be useful in practice. In particular, our bounds suggest that for small values of L, one can solve a large dynamic program to derive the optimal policy (whose size may be exponential in L), while for larger values of L one can simply use the constant-order policy. We again note that since our results only hold for the infinite-horizon problem, and will use critically certain stationarity properties that only hold in this regime, our results are not directly comparable to those of [75], whose bounds also hold for finite-horizon problems. Closing this gap, and proving tighter bounds for the finite-horizon problem, remains an interesting open question.

2.2.4.1 Example: exponentially distributed demand

In this section, for the special case of exponentially distributed demand, we further compute all expressions appearing in our bound in closed form, and numerically evaluate them, demonstrating good performance for a wide range of parameter values. Thus suppose demand is exponentially distributed with rate λ, i.e. mean λ^{-1}. In this case, it is well-known that for $r \in [0, \mathbb{E}[D])$, $\mathbb{E}[I_\infty^r]$ is the expected steady-state waiting

time in a corresponding $M/D/1$ queue, and equals $\frac{r^2\lambda}{2(1-r\lambda)}$ (cf. [81]). For $b, h > 0$, let

$$\tau_{b,h} \triangleq \sqrt{\frac{h}{2b+h}} \quad , \quad \gamma_{b,h} \triangleq (1 - \tau_{b,h}) \exp(\tau_{b,h}).$$

It may be easily demonstrated that $\gamma_{b,h} \in (0, 1)$. In that case, when demand is exponentially distributed, Theorem 2 is equivalent to the following bound, for which we provide a complete derivation in the appendix (Section 2.5).

Corollary 1 (Case of exponentially distributed demand) *Suppose D is exponentially distributed with rate λ. Then $C(\pi_{r_\infty}) = \lambda^{-1}(\sqrt{h(2b+h)} - h)$, and for all $L \geq 1$,*

$$\frac{C(\pi_{r_\infty})}{OPT(L)} \leq 1 + \left(\tau_{b,h} + (\tau_{b,h}^{-1} - 1)\big(e(L+1)\big)^{-1} \right)\left((1 - \gamma_{b,h}) \log(1 + bh^{-1}) \right)^{-1} \gamma_{b,h}^{L+1}. \quad (9)$$

Note that (9) does not depend on λ, which follows from the scaling properties of the exponential distribution. We now numerically evaluate (9) under different lost-demand penalty and lead time scenarios, with the holding cost fixed to 1, and present the results in Table 1. For each b, we also give the value of the best constant-order policy, $C(\pi_{r_\infty})$, further assuming $\mathbb{E}[D] = 1$ (i.e. $\lambda = 1$).

Table 1: When $h = \lambda = 1$, values of (9) and $C(\pi_{r_\infty})$ under different b and L

(9)	L=1	L=4	L=10	L=20	L=30	L=50	L=70	L=100	$C(\pi_{r_\infty})$
b=1/9	1.64	1.01	1.00	1.00	1.00	1.00	1.00	1.00	0.11
b=1/4	2.13	1.08	1.00	1.00	1.00	1.00	1.00	1.00	0.22
b=1	3.36	1.89	1.15	1.01	1.00	1.00	1.00	1.00	0.73
b=4	6.42	3.99	2.62	1.72	1.34	1.08	1.02	1.00	2.00
b=9	12.26	6.77	4.43	3.12	2.45	1.73	1.38	1.15	3.36
b=19	26.56	12.88	7.70	5.36	4.33	3.22	2.58	1.98	5.25
b=39	62.26	27.60	14.86	9.62	7.62	5.75	4.75	3.81	7.89
b=99	204.5	85.21	41.77	24.43	18.20	12.92	10.49	8.49	13.11

We note that when $\frac{b}{h}$ is small (i.e. less than or equal to 1), our bounds demonstrate an excellent performance by the constant-order policy even for lead times as small as 10. When $\frac{b}{h}$ is moderate (i.e. less than or equal to 9), our bounds demonstrate a similarly good performance for lead times on the order of 70. Even when $\frac{b}{h}$ is very large, our bounds still imply non-trivial performance guarantees, e.g. the constant-order policy is always within a factor of 4 of optimal when $b = 39$ and $L = 100$. Combining (9) with our explicit evaluation of $C(\pi_{r_\infty})$ yields tight bounds on $\mathrm{OPT}(L)$ whenever (9) is close to 1. For example, for $b = \lambda = 1$, our bounds imply that $\mathrm{OPT}(10) \in [.63, .73]$. Given the complexity of computing $\mathrm{OPT}(L)$ exactly using dynamic programming as L grows (cf. [214, 36]), and the very large values of L required for the earlier results of [75] to apply, we believe that our bounds provide the first window into the behavior of $\mathrm{OPT}(L)$ for moderate, but realistic, values of L, b, h.

2.2.4.2 Impact of the ratio $\frac{b}{h}$ on our bounds

In this section, we discuss the dependence of our demonstrated exponential rate of convergence γ on the ratio of the lost-sales penalty to the holding cost. Indeed, the numerical results of Section 2.2.4.1 suggest a degradation in performance as $\frac{b}{h}$ grows, and we now formalize this. In particular, we show that γ is non-decreasing in $\frac{b}{h}$. Note that by a simple scaling argument, for any fixed demand distribution $\mathcal{D}$, r_∞ (and thus γ) is a function of $\frac{b}{h}$ only, as opposed to the particular values of b, h. To make this dependence explicit, let $\gamma(\varrho)$ denote the value of γ when $\frac{b}{h} = \varrho$ (with the dependence on $\mathcal{D}$ implicit).

Lemma 1 *Under the same assumptions as Theorem 2, for any fixed demand distribution $\mathcal{D}$, $\gamma(\varrho)$ is non-decreasing in ϱ.*

We include a proof of Lemma 1 in the appendix (Section 2.5). This result suggests that the optimality gap of the constant-order policy may be larger when $\frac{b}{h}$ is large.

Interestingly, this is exactly the regime in which [93] proved that order-up-to policies are nearly optimal. More formally understanding this connection remains an interesting open question.

2.3   Proof of Theorem 2

In this section, we prove Theorem 2. Recall that π^* denotes some fixed stationary and convergent policy which is optimal for Problem 2, where the existence of such a policy follows from our assumptions. Let $(\mathcal{I}^*, \chi^*)$ denote a vector distributed as the stationary measure of the corresponding Markov chain, and $\{D_i, i \geq 1\}$ an i.i.d. sequence of demands, distributed as D, independent of $(\mathcal{I}^*, \chi^*)$. Let $\delta_{i,j}$ equal 1 if $i = j$, and 0 otherwise. It follows from stationarity, the inventory dynamics, and a straightforward induction that $\mathcal{I}^* \sim \max_{j=0,\dots,L} \left(\sum_{i=1}^{j} (\chi^*_{L+1-i} - D_{L+1-i}) + \delta_{j,L} \mathcal{I}^* \right)$. Thus

$$\mathbb{E}[\mathcal{I}^*] = \mathbb{E}\left[\max_{j=0,\dots,L} \left(\sum_{i=1}^{j} (\chi^*_{L+1-i} - D_{L+1-i}) + \delta_{j,L} \mathcal{I}^* \right) \right]. \tag{10}$$

The crux of our argument consists of two simple observations. First, we conclude the following from stationarity and the manner in which the pipeline vector is updated.

Observation 1 χ_i^* *has the same distribution for all* $i \in [1, L]$*, and* $\mathbb{E}[\chi_i^*] = \mathbb{E}[\chi_1^*]$ *for all* $i \in [1, L]$*.*

Second, we note that the right-hand side of (10) is a jointly convex function of χ^* and $\mathcal{I}^*$, which will allow us to apply the multi-variate Jensen's inequality (cf. [50]). In particular, for fixed $\mathbf{d} \in \mathbb{R}^L$, let us define

$$\mathfrak{f}_{\mathbf{d}}(\chi_1, \dots, \chi_L, I) \triangleq \max_{j=0,\dots,L} \left(\sum_{i=1}^{j} (\chi_{L+1-i} - d_{L+1-i}) + \delta_{j,L} I \right).$$

Observation 2 *For each fixed* $\mathbf{d} \in \mathbb{R}^L$*,* $\mathfrak{f}_{\mathbf{d}}(\chi_1, \dots, \chi_L, I)$ *is a jointly convex function of* $(\chi_1, \dots, \chi_L, I)$ *over* $\mathbb{R}^{L+1}$*. Combining with Observation 1, the multi-variate Jensen's inequality, and the i.i.d. property of* $\{D_i, i \geq 1\}$*, we conclude that*

$$\mathbb{E}[\mathcal{I}^*] \geq \mathbb{E}\left[\max_{j=0,\dots,L} \left(j\mathbb{E}[\chi_1^*] - \sum_{i=1}^{j} D_i + \delta_{j,L} \mathbb{E}[\mathcal{I}^*] \right) \right]. \tag{11}$$

We note that the observation of such a convexity in terms of the on-hand inventory and pipeline vector is not new. Indeed, the so called L-natural-convexity of the relevant cost-to-functions has been studied extensively (cf. [113, 138, 215, 36]), and used to obtain both structural results and algorithms. In contrast, here we use convexity to relate the expected inventory under an optimal policy to the expected inventory under a particular constant-order policy, intuitively that which orders $\mathbb{E}[\chi_1^*]$ in every period. Very similar ideas and arguments have appeared previously in the queueing theory literature, to demonstrate the extremality (with regards to expected waiting times) of certain queueing systems with constant service (or inter-arrival) times (cf. [96, 82]). We also note that although a related idea appears in the proofs of [75], the fact that they do not work in the stationary regime results in their obtaining much weaker results, since outside of stationarity one can no longer assume that all pipeline vector components have the same mean.

Before proceeding, let us define several additional notations. In particular, for $r \in \left[0, \mathbb{E}[D]\right]$ and $L \geq 1$, let

$$I_L^r \triangleq \max_{j=0,\ldots,L} \left(jr - \sum_{i=1}^{j} D_i \right) \quad , \quad C_L(r) \triangleq h\mathbb{E}[I_L^r] + b\mathbb{E}[D] - br,$$

and $r_L \in \arg\min_{0 \leq r \leq \mathbb{E}[D]} C_L(r)$ denote the infimum of the set of minimizers of $C_L(r)$. We note that I_L^r is distributed as the waiting time of the L-th customer in the corresponding $GI/GI/1$ queue (initially empty) with interarrival distribution D and processing time r. We also note that for $r \in [0, \mathbb{E}[D])$, I_∞^r is the weak limit, as $L \to \infty$, of I_L^r. Similarly, $C(\pi_r) = \lim_{L \to \infty} C_L(r)$, and $C_L(r)$ is monotone increasing in L.

We now combine (11) with (6), non-negativity, and definitions to bound the optimality gap of the constant-order policy.

Lemma 2 $OPT(L) \geq C_L(r_L)$, and

$$C(\pi_{r_\infty}) - OPT(L) \leq h \left(\mathbb{E}\left[I_\infty^{r_\infty}\right] - \mathbb{E}\left[I_L^{r_\infty}\right] \right) + h \left(\mathbb{E}\left[I_L^{r_\infty}\right] - \mathbb{E}\left[I_L^{r_L}\right] \right) - b \left(r_\infty - r_L \right). \quad (12)$$

Proof. Combining (11) with the nonnegativity of $\mathbb{E}[\mathcal{I}^*]$, we conclude that $\mathbb{E}[\mathcal{I}^*] \geq \mathbb{E}\left[I_L^{\mathbb{E}[\chi_1^*]}\right]$. Thus by (6), $\mathrm{OPT}(L) \geq h\mathbb{E}\left[I_L^{\mathbb{E}[\chi_1^*]}\right] + b\mathbb{E}[D] - b\mathbb{E}[\chi_1^*]$. Combining with (5) and the definition of r_L, we conclude that $\mathrm{OPT}(L) \geq C_L(r_L)$. It then follows from (7) that

$$
\begin{aligned}
C\left(\pi_{r_\infty}\right) - \mathrm{OPT}(L) &\leq \ (h\mathbb{E}\left[I_\infty^{r_\infty}\right] + b\mathbb{E}[D] - br_\infty) - (h\mathbb{E}\left[I_L^{r_L}\right] + b\mathbb{E}[D] - br_L) \\
&= h\left(\mathbb{E}\left[I_\infty^{r_\infty}\right] - \mathbb{E}\left[I_L^{r_\infty}\right]\right) + h\left(\mathbb{E}\left[I_L^{r_\infty}\right] - \mathbb{E}\left[I_L^{r_L}\right]\right) - b\left(r_\infty - r_L\right),
\end{aligned}
$$

completing the proof. ∎

We proceed by bounding the terms appearing in the right-hand side of (12) separately. We begin by recalling a classical result of [116], which uses the celebrated Spitzer's identity to bound the difference between the expected waiting time of the Lth job to arrive to a single-server queue (initially empty), and the steady-state expected waiting time. As this difference is exactly $\mathbb{E}\left[I_\infty^r\right] - \mathbb{E}\left[I_L^r\right]$, the result will allow us to bound the relevant term of (12). We state Kingman's results as customized to our own setting, notations, and assumptions.

Lemma 3 (Theorems 1, 4, [116]) *For all $r \in \left[0, \mathbb{E}[D]\right]$ and $L \geq 1$,*

$$
\mathbb{E}[I_L^r] = \sum_{n=1}^{L} \frac{1}{n} \mathbb{E}\left[\left(nr - \sum_{i=1}^{n} D_i\right)^+\right].
$$

If in addition $r < \mathbb{E}[D]$, then

$$
\mathbb{E}[I_\infty^r] = \sum_{n=1}^{\infty} \frac{1}{n} \mathbb{E}\left[\left(nr - \sum_{i=1}^{n} D_i\right)^+\right].
$$

Also,

$$
\mathbb{E}\left[I_\infty^{r_\infty}\right] - \mathbb{E}\left[I_L^{r_\infty}\right] \leq \ \left((1-\gamma)e\vartheta(L+1)\right)^{-1}\gamma^{L+1}.
$$

To bound the remaining term $h\left(\mathbb{E}\left[I_L^{r_\infty}\right] - \mathbb{E}\left[I_L^{r_L}\right]\right) - b\left(r_\infty - r_L\right)$, we begin by proving that $r_\infty \leq r_L$ for all L. This makes sense at an intuitive level, since r_L is minimizing a function which "penalizes less" for carrying inventory. However, in spite of this clear intuition, the proof is not entirely trivial, as one function dominating another does not necessarily imply a similar comparison of the appropriate argmins.

Lemma 4 $r_\infty \leq r_L$ for all $L \geq 1$.

Proof. Suppose for contradiction that there exists $L \in [1, \infty)$ such that $r_L < r_\infty$. Note that in this case, both $r_L, r_\infty < \mathbb{E}[D]$, and thus by Lemma 3 both $\mathbb{E}[I_\infty^{r_\infty}], \mathbb{E}[I_\infty^{r_L}] < \infty$. From definitions and the associated respective optimality of r_L, r_∞, we conclude that

$$h\mathbb{E}[I_\infty^{r_\infty}] + b\mathbb{E}[D] - br_\infty \leq h\mathbb{E}[I_\infty^{r_L}] + b\mathbb{E}[D] - br_L,$$

$$h\mathbb{E}[I_L^{r_L}] + b\mathbb{E}[D] - br_L \leq h\mathbb{E}[I_L^{r_\infty}] + b\mathbb{E}[D] - br_\infty.$$

Summing these two inequalities together implies that

$$\mathbb{E}[I_\infty^{r_\infty}] + \mathbb{E}[I_L^{r_L}] \leq \mathbb{E}[I_\infty^{r_L}] + \mathbb{E}[I_L^{r_\infty}],$$

which, by Lemma 3, is equivalent to

$$\sum_{n=L+1}^{\infty} \frac{1}{n}\mathbb{E}\left[\left(nr_\infty - \sum_{i=1}^{n} D_i\right)^+\right] \leq \sum_{n=L+1}^{\infty} \frac{1}{n}\mathbb{E}\left[\left(nr_L - \sum_{i=1}^{n} D_i\right)^+\right] < \infty. \quad (13)$$

Here we note that the desired result intuitively follows from (13) and the monotonicity of the relevant functions, i.e. the fact that $x > y$ implies $\mathbb{E}\left[(nx - \sum_{i=1}^{n} D_i)^+\right] \geq \mathbb{E}\left[(ny - \sum_{i=1}^{n} D_i)^+\right]$. However, we must rule out certain subtle problems that could potentially arise from the function $\mathbb{E}\left[(nr - \sum_{i=1}^{n} D_i)^+\right]$ not being *strictly* monotonic in r, and proceed as follows. Definitions, non-negativity, and the fact that $r_L < r_\infty$, together imply that

$$\left(nr_\infty - \sum_{i=1}^{n} D_i\right)^+ = \mathbb{I}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right)\left(nr_\infty - \sum_{i=1}^{n} D_i\right)$$

$$\geq \mathbb{I}\left(nr_L \geq \sum_{i=1}^{n} D_i\right)\left(nr_\infty - \sum_{i=1}^{n} D_i\right).$$

Combining with (13), we conclude that

$$\sum_{n=L+1}^{\infty} \frac{1}{n}\mathbb{E}\left[\mathbb{I}\left(nr_L \geq \sum_{i=1}^{n} D_i\right)\left(nr_\infty - \sum_{i=1}^{n} D_i\right)\right]$$

$$\leq \sum_{n=L+1}^{\infty} \frac{1}{n}\mathbb{E}\left[\mathbb{I}\left(nr_L \geq \sum_{i=1}^{n} D_i\right)\left(nr_L - \sum_{i=1}^{n} D_i\right)\right] < \infty.$$

34

It follows that

$$\sum_{n=L+1}^{\infty} \frac{1}{n} \mathbb{E}\left[\mathbb{I}\left(nr_L \geq \sum_{i=1}^{n} D_i\right)(nr_\infty - nr_L)\right] \leq 0.$$

However, since by assumption $nr_\infty - nr_L > 0$, it follows from non-negativity that

$$\sum_{n=L+1}^{\infty} \mathbb{E}\left[\mathbb{I}\left(nr_L \geq \sum_{i=1}^{n} D_i\right)\right] = 0,$$

and thus $\mathbb{P}\left(nr_L \geq \sum_{i=1}^{n} D_i\right) = 0$ for all $n \geq L+1$. Noting that $\mathbb{P}\left(nr_L \geq \sum_{i=1}^{n} D_i\right) \geq \mathbb{P}^n\left(r_L \geq D_1\right)$, we conclude that $\mathbb{P}\left(r_L \geq D_1\right) = 0$. It follows that $\mathbb{E}[I_\infty^{r_L}] = \mathbb{E}[I_L^{r_L}] = 0$, and thus by (7), $C\left(\pi_{r_L}\right) = C_L\left(r_L\right)$. Combining with Lemma 2, which implies that $C_L\left(r_L\right) \leq C\left(\pi_{r_\infty}\right)$, and the optimality of r_∞, we conclude $r_L \in \arg\min_{0 \leq r \leq \mathbb{E}[D]} C\left(\pi_r\right)$. However, as r_∞ is by definition the infimum of $\arg\min_{0 \leq r \leq \mathbb{E}[D]} C\left(\pi_r\right)$, the fact that $r_L < r_\infty$ thus yields a contradiction, completing the proof. ∎

Before proceeding, we also derive a certain critical inequality, which we will use to show that the term $h\left(\mathbb{E}\left[I_L^{r_\infty}\right] - \mathbb{E}\left[I_L^{r_L}\right]\right)$ and the term $b\left(r_\infty - r_L\right)$ essentially "cancel out". This inequality follows from the first-order optimality conditions of the convex optimization problem associated with r_∞, but requires some care, as the relevant functions are potentially non-differentiable, and the desired statement in principle involves an interchange of expectation and differentiation.

Lemma 5 $\sum_{n=1}^{\infty} \mathbb{P}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right) \geq \frac{b}{h}$.

Proof. Since $r_\infty < \mathbb{E}[D]$, there exists $\delta > 0$ such that $r_\infty + \epsilon < \mathbb{E}[D]$ for all $\epsilon \in [0, \delta]$. Let us fix any such $\epsilon > 0$. The definition and associated optimality of r_∞ implies that $\mathbb{E}[\pi_{r_\infty}] \leq \mathbb{E}[\pi_{r_\infty+\epsilon}]$. Combining with Lemma 3 and (7), we conclude that

$$h\sum_{n=1}^{\infty} \frac{1}{n} \mathbb{E}\left[\mathbb{I}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right)\left(nr_\infty - \sum_{i=1}^{n} D_i\right)\right]$$

is at most

$$h\sum_{n=1}^{\infty} \frac{1}{n} \mathbb{E}\left[\mathbb{I}\left(n(r_\infty + \epsilon) \geq \sum_{i=1}^{n} D_i\right)\left(n(r_\infty + \epsilon) - \sum_{i=1}^{n} D_i\right)\right] - b\epsilon.$$

Combining with the fact that

$$\mathbb{I}\left(nr_\infty \geq \sum_{i=1}^n D_i\right)\left(nr_\infty - \sum_{i=1}^n D_i\right) \geq \mathbb{I}\left(n(r_\infty + \epsilon) \geq \sum_{i=1}^n D_i\right)\left(nr_\infty - \sum_{i=1}^n D_i\right),$$

it follows that

$$h\sum_{n=1}^\infty \frac{1}{n}\mathbb{E}\left[\mathbb{I}\left(n(r_\infty + \epsilon) \geq \sum_{i=1}^n D_i\right)\left(nr_\infty - \sum_{i=1}^n D_i\right)\right]$$

is at most

$$h\sum_{n=1}^\infty \frac{1}{n}\mathbb{E}\left[\mathbb{I}\left(n(r_\infty + \epsilon) \geq \sum_{i=1}^n D_i\right)\left(n(r_\infty + \epsilon) - \sum_{i=1}^n D_i\right)\right] - b\epsilon.$$

Equivalently (as all relevant sums are finite)

$$\sum_{n=1}^\infty \mathbb{P}\left(n(r_\infty + \epsilon) \geq \sum_{i=1}^n D_i\right) \geq \frac{b}{h}.$$

As this holds for all sufficiently small ϵ, the only remaining step is to demonstrate validity at $\epsilon = 0$. By monotonicity, for each fixed n and all $\epsilon \in [0, \delta]$,

$$\mathbb{P}\left(n(r_\infty + \epsilon) \geq \sum_{i=1}^n D_i\right) \leq \mathbb{P}\left(n(r_\infty + \delta) \geq \sum_{i=1}^n D_i\right).$$

Furthermore, since $r_\infty + \delta < \mathbb{E}[D]$, for any fixed $\nu > 0$, there exists $M_\nu < \infty$ (depending only on $\nu, \mathcal{D}, r_\infty, \delta$) such that

$$\sum_{n=M_\nu+1}^\infty \mathbb{P}\left(n(r_\infty + \delta) \geq \sum_{i=1}^n D_i\right) \leq \nu.$$

Indeed, the above follows from a standard argument (the details of which we omit) in which each term is bounded using Chernoff's inequality, and the terms are summed as an infinite series (cf. [48]). Combining the above, we conclude that for all $\nu > 0$, and $\epsilon \in (0, \delta]$,

$$\sum_{n=1}^{M_\nu} \mathbb{P}\left(n(r_\infty + \epsilon) \geq \sum_{i=1}^n D_i\right) \geq \frac{b}{h} - \nu.$$

As $\mathbb{P}\left(nx \geq \sum_{i=1}^n D_i\right)$ is a right-continuous function of x (by the right-continuity of cumulative distribution functions), it follows that $\sum_{n=1}^{M_\nu} \mathbb{P}\left(nx \geq \sum_{i=1}^n D_i\right)$ is similarly right-continuous in x. Right-continuity at $\epsilon = 0$ follows, and we conclude that

$$\sum_{n=1}^{M_\nu} \mathbb{P}\left(nr_\infty \geq \sum_{i=1}^n D_i\right) \geq \frac{b}{h} - \nu.$$

As this holds for all ν, letting $\nu \downarrow 0$ completes the proof. ∎

With Lemmas 3, 4, and 5 in hand, we now complete the proof of our main result, Theorem 2.

Proof. [Proof of Theorem 2] Recall that the remaining term on the right-hand side of (12) which we are yet to bound is

$$h\left(\mathbb{E}\left[I_L^{r_\infty}\right] - \mathbb{E}\left[I_L^{r_L}\right]\right) - b\left(r_\infty - r_L\right). \tag{14}$$

We first bound $\mathbb{E}\left[I_L^{r_\infty}\right] - \mathbb{E}\left[I_L^{r_L}\right]$, which by Lemma 3 equals

$$\sum_{n=1}^{L} \frac{1}{n}\mathbb{E}\left[\left(nr_\infty - \sum_{i=1}^{n} D_i\right)\mathbb{I}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right)\right]$$
$$-\sum_{n=1}^{L} \frac{1}{n}\mathbb{E}\left[\left(nr_L - \sum_{i=1}^{n} D_i\right)\mathbb{I}\left(nr_L \geq \sum_{i=1}^{n} D_i\right)\right].$$

Combining with Lemma 4 (i.e. the fact that $r_L \geq r_\infty$), which implies that

$$\left(nr_L - \sum_{i=1}^{n} D_i\right)\mathbb{I}\left(nr_L \geq \sum_{i=1}^{n} D_i\right) \geq \left(nr_L - \sum_{i=1}^{n} D_i\right)\mathbb{I}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right),$$

we conclude that $\mathbb{E}\left[I_L^{r_\infty}\right] - \mathbb{E}\left[I_L^{r_L}\right]$ is at most

$$\sum_{n=1}^{L} \frac{1}{n}\mathbb{E}\left[\left(nr_\infty - \sum_{i=1}^{n} D_i\right)\mathbb{I}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right)\right]$$
$$-\sum_{n=1}^{L} \frac{1}{n}\mathbb{E}\left[\left(nr_L - \sum_{i=1}^{n} D_i\right)\mathbb{I}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right)\right],$$

which itself equals

$$-\left(r_L - r_\infty\right)\sum_{n=1}^{L}\mathbb{P}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right).$$

It follows that (14) is at most

$$\left(r_L - r_\infty\right)\left(b - h\sum_{n=1}^{L}\mathbb{P}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right)\right). \tag{15}$$

Note that Lemma 5 implies that

$$\sum_{n=L+1}^{\infty}\mathbb{P}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right) \geq \frac{b}{h} - \sum_{n=1}^{L}\mathbb{P}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right). \tag{16}$$

37

Combining (15) and (16), we conclude that (14) is at most

$$(r_L - r_\infty)\, h \sum_{n=L+1}^{\infty} \mathbb{P}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right).$$

It follows from the well-known Chernoff's inequality (cf. [48]) that

$$\mathbb{P}\left(nr_\infty \geq \sum_{i=1}^{n} D_i\right) \leq \gamma^n.$$

By summing the associated geometric series, and combining with the fact that by definition $r_L \leq \mathbb{E}[D]$, we conclude that (14) is at most

$$h(\mathbb{E}[D] - r_\infty)(1 - \gamma)^{-1}\gamma^{L+1}.$$

Combining the above bound for (14) with Lemma 3, plugging into (12), and applying the fact that $\text{OPT}(L) \geq g$, completes the proof. ∎

2.4   Conclusion

In this chapter, we proved that for a family of challenging inventory models (i.e. lost sales models with large lead times), the optimality gap of the simple constant-order policy converges *exponentially fast* to zero as the lead time grows with the other problem parameters held fixed, and derived effective explicit bounds for this optimality gap. This takes a large step towards answering several open questions of [75], who recently proved the asymptotic optimality of the constant-order policy in this setting, but whose bounds on the rate of convergence were impractical, involving a massive prefactor and an inverse polynomial dependence on the relevant error term. We also demonstrated that the corresponding rate of exponential decay is at least as fast as the exponential rate of convergence of the expected waiting time in a certain single-server queue to its steady-state value, which we proved to be monotone in the ratio of the lost-sales penalty to the holding cost. For the special case of exponentially distributed demand, we further computed all expressions appearing in

our bound in closed form, and numerically evaluated these bounds, demonstrating good performance for a wide range of parameter values.

This work leaves many interesting directions for future research. First, it would be interesting to investigate the tightness of our exponential bound, e.g. to determine whether our exponential rate captures the true exponential rate of convergence of the optimality gap of the constant-order policy. Although one can come up with pathological examples for which this is not true, e.g. discrete demand distributions with probability at least $\frac{b}{b+h}$ at 0 (for which $Q = 0$, π_0 is optimal amongst all policies for all $L \geq 0$, yet $\gamma > 0$), we conjecture that under mild assumptions γ indeed captures the true rate of convergence of the optimality gap. On a related note, it is an open question to bridge the gap between our own results, which hold only in the stationary regime, and the results of [75], which also hold for the corresponding finite-horizon problem. Although (as noted by the authors) the arguments of [75] could likely be rederived using convexity-type (as opposed to coupling) arguments, resolving the fundamental question of whether (and precisely in what sense) the optimality gap decays exponentially quickly for finite-horizon problems seems to entail overcoming several challenging problems related to quantifying the "rate of convergence to stationarity" of optimal policies for lost sales inventory models, i.e. both the rate at which the finite-horizon problem "converges" to the infinite-horizon problem, and the rate at which the Markov chains associated with stationary optimal policies for the infinite-horizon problem converge to their steady-state behavior. Although recent progress has been made on several related questions (cf. [95]), a complete resolution seems beyond the reach of current techniques.

Second, it is an interesting open challenge to analyze the performance of more sophisticated policies for lost sales inventory models, e.g. affine policies, which should exhibit even better performance. For example, it is an open question whether the optimality gap of a more sophisticated (but still simple and efficient) policy can be

made to decay (exponentially) faster as the lead time grows. Another interesting question involves the formal construction and analysis of "hybrid" algorithms, which e.g. solve large dynamic programs when L is small and transition to using simpler policies when L is large, or use base-stock policies when $\frac{b}{h}$ is large (relative to L) and a constant-order policy when $\frac{b}{h}$ is small. More generally, it is an open challenge to classify the family of all algorithms which are asymptotically optimal as the lead time grows, and to better understand the relative performance of these policies. For example, it is an open question whether the order-up-to policy of [93], or the dual-balancing policy of [123], exhibit such asymptotic optimality.

Third, it would be interesting to prove that a similar phenomena occurs in more general inventory settings. As a first step, one could extend our results to models in which demand is not i.i.d., models with a fixed ordering cost, and models with integrality constraints. In these settings, although one would expect that the constant-order policy may no longer be asymptotically optimal, one can ask what analogous simple policy should have the asymptotic optimality property. We also believe that our methodology can be able to prove that simple policies work well for considerably more complex inventory models. For example, in the next chapter, we demonstrate that the so-called tailored base-surge policy, which combines the constant-order policy with a base-stock policy, is asymptotically optimal for the more sophisticated dual-sourcing inventory model, a natural generalization of the lost sales model considered here (cf. [105]).

On a final note, our results and methodology (combined with that of [75]) provide a fundamentally new approach to lost sales inventory models with positive lead times. We believe that our approach, combined with other recent developments in inventory theory (e.g. the efficient solution of related dynamic programs), represents a considerable step towards making these models solvable in practice. Such progress

may ultimately help to free researchers from having to use backlogged demand inventory models as approximations to lost-sales inventory models, even when such an approximation is not appropriate, which has been recognized as a major problem in the inventory theory literature (cf. [24]).

2.5 Appendix

Proof. [Proof of Corollary 1] Suppose D is exponentially distributed with rate λ. It is well-known that in this case, for all $r \in [0, \lambda^{-1})$, $\mathbb{E}[I_\infty^r] = \frac{r^2\lambda}{2(1-r\lambda)}$ (cf. [81]). It follows from (7) that for all $r \in [0, \lambda^{-1})$,

$$C(\pi_r) \quad = \quad h\mathbb{E}[I_\infty^r] + b\mathbb{E}[D] - br \quad = \quad h\frac{r^2\lambda}{2(1-r\lambda)} + b\lambda^{-1} - br.$$

Recall that $C(\pi_r)$ is a convex function of r on $[0, \lambda^{-1})$, and note that

$$\frac{d}{dr}C(\pi_r) = \frac{h}{2}\left((\lambda r - 1)^{-2} - 1\right) - b. \tag{17}$$

As it is easily verified that the right-hand side of (17) strictly increases from $-b$ to ∞ on $[0, \lambda^{-1})$, it follows that r_∞ must be the unique solution to the equation $\frac{d}{dr}C(\pi_r) = 0$ on $[0, \lambda^{-1})$. It then follows from a straightforward calculation (the details of which we omit) that

$$r_\infty = \lambda^{-1}(1 - \tau_{b,h}) \quad , \quad C(\pi_{r_\infty}) = \lambda^{-1}(\sqrt{h(2b+h)} - h).$$

As $\mathbb{E}[\exp(-\theta D)] = \lambda(\lambda + \theta)^{-1}$ for all $\theta \geq 0$, we conclude that

$$\phi(\theta) = \exp\left(\lambda^{-1}(1 - \tau_{b,h})\theta\right)\lambda(\lambda + \theta)^{-1}.$$

As above, it follows from another straightforward calculation (the details of which we omit) that ϑ equals the unique solution to $\frac{d}{d\theta}\phi(\theta) = 0$ on $[0, \infty)$, and thus

$$\vartheta = \tau_{b,h}\lambda(1 - \tau_{b,h})^{-1} \quad , \quad \gamma = \gamma_{b,h} = (1 - \tau_{b,h})\exp(\tau_{b,h}).$$

Finally, let us compute Q and g. Noting that Q is the $\frac{b}{b+h}$ quantile of the demand distribution, i.e., $1 - \exp(-\lambda Q) = \frac{b}{b+h}$, implies $Q = \lambda^{-1} \log(1 + bh^{-1})$. It follows that

$$
\begin{aligned}
g &= h \int_0^Q (Q - x)\lambda \exp(-\lambda x)\, dx + b \int_Q^\infty (x - Q)\lambda \exp(-\lambda x)\, dx \\
&= (h + b) \int_0^Q (Q - x)\lambda \exp(-\lambda x)\, dx + b\left(\mathbb{E}[D] - Q\right) \\
&= (h + b)\left(Q - \lambda^{-1} + \lambda^{-1} \exp(-\lambda Q)\right) + b\left(\lambda^{-1} - Q\right) \\
&= h\lambda^{-1} \log\left(1 + bh^{-1}\right).
\end{aligned}
$$

Combining the above with a straightforward calculation (the details of which we omit) completes the proof. $\blacksquare$

Proof. [Proof of Lemma 1] Suppose $\frac{b_1}{h_1} < \frac{b_2}{h_2}$, and let

$$
r_\infty^i \in \arg\min_{0 \le r \le \mathbb{E}[D]} \left(h_i \mathbb{E}[I_\infty^r] + b_i \mathbb{E}[D] - b_i r\right), \quad i = 1, 2.
$$

From the respective optimality of r_∞^1, r_∞^2, we conclude that

$$
\begin{aligned}
\mathbb{E}[I_\infty^{r_\infty^1}] + \tfrac{b_1}{h_1}\mathbb{E}[D] - \tfrac{b_1}{h_1}r_\infty^1 &\le \mathbb{E}[I_\infty^{r_\infty^2}] + \tfrac{b_1}{h_1}\mathbb{E}[D] - \tfrac{b_1}{h_1}r_\infty^2, \\
\mathbb{E}[I_\infty^{r_\infty^2}] + \tfrac{b_2}{h_2}\mathbb{E}[D] - \tfrac{b_2}{h_2}r_\infty^2 &\le \mathbb{E}[I_\infty^{r_\infty^1}] + \tfrac{b_2}{h_2}\mathbb{E}[D] - \tfrac{b_2}{h_2}r_\infty^1.
\end{aligned}
$$

Summing these two inequalities together implies

$$
\left(\frac{b_2}{h_2} - \frac{b_1}{h_1}\right)\left(r_\infty^2 - r_\infty^1\right) \ge 0.
$$

It follows that $r_\infty^2 \ge r_\infty^1$, and for all $\theta \ge 0$,

$$
\exp(\theta r_\infty^2)\mathbb{E}[\exp(-\theta D)] \ge \exp(\theta r_\infty^1)\mathbb{E}[\exp(-\theta D)].
$$

Combining with the definition of γ completes the proof. $\blacksquare$

CHAPTER III

ASYMPTOTIC OPTIMALITY OF TAILORED BASE-SURGE POLICIES IN DUAL-SOURCING INVENTORY SYSTEMS

This chapter is based on [203].

3.1 Introduction and literature review

Companies face the challenge of optimizing their sourcing strategies in a globalized world, and how to best utilize different sources effectively is a billion dollar industry. In practice, many companies (such as Caterpillar, cf. [157]), often adopt *dual-sourcing* strategies for making such decisions. Under a dual-sourcing strategy, the companies usually purchase their materials from a regular supplier at a lower cost, but they are also able to obtain materials from an expedited supplier at a higher cost under emergency circumstances. For example, in the summer of 2003, Amazon used FedEx to deliver the new Harry Potter more promptly and maintained regular shipping via UPS (cf. [115], [191]). [3] describes an example of a $10 billion high-tech U.S. company that has two suppliers, one in Mexico and one in China. The one in Mexico has shorter lead time but higher per-unit ordering cost; the one in China has longer lead time (5 to 10 times longer) but lower per-unit ordering cost. The company takes advantage of the dual-sourcing strategy to meet the demand more responsively (from Mexico) as well as less expensively (from China).

Although dual-sourcing is attractive, and very relevant to practice, optimizing a dual-sourcing inventory system is notoriously challenging. Such inventory systems have been studied now for over forty years, but the structure of the optimal policy

remains poorly understood, with the exception of when the system is consecutive, i.e., the lead time difference between the two sources is exactly one. More specifically, the earliest studies of periodic review dual-sourcing inventory models include [9], [45], and [143], which showed that base-stock (also known as order-up-to) policies are optimal when the lead times of the two sources are zero and one respectively. [67] extended the result to general lead time settings as long as the lead time difference remains one. [199] showed that the optimal policy is no longer a simple base-stock policy when the lead time difference is beyond one and the structure of the optimal policy can be quite complex. Furthermore, it is well known that a dual-sourcing inventory system can be regarded as a generalization of a lost-sales inventory system (cf. [183]). Indeed, the intractability of both the dual-sourcing and lost-sales inventory models has a common source - as the lead time grows, the state-space of the natural dynamic programming (DP) formulation grows exponentially, rendering such techniques impractical. This issue is typically referred to as the "curse of dimensionality" (cf. [113], [138], [214]), and we refer the reader to [75] and [202] for a relevant discussion in the context of lost-sales inventory models.

There is a vast literature investigating periodic review dual-sourcing inventory models as well as their variants (cf. [134]), including: models with multiple suppliers (cf. [209], [134], [61]); models with two suppliers, one with higher variable costs but lower setup costs, and one with lower variable costs but higher setup costs (cf. [66]); models with a long-term contract supplier and a spot market (cf. [207], [40]); models for which the unmet demand must be satisfied from the expediting source (cf. [92]); models with expediting and advance demand information (cf. [4]); models allowing emergency orders within the regular review period (cf. [188]); models with correlated demand (cf. [176]); models considering capacity cost and flexibility for sourcing decisions (cf. [29]); multi-echelon models with expediting (cf. [121]); and models with joint inventory-pricing control (cf. [76], [210]). For continuous review

44

models, we refer the reader to [135], [136], [30], [185], and the references therein.

As an exact solution seems out of reach, the operations research and management communities have instead investigated certain structural properties of the optimal policy (cf. [91]), and exerted considerable effort towards constructing various heuristic policies. [191] proposed the family of dual index (DI) policies, which have two base-stock levels, one for the regular source and one for the express source, and "orders up" to bring appropriate notions of inventory position up to these levels. [173] analyzed the closely related class of single index (SI) policies, for which the relevant notions of inventory position are different. Both families of policies seem to perform well in numerical studies. [183] considered two generalized classes of policies: one with an order-up-to structure for the express source, and one with an order-up-to structure for the regular source. Their numerical experiment showed that such policies can outperform DI policies. In the presence of production capacity costs, [29] studied dual-sourcing smoothing policies, under which the order quantities from both sources in each period are convex combinations of observed past demands. They analyzed such polices under normally distributed demand, and their numerical results showed that these policies performed better for higher capacity costs and longer lead time differences (between the two sources).

A simple and natural policy that is implemented in practice, which will be the subject of our own investigations, is the so-called *Tailored Base-Surge* (TBS) policy. It was first proposed and analyzed in [3], where we note that closely related standing order policies had been studied previously (cf. [162, 106]). Under such a TBS policy, a constant order is placed at the regular source in each period to meet a *base* level of demand, while the orders placed at the express source follow an order-up-to rule to manage demand *surges*. We refer to Mini-Case 6 in [131] for more about the motivation and background of TBS policies. Note that dual-sourcing inventory systems

45

in which a constant-order policy is implemented for the regular source are essentially equivalent to single-sourcing inventory systems with constant returns, which have been investigated in the literature (cf. [64], [46]).

[3] analyzed TBS policies in a continuous review model, and their focus was to find the best TBS policy. Numerical results in [118], [163] showed that TBS policies are comparable to DI policies, and outperform DI policies for some problem instances. [3] conjectured that this policy performs more effectively as the lead time difference between the two sources grows. [105] analyzed a periodic review model and studied the performance of TBS policy. They provided an explicit bound on the performance of TBS policies compared to the optimal one when the demand had a specific structure, and provided numerical experiments suggesting that the performance of the TBS policy improves as the lead time difference grows large.

However, to date there is no theoretical justification for the good behavior of TBS policies as the lead time difference grows large, and giving a solid theoretical foundation to this observed phenomena remains a major open question. We note that until recently, a similar state of affairs existed regarding the good performance of constant-order policies as the lead time grows large in single-source lost-sales inventory models. However, using tools from applied probability, queueing theory, and convexity, this phenomena was recently explained in [75] and [202], in which it was proven that a simple constant-order policy is asymptotically optimal in this setting as the lead time of the single source grows large. The intuition here is that as the lead time grows large, so much randomness is introduced into the system between when an order is placed and when that order is received, that it is essentially impossible for any algorithm to meaningfully use the state information to make significantly better decisions. Thus a policy which ignores the state information (i.e. constant-order policy) performs nearly as well as an optimal policy. We note that the results of [202] (i.e., chapter 2 in this thesis) further demonstrate that the optimality gap of the constant-order

policy actually shrinks exponentially fast to zero as the lead time grows large, and provide explicit and effective bounds even for moderate-to-small lead times.

3.1.1 Outline of chapter

The rest of the chapter is organized as follows. We formally define the dual-sourcing problem in Section 3.2.1, and describe the TBS policy in Section 3.2.2. We state our main result in Section 3.2.3, and prove our main result in Section 3.3. We summarize our main contributions and propose directions for future research in Section 3.4. We also include a technical appendix in Section 3.5.

3.2 *Main results*

3.2.1 Model description, problem statement and assumptions

In this section, we formally define our dual-sourcing inventory problem, closely following the definitions given in [183]. Note that the general framework of dual-sourcing inventory model is already introduced in Section 1.2.2 and we recall the notations here. Let $\{D_t\}_{t \geq 1}, \{D'_t\}_{t \geq 1}$ be mutually independent sequence of nonnegative independent and identically distributed (i.i.d.) demand realizations, distributed as the random variable (r.v.) D. Let T be the time horizon, $L \geq 1$ be the deterministic lead time of the regular source (R), and $L_0 \geq 0$ the deterministic lead time of the express source (E), where $L \geq L_0 + 1$. Let c_R, c_E be the unit purchase costs of the regular and express sources, and h, b be the unit holding and backorder costs respectively, with $c \triangleq c_E - c_R > 0$. In addition, let I_t denote the on-hand inventory at the start of period t (before any orders or demands are received), and $q_t^R(q_t^E)$ denote the order placed from R(E) at the beginning of period t. Note that due to the leadtimes, the order received from R(E) in period t is $q_{t-L}^R(q_{t-L_0}^E)$. As we will be primarily interested in the corresponding long-run-average problem, we without loss of generality (w.l.o.g.) suppose that the initial conditions are such that the initial inventory is 0, and no initial orders have been placed from either R or E. As a notational convenience, we

47

define $q_k^R = 0, k = -(L-1), \ldots, 0$; and $q_k^E = 0, k = -(L_0-1), \ldots, 0$. For $t = 1, \ldots, T$, recall that the events in period t are ordered as follows.

- Ordering decisions from R and E are made (i.e. q_t^E, q_t^R are chosen);

- New inventory $q_{t-L}^R + q_{t-L_0}^E$ is delivered and added to the on-hand inventory;

- The demand D_t is realized, costs for period t are incurred, and the inventory is updated.

Note that the on-hand inventory is updated according to $I_{t+1} = I_t + q_{t-L}^R + q_{t-L_0}^E - D_t$, and may be negative since backorder is allowed. We require that the new orders q_t^R and q_t^E are non-negative measurable (and thus deterministic) functions of the realized demands, inventory levels, and ordering quantities in periods $1, \ldots, t-1$, as well as the problem primitives $D, L, L_0, c_R, c_E, h, b$ and the current time t. We call the corresponding family of policies *admissible*, and denote this family by Π. We note that any policy $\pi \in \Pi$ can in principle be implemented on such a problem of any time horizon (even infinite). Let $G(y)$ be the sum of the holding and backorder costs when the inventory level equals y in the end of a time period, i.e. $G(y) \triangleq hy^+ + by^-$, where $x^+ \triangleq \max(x,0)$, $x^- \triangleq \max(-x,0)$. Here we note that G is convex and Lipschitz, and for $x, y \in \mathbb{R}$,

$$|G(x) - G(y)| \leq \max(b,h)|x-y|, \quad \text{and} \quad |G(x)| \geq \min(b,h)|x|. \qquad (18)$$

Let C_t be the sum of the ordering, holding and backorder costs incurred in time period t, i.e. $C_t \triangleq c_R q_t^R + c_E q_t^E + G(I_t + q_{t-L}^R + q_{t-L_0}^E - D_t)$. To denote the dependence of the cost on the policy π, we use the notation C_t^π. Let $C(\pi)$ denote the long-run average cost incurred by a policy π, i.e. $C(\pi) \triangleq \limsup_{T \to \infty} \frac{\sum_{t=1}^T \mathbb{E}[C_t^\pi]}{T}$. The value of the corresponding long-run average cost dual-sourcing inventory optimization problem is denoted by $\mathrm{OPT}(L) \triangleq \inf_{\pi \in \Pi} C(\pi)$.

Before proceeding, it will be useful to review a certain well-known reduction between the setting in which $L_0 > 0$ and the setting in which $L_0 = 0$ (cf. [183]), where we note that similar reductions are known to hold for many classical inventory problems with backlogging (cf. [113, 170]). Let us define the so-called *expedited inventory position* at time $t \geq 1$ as $\hat{I}_t \triangleq I_t + \sum_{k=t-L_0}^{t-1} q_k^E + \sum_{k=t-L}^{t-L+L_0} q_k^R$, which corresponds to the net inventory at the start of period t plus all orders to be received in periods $t, \ldots, t + L_0$ (which were placed before period t), and the *truncated regular pipeline* at time t as the $(L - L_0 - 1)$-dimensional vector $\mathcal{R}^t \triangleq (q_{t-L+L_0+1}^R, \ldots, q_{t-1}^R)$, with $\mathcal{R}_k^t = q_{t-L+L_0+k}^R, k = 1, \ldots, L - L_0 - 1$. Let $\hat{\Pi}$ denote those policies belonging to π with the additional restriction that the new orders q_t^R, q_t^E are measurable functions of only $\hat{I}_t, \mathcal{R}^t$, as well as the problem primitives $D, L, L_0, c_R, c_E, h, b$ and current time t.

Lemma 6 ([183] Lemma 2.1) $\inf_{\pi \in \Pi} C(\pi) = \inf_{\pi \in \hat{\Pi}} C(\pi)$, *i.e. one may w.l.o.g. restrict oneself to policies belonging to $\hat{\Pi}$.*

For the remainder of the chapter, we thus consider the relevant optimization only over policies belonging to $\hat{\Pi}$, i.e.

$$\inf_{\pi \in \hat{\Pi}} C(\pi). \tag{19}$$

Recall that a stationary Markov policy is one that places orders only based on the current state information (i.e. $\hat{I}_t$ and $\mathcal{R}^t$), but independent of the current time period t and process history. It is generally well-known that for many inventory problems of interest, the relevant long-run-average optimization problems admit optimal stationary Markov policies, where such results typically follow from the general theory of infinite-horizon Markov decision processes (cf. [175, 172]). Explicit sufficient conditions for a rich family of inventory models to admit such an optimal policy were given in [95], where these conditions were verified to hold for many models of interest (e.g. lost-sales inventory models with positive lead times). Furthermore, it was commented in [183] that although [95] does not explicitly verify that their conditions

hold for the dual-sourcing problem, the relevant results still hold, and a proof was sketched under some additional technical conditions. A proof was also sketched in [91] under the assumption that demand is bounded. For simplicity and clarity of exposition, in the remainder of this chapter we simply assume the existence of such an optimal stationary Markov policy for Problem (19). Furthermore, we also assume that of these optimal stationary policies, there exists at least one whose corresponding induced Markov chain converges in distribution and in expectation to a unique stationary measure. Again, the existence of such optimal policies is to be expected from the basic theory of Markov chains and Markov decision processes, and we refer the reader to [175, 172, 156, 7, 130, 155], as well as the excellent recent survey of [5], for further details. We also note that the related work of [202] made a similar assumption in the context of lost-sales models. Although our precise assumptions could in principle be relaxed, e.g. to only requiring that for each $\epsilon > 0$ the stated assumptions hold for some (possibly randomized) policy which is ϵ-close to optimal (as opposed to an exactly optimal deterministic policy), we do not pursue such a generalization here for the sake of brevity and clarity of exposition.

Assumption 1 *Problem (19) has an optimal stationary Markov policy for all L, whose corresponding induced Markov chain converges in distribution and in expectation to a stationary measure when the initial inventory is 0, and no initial orders have been placed from either R or E. Also, we require that D is non-negative and integrable, satisfying $\mathbb{E}[D] < \infty$, and non-degenerate (i.e. not w.p.1 equal to its mean).*

Let $\pi^{L,*}$ be such an optimal policy, $\left(\hat{I}^{L,*}, \mathcal{R}^{L,*}\right)$ be a vector distributed as the stationary measure of the corresponding Markov Chain (with all r.v.s constructed on a common probability space with the appropriate joint distribution, independent of $\{D_t\}_{t \geq 1}, \{D'_t\}_{t \geq 1}$), and $D_{[t]} \triangleq (D_1, \ldots, D_t)$, where $D_{[0]}$ denotes the empty set $\emptyset$. Let $q_t^{L,*} = q_t^{L,*}(\hat{I}^{L,*}, \mathcal{R}^{L,*}, D_{[t-1]})$ denote the quantity ordered from E by π^* in period t if the initial expedited inventory position equals $\hat{I}^{L,*}$, the initial truncated regular

50

pipeline equals $\mathcal{R}^{L,*}$, and the first $t-1$ demands equal $D_{[t-1]}$. As we will be interested primarily in the setting that $L \to \infty$ with L_0, c_R, c_E, b, h held fixed, we will generally supress notationally dependence on these parameters, only making the dependence on L explicit. For two r.v.s X, Y, let $X \sim Y$ denote equivalence in distribution between X and Y. It follows from stationarity that

$$\text{OPT}(L) = c_R \mathbb{E}\left[\mathcal{R}_1^{L,*}\right] + c_E \mathbb{E}\left[q_1^{L,*}\right] + \mathbb{E}\left[G\left(\hat{I}^{L,*} + q_1^{L,*} - \sum_{i=1}^{L_0+1} D_i\right)\right]; \qquad (20)$$

$$\mathbb{E}[\mathcal{R}_1^{L,*}] + \mathbb{E}\left[q_1^{L,*}\right] = \mathbb{E}[D_1]. \qquad (21)$$

Combining (20) and (21), and w.l.o.g. assuming $c_R = 0$ (cf. [183]), we have

$$\text{OPT}(L) = c\left(\mathbb{E}[D] - \mathbb{E}[\mathcal{R}_1^{L,*}]\right) + \mathbb{E}\left[G\left(\hat{I}^{L,*} + q_1^{L,*} - \sum_{i=1}^{L_0+1} D_i\right)\right]. \qquad (22)$$

3.2.2 TBS policy

In this section, we formally introduce the family of TBS policies, and characterize the "best" TBS policy. As a notational convenience, let us define all empty sums to equal zero, empty products to equal one, and $\mathbb{I}(A)$ denote the indicator of the event A. A TBS policy $\pi_{r,S}$ with parameters (r, S) is defined (cf. [105]) as the policy that places a constant order r from R in every period, and follows an order-up-to rule from E which in each period raises the expedited inventory position to S (if it is below S), and otherwise orders nothing. More formally, under this policy $q_t^R = r$, and $q_t^E = \max(0, S - \hat{I}_t)$, for all t.

Let $I^\infty(r) \triangleq \sup_{j \geq 0}\left(jr - \sum_{i=1}^{j} D_i\right)$. In that case, it follows from the results of [105] that

$$C(\pi_{r,S}) = c(\mathbb{E}[D] - r) + \mathbb{E}\left[G\left(I^\infty(r) + S - \sum_{i=1}^{L_0+1} D_i'\right)\right]. \qquad (23)$$

Note that for each r, the minimization problem $\inf_{S \in \mathbb{R}} C(\pi_{r,S})$ is equivalent to a standard one-period newsvendor problem. Furthermore, defining $F^\infty(r) \triangleq \min_{S \in \mathbb{R}} C(\pi_{r,S})$,

it is proven in [105] that $F^\infty(r)$ is convex in r on $(-\infty, \mathbb{E}[D])$. Combining the above with standard results for single-server queues (cf. [7]) and (18), we conclude that there exists at least one pair (r^*, S^*) such that $r^* \in \arg\min_{0 \le r \le \mathbb{E}[D]} F^\infty(r)$ and $S^* \in \arg\min_{S \in \mathbb{R}} C(\pi_{r^*,S})$; that this pair defines the TBS policy with least long-run-average cost; and that this pair can be computed efficiently by solving a convex program which is independent of the larger lead time L.

3.2.3 Main result

Our main result shows that the best TBS policy is asymptotically optimal as $L \to \infty$.

Theorem 3 *Under Assumption 1,* $\lim_{L \to \infty} OPT(L) = C(\pi_{r^*,S^*})$.

3.3 Proof of Theorem 3

3.3.1 Lower bound for the optimal cost

In this section, we prove a lower bound for $OPT(L)$ by extending the steady-state and convexity approach of [202] to the dual-sourcing setting. We note that here our lower bound will involve a non-trivial optimization over measurable functions, in contrast to the bounds used in [202] which were of a static nature. From stationarity, for each $k = 1, \ldots, L - L_0$,

$$\hat{I}^{L,*} + \sum_{i=1}^{k-1}(q_i^{L,*} + \mathcal{R}_i^{L,*} - D_i) + q_k^{L,*} - \sum_{i=k}^{k+L_0} D_i \sim \hat{I}^{L,*} + q_1^{L,*} - \sum_{i=1}^{L_0+1} D_i;$$

and for each $k = 1, \ldots, L - L_0 - 1$,

$$\mathbb{E}[\mathcal{R}_k^{L,*}] = \mathbb{E}[\mathcal{R}_1^{L,*}] \stackrel{\Delta}{=} r_L.$$

Combining the above with (22) implies that for any $\alpha \in (0,1)$, $OPT(L)$ equals

$$c\left(\mathbb{E}[D] - r_L\right) + \frac{1-\alpha}{1-\alpha^L} \sum_{k=1}^{L} \alpha^{k-1} \mathbb{E}\left[G\left(\hat{I}^{L,*} + q_1^{L,*} - \sum_{i=1}^{L_0+1} D_i\right)\right]$$

$$\ge \quad c\left(\mathbb{E}[D] - r_L\right)$$

$$+ (1-\alpha) \sum_{k=1}^{L-L_0} \alpha^{k-1} \mathbb{E}\left[G\left(\hat{I}^{L,*} + \sum_{i=1}^{k-1}(q_i^{L,*} + \mathcal{R}_i^{L,*} - D_i) + q_k^{L,*} - \sum_{i=k}^{k+L_0} D_i\right)\right].$$

52

Here we have introduced the discount factor α to implement the so-called "vanishing discount factor" approach to analyzing infinite-horizon Markov decision processes (MDP) (cf. [95]), which will allow for a simpler analysis when we pass to the limit as $L \to \infty$. Indeed, this discount factor will help us to analyze the lower bound which arises when we apply Jensen's inequality, as this lower bound will itself involve the solution to a non-trivial multi-stage dynamic optimization problem. We note that the lower bound which arose when related techniques were applied to single-sourcing systems with lost sales in [202] only involved a static optimization problem, and thus no such discount factor was introduced.

It then follows from the independence structure of the relevant r.v.s, and the measurability properties of $q_k^{L,*}$, that for each $k = 1, \ldots, L - L_0$,

$$\mathbb{E}\left[\hat{I}^{L,*} + \sum_{i=1}^{k-1}(q_i^{L,*} + \mathcal{R}_i^{L,*} - D_i) + q_k^{L,*} - \sum_{i=k}^{k+L_0} D_i \,\middle|\, D_{[k+L_0]}\right]$$

equals

$$\mathbb{E}[\hat{I}^{L,*}] + \sum_{i=1}^{k-1}(\mathbb{E}[q_i^{L,*}|D_{[i-1]}] + r_L - D_i) + \mathbb{E}[q_k^{L,*}|D_{[k-1]}] - \sum_{i=k}^{k+L_0} D_i.$$

Further combining with the convexity of G and Jensen's inequality for conditional expectations (which applies due to Assumption 1), we obtain the following result.

Proposition 1 *For any $\alpha \in (0,1)$ and $L \geq L_0 + 1$, $OPT(L) - c\left(\mathbb{E}[D] - r_L\right)$ is at least*

$$(1 - \alpha) \sum_{k=1}^{L-L_0} \alpha^{k-1} \mathbb{E}\left[G\left(\mathbb{E}[\hat{I}^{L,*}] - (L_0 + 1)r_L + \sum_{i=1}^{k-1}\left(\mathbb{E}[q_i^{L,*}|D_{[i-1]}] - (D_i - r_L)\right)\right.\right.$$
$$\left.\left. + \mathbb{E}[q_k^{L,*}|D_{[k-1]}] - \sum_{i=k}^{k+L_0}(D_i - r_L)\right)\right].$$
$$(24)$$

Note that (24) is the discounted cost incurred (during periods $L_0 + 1, \ldots, L$) by the policy ordering $\mathbb{E}[q_i^{L,*}|D_{[i-1]}]$ in period i, of a single-sourcing L-period backlog inventory problem with unit holding cost h, backorder cost b, zero ordering cost,

discount factor α, i.i.d. demand distributed as $D - r_L$ (which we note can be positive *or negative*), lead time L_0, and initial inventory position (initial net inventory plus all entries of the initial pipeline vector) $\mathbb{E}[\hat{I}^{L,*}] - (L_0 + 1)r_L$ (cf. [113]), normalized by $(1 - \alpha)$. Such models, and their optimal policies, have been studied in-depth in the literature (cf. [113, 213, 64]), and are well-understood (especially for the case of non-negative demand, cf. [213]). Let $\hat{\Pi}$ denote the family of all feasible non-anticipative policies for the aforementioned inventory problem (as it is typically defined, cf. [213]), i.e. those policies for which new orders are non-negative measurable functions of the realized demands, inventory levels, and ordering quantities in periods $1, \ldots, t - 1$, as well as the current time t. For $\pi \in \hat{\Pi}$, initial inventory position $x \in \mathbb{R}$, $r \in \mathbb{R}$, and $i \geq 1$, let $C_i^\pi(r, x)$ denote the cost incurred by policy π in the aforementioned inventory problem in period $i + L_0$, if the demand in each period is i.i.d. distributed as $D - r$ (with the leadtime L_0 and costs b, h as above). For $x \in \mathbb{R}, r \in \mathbb{R}, \alpha \in (0, 1), n \geq 1$, let us define

$$V_\alpha^n(r, x) \overset{\Delta}{=} \inf_{\pi \in \hat{\Pi}} \mathbb{E}\left[\sum_{i=1}^n \alpha^{i-1} C_i^\pi(r, x)\right]; \tag{25}$$

and

$$V_\alpha^\infty(r, x) \overset{\Delta}{=} \inf_{\pi \in \hat{\Pi}} \mathbb{E}\left[\sum_{i=1}^\infty \alpha^{i-1} C_i^\pi(r, x)\right]. \tag{26}$$

As a notational convenience, we define

$$V_\alpha^0(r, x) = 0, \quad V_\alpha^n(r, -\infty) \overset{\Delta}{=} \inf_{x \in \mathbb{R}} V_\alpha^n(r, x), \quad V_\alpha^\infty(r, -\infty) \overset{\Delta}{=} \inf_{x \in \mathbb{R}} V_\alpha^\infty(r, x).$$

Then combining the above, we derive the following lower bound for $\mathrm{OPT}(L)$.

Lemma 7 *Under Assumption 1, for all $\alpha \in (0, 1)$ and $L \geq L_0 + 1$,*

$$OPT(L) \geq c(\mathbb{E}[D] - r_L) + (1 - \alpha)V_\alpha^{L-L_0}(r_L, -\infty). \tag{27}$$

The remainder of the proof involves demonstrating a certain interchange-of-limit results for the right-hand-side (r.h.s.) of (27). We now briefly describe the associated logic informally, and formalize all arguments in the next section. Let $r_\infty \overset{\Delta}{=}$

$\limsup_{L\to\infty} r_L$. Examining the r.h.s. of (27) along a subsequence on which r_L converges to r_∞ and taking limits, one shows that for any fixed $\alpha \in (0,1)$,

$$\lim_{L\to\infty} OPT(L) \geq c(\mathbb{E}[D] - r_\infty) + (1-\alpha)V_\alpha^\infty(r_\infty, -\infty).$$

One then shows that the infinite-horizon problem associated with $V_\alpha^\infty(r_\infty, -\infty)$ has an optimal policy which is stationary, Markov, and of order-up-to type, say to level $S_\alpha^\infty(r_\infty)$. It follows that for any fixed $\alpha \in (0,1)$, $V_\alpha^\infty(r_\infty, -\infty)$ is the expected infinite-horizon discounted cost incurred by the TBS policy with parameters $r_\infty, S_\alpha^\infty(r_\infty)$, but possibly initialized *not according to the stationary distribution of the associated inventory process, but in the state which minimizes this discounted expected infinite-horizon cost.* One then appropriately bounds the difference in cost incurred under these two different initializations uniformly in α. Finally, letting $\alpha \uparrow 1$ will demonstrate that as $L \to \infty$, there exist TBS policies performing arbitrarily close to optimal. This will imply asymptotic optimality of the best TBS policy (with parameters r^*, S^*), completing the proof of our main result Theorem 3.

3.3.2 Interchange of limits and proof of Theorem 3

We now complete the proof of Theorem 3 by formalizing the interchange-of-limits argument sketched at the end of Section 3.3.1. Such interchange arguments are standard in the literature on Markov decision processes and infinite-horizon inventory control problems (cf. [99, 175, 172, 64, 59, 95]). We note that the somewhat non-standard aspect of the interchange of limits which we must demonstrate is that the demand in each period is distributed as $D - r_L$, and thus *may be negative.* As such, the original arguments showing such an interchange for the analogous inventory models when demand is non-negative (cf. [99]) do not directly apply. The possibility of negative demand also makes the verification of the conditions of general theorems which validate such an interchange (cf. [175, 172]) somewhat involved, even when these theorems are customized to the inventory setting (cf. [147, 95]). We note that

the verification of closely related interchange-of-limits results have arisen recently in the context of analyzing inventory systems with returns, which reduce to standard inventory systems where demand can be positive or negative (cf. [64]). However, those results (which verify the technical conditions of [175]) do not seem to extend immediately to our case, and further seem to require that the demand and ordering quantities take integer values. In light of the above, and for the sake of clarity and completeness, we now provide a self-contained proof of the desired interchange, which (combined with Lemma 7) will complete the proof of our main result Theorem 3.

We begin by stating some well-known properties of $V_\alpha^n(r,x)$ and $V_\alpha^\infty(r,x)$, which follow from the results of [105], [113] and [170]. We note that although in some cases the proofs there are only explicitly given for the case of non-negative demand, as noted in [88] and [64], the arguments carry over to the general case (in which demand may be negative) with only trivial modification.

Lemma 8 ([105, 170]) *For all $\alpha \in (0,1), r, x \in \mathbb{R}$, and $n \geq 2$,*

$$V_\alpha^n(r,x) = \inf_{y \geq x} \left(\mathbb{E}\left[G\big(y - \sum_{k=n}^{L_0+n} (D_k - r)\big) \right] + \alpha \mathbb{E}\left[V_\alpha^{n-1}\big(r, y - (D_{L_0+n} - r)\big) \right] \right).$$

Furthermore, $V_\alpha^n(r,x)$ is: a convex (and thus also continuous) function of x on $\mathbb{R}$ for each fixed n, r; a continuous function of r on $\mathbb{R}$ for each fixed n, x; an increasing function of x on $\mathbb{R}$ for each fixed n, r; and an increasing function of n on Z^+ for each fixed x, r. In addition, the infinite-horizon problem stated in the r.h.s. of (26) admits an optimal stationary Markov policy.

In preparation for demonstrating the desired interchange-of-limits, we now appropriately bound the optimal value, and set of minimizers, of $V_\alpha^n(r,x)$, uniformly in n. For $\alpha \in (0,1)$ and $r \in \mathbb{R}$, let $\overline{S}_\alpha^n(r)$ denote the supremum of the set of minimizers (with respect to x) of $V_\alpha^n(r,x)$, where we note that a straightforward contradiction demonstrates that $\overline{S}_\alpha^n(r) \in (-\infty, \infty)$ for each α, n, r; and it follows from Lemma 8

that $V_\alpha^n(r, -\infty) = V_\alpha^n\big(r, \overline{S}_\alpha^n(r)\big)$. Then we prove the following uniform bounds, and defer the proof to the appendix (Section 3.5).

Lemma 9 *For $\alpha \in (0,1)$ and $r, x \in \mathbb{R}$, $\sup_{n \geq 1} V_\alpha^n(r, x) < \infty$. Also, for each $\alpha \in (0,1)$, there exist finite-valued strictly positive functions (of r) $\overline{S}_\alpha(r)$, $\epsilon_\alpha(r)$, which are continuous (in r) on $\mathbb{R}$, with the following properties. For all $n \geq 1$: $|\overline{S}_\alpha^n(r)| < \overline{S}_\alpha(r)$; and for all $y \notin [-\overline{S}_\alpha(r), \overline{S}_\alpha(r)]$,*

$$\mathbb{E}\left[G\Big(y - \sum_{k=n}^{L_0+n}(D_k - r)\Big)\right] + \alpha\mathbb{E}\left[V_\alpha^{n-1}\big(r, y - (D_{L_0+n} - r)\big)\right] \geq V_\alpha^n\big(r, \overline{S}_\alpha^n(r)\big)) + \epsilon_\alpha(r).$$

From Lemma 9 we derive the following two corollaries, whose proofs we defer to the appendix (Section 3.5).

Corollary 2 *For $\alpha \in (0,1)$,*

$$\lim_{L \to \infty} OPT(L) \geq c(\mathbb{E}[D] - r_\infty) + (1 - \alpha)\lim_{L \to \infty} V_\alpha^L\big(r_\infty, -2\overline{S}_\alpha(r_\infty)\big).$$

Corollary 3 *For all $\alpha \in (0,1)$ and $r, x \in \mathbb{R}$, $V_\alpha^\infty(r, x) = \lim_{n \to \infty} V_\alpha^n(r, x)$. Furthermore, for all $\alpha \in (0,1)$ and $r \in \mathbb{R}$, $V_\alpha^\infty(r, x)$ is a finite-valued, convex, and non-decreasing function of x on $\mathbb{R}$. Letting $S_\alpha^\infty(r)$ denote the supremum of the set of minimizers (in x) of $V_\alpha^\infty(r, x)$, it holds that $|S_\alpha^\infty(r)| \leq \overline{S}_\alpha(r)$, and the infinite-horizon problem stated in the r.h.s. of (26) admits an optimal stationary base-stock policy, with order-up-to level $S_\alpha^\infty(r)$.*

We now formally define the Markov process representing the inventory position process under such an optimal stationary base-stock policy. Let $S_\alpha \triangleq S_\alpha^\infty(r_\infty)$, and $\{X_k^\alpha, k \geq 1\}$ denote the following Markov process. X_1^α equals S_α. For all $k \geq 1$, $X_{k+1}^\alpha = \max\big(X_k^\alpha + r_\infty - D_k, S_\alpha\big)$. Let $W_k \triangleq \sum_{j=1}^k (r_\infty - D_j)$, $Z_k \triangleq \max_{i \in [0, k-1]} W_i$, $Z_\infty \triangleq \sup_{i \geq 0} W_i$, $M_k \triangleq \mathbb{E}[Z_k]$, $M_\infty \triangleq \mathbb{E}[Z_\infty]$. It follows from the well-known analysis of the single-server queue using Lindley's recursion (cf. [7]) that $X_k^\alpha \sim S_\alpha + Z_k$; and $X_\infty^\alpha \triangleq \lim_{k \to \infty} X_k^\alpha$ is a well-defined r.v. distributed as $S_\alpha + Z_\infty$. Combining these definitions with Lemma 8 and Corollaries 2 and 3, we conclude the following.

Corollary 4 *For* $\alpha \in (0, 1)$,

$$\lim_{L \to \infty} OPT(L) \geq c(\mathbb{E}[D] - r_\infty) + (1 - \alpha) \sum_{k=1}^{\infty} \alpha^{k-1} \mathbb{E}\Big[G\big(S_\alpha + Z_k - \sum_{i=1}^{L_0+1}(D_i' - r_\infty)\big)\Big].$$

We now briefly review some useful (and generally well-known) properties of Z_k, which we will use to complete the proof of our main results, and defer the proof to the appendix (Section 3.5).

Lemma 10 *If* $r_\infty < \mathbb{E}[D]$, *then* $M_\infty < \infty$. *Also, there exists* $\theta^* > 0$ *such that* $\gamma_* \triangleq \mathbb{E}\big[\exp\big(\theta^*(r_\infty - D)\big)\big] \in (0, 1)$, *and* $M_\infty - M_n \leq \big(\theta^*(1 - \gamma_*)\big)^{-1}\gamma_*^n$ *for all* $n \geq 2$. *Alternatively, if* $r_\infty \geq \mathbb{E}[D]$, *then* Z_∞ *is almost surely infinite and* $M_\infty = \infty$. *In both cases,* $\{M_k, k \geq 1\}$ *is strictly increasing,* $M_\infty = \lim_{k \to \infty} M_k$, *and for all* $i \geq j \geq 1$, $M_i - M_j = \sum_{k=j}^{i-1} k^{-1}\mathbb{E}[\max(0, W_k)]$. *Also, if* $r_\infty = \mathbb{E}[D]$, *there exists a strictly positive finite constant* C^* *(depending only on* D*) such that for all* $i \geq j + 1 \geq 3$, $M_i - M_j \geq C^*\big(i^{\frac{1}{2}} - j^{\frac{1}{2}}\big)$.

We now combine Corollary 4 with Lemma 10 to complete the proof of our main result.

Proof. [Proof of Theorem 3] First, we prove that $r_\infty < \mathbb{E}[D]$, which we will need to demonstrate the desired uniform convergence in α. We note that this will require somewhat subtle arguing, since we must show that otherwise (i.e. if $r_\infty = \mathbb{E}[D]$) the cost associated with the relevant inventory process grows too quickly (as $\alpha \uparrow 1$) even when initialized to S_α, which may be converging to $-\infty$ as $\alpha \uparrow 1$. Suppose for contradiction that $r_\infty = \mathbb{E}[D]$. In this case, it follows from Corollary 4, (18), and Jensen's inequality that for all $\alpha \in (0, 1)$,

$$\lim_{L \to \infty} OPT(L) \geq \min(b, h) \inf_{S \in \mathbb{R}} \sum_{k=1}^{\infty} (1 - \alpha)\alpha^{k-1}\big|S + M_k\big|. \tag{28}$$

Let G_α denote a geometrically distributed r.v. with success probability $1 - \alpha$, independent of $\{Z_k, k \geq 1\}$, and $m_\alpha \triangleq \lceil -\frac{1}{\log_2(\alpha)} \rceil$ denote a median of G_α. Note that

the memoryless property implies $\mathbb{P}(G_\alpha \geq 2m_\alpha) = \frac{1}{4}$, and that we may interpret the r.h.s. of (28) as an appropriate single-stage newsvendor problem (with ordering level S and demand distributed as M_{G_α}). We conclude from Lemma 10, well-known results for the newsvendor problem (cf. [213]), and the memoryless property that for all sufficiently large $\alpha \in (0,1)$, $\lim_{L\to\infty} OPT(L)$ is at least

$$\min(b,h)\mathbb{E}\big[\big|M_{m_\alpha}-M_{G_\alpha}\big|\big] \geq \frac{1}{4}\min(b,h)\big(M_{2m_\alpha}-M_{m_\alpha}\big) \geq \frac{1}{4}\min(b,h)C^*\big((2m_\alpha)^{\frac{1}{2}}-m_\alpha^{\frac{1}{2}}\big).$$

As it is easily verified that $\lim_{\alpha\uparrow 1}\big((2m_\alpha)^{\frac{1}{2}}-m_\alpha^{\frac{1}{2}}\big) = \infty$, we conclude that if $r_\infty = \mathbb{E}[D]$, then $\lim_{L\to\infty} OPT(L) = \infty$. However, in this case a contradiction is easily reached by considering the TBS policy $\pi_{0,0}$, which incurs long-run average cost $C(\pi_{0,0}) < \infty$ for all $L \geq L_0 + 1$, completing the proof that $r_\infty < \mathbb{E}[D]$.

In that case, it follows from (23) that for all $\alpha \in (0,1)$, $C(\pi_{r_\infty,S_\alpha+(L_0+1)r_\infty})$ equals

$$c(\mathbb{E}[D] - r_\infty) + \mathbb{E}\left[G\left(S_\alpha + Z_\infty + (L_0+1)r_\infty - \sum_{i=1}^{L_0+1} D_i'\right)\right]$$

$$= c(\mathbb{E}[D] - r_\infty) + (1-\alpha)\sum_{k=1}^\infty \alpha^{k-1}\mathbb{E}\left[G\left(S_\alpha + Z_\infty - \sum_{i=1}^{L_0+1}(D_i' - r_\infty)\right)\right].$$

Combining with Corollary 4, Lemma 10, and (18), we conclude that for all $\alpha \in (0,1)$, $C(\pi_{r^*,S^*}) - \lim_{L\to\infty} OPT(L)$ is at most

$$C(\pi_{r^*,S^*}) - \lim_{L\to\infty} OPT(L) \geq (1-\alpha)\sum_{k=1}^\infty \alpha^{k-1}\left(\mathbb{E}\left[G\left(S_\alpha + Z_\infty - \sum_{i=1}^{L_0+1}(D_i' - r_\infty)\right)\right]\right.$$

$$\left. - \mathbb{E}\left[G\left(S_\alpha + Z_k - \sum_{i=1}^{L_0+1}(D_i' - r_\infty)\right)\right]\right). \tag{29}$$

Note that (29) is at most

$$\max(b,h)(1-\alpha)\left(M_\infty + \sum_{k=1}^\infty \alpha^{k-1}\big(\theta^*(1-\gamma_*)\big)^{-1}\gamma_*^k\right)$$

$$= \max(b,h)\left((1-\alpha)M_\infty + \gamma_*\big(\theta^*(1-\gamma_*)\big)^{-1}\frac{1-\alpha}{1-\gamma_*\alpha}\right), \tag{30}$$

which converges to 0 as $\alpha \uparrow 1$ completes the proof. ∎

3.4   Conclusion

In this chapter, we proved that when the lead time of the express source is held fixed, a simple TBS policy is asymptotically optimal for the dual-sourcing inventory problem as the lead time of the regular source grows large. Our results provide a solid theoretical foundation for several conjectures and numerical experiments appearing previously in the literature regarding the good empirical performance of such policies. Furthermore, the simple TBS policy performs nearly optimally exactly when standard DP-based methodologies become intractable due to the curse of dimensionality. In addition, since the "best" TBS policy can be computed by solving a convex program that does not depend on the lead time of the regular source, and is easy to implement, our results lead directly to very efficient algorithms with asymptotically optimal performance guarantees. Perhaps most importantly, since many companies are already implementing such TBS policies, our results provide strong theoretical support for the widespread use of TBS policies in practice.

This work leaves many interesting directions for future research. First, it would be interesting to investigate the rate of convergence to optimality of TBS policies as the lead time grows large, especially in light of their use in practical settings. Such an analysis would seem to involve estimates for the rate of convergence of finite horizon inventory optimization problems to their infinite horizon counterparts, which has been previously investigated for related systems (cf. [89, 90]). We do note that a detailed analysis of the proofs of our main results in principle yields a computable bound for this rate. For example, one can minimize the r.h.s. of (27) over $r_L \in \big[0, \mathbb{E}[D]\big]$, and maximize over $\alpha \in (0, 1)$, to yield a (relatively) easy-to-compute bound for each $L \geq L_0 + 1$ (which can be further improved by more carefully optimizing our approach and bounds for this purpose). However, a more precise theoretical analysis of the performance of TBS policies for small-to-moderate lead times, analogous to the exponential rate of convergence to optimality of the constant-order policy for

lost-sales inventory models identified in [202] (i.e., chapter 2 in this thesis), seems to require fundamentally new ideas.

Second, and related to the aforementioned discussion as regards the rate of convergence to optimality of TBS policies, it would be interesting to identify other more sophisticated algorithms which perform better for small-to-moderate lead times, yet remain efficient to implement. Indeed, it remains an interesting open question to better understand the trade-off between algorithmic run-time and acheivable performance guarantees in this context, i.e. how complex an algorithm is required to "exploit" the weak correlations which persist even as the lead time grows large. In the context of dual-sourcing, potential algorithms here include: the so-called dual-sourcing smoothing policies recently studied in [29]; affine policies more generally (cf. [14], [20]), of which dual-sourcing smoothing policies are a special case; the single index and dual index policies discussed earlier; or the dual-balancing policies analyzed in [123]. On a related note, it would be quite interesting to analyze "hybrid" algorithms, which could e.g. solve a large dynamic program when the lead time is small, and gradually transition to using simpler heuristics as the lead time grows large; or combine different heuristics depending on the specific problem parameters.

On a final note, combined with the results of [75] and [202], our methodology lays the foundations for a completely new approach to analyzing inventory models with large lead times. So far, this approach has been successful in yielding key insights and efficient algorithms for two settings previously believed intractable: lost sales models with large lead times, and dual-sourcing models with large lead time gap. We believe that our techniques have the potential to make similar progress on many other difficult supply chain optimization problems of practical relevance in which there is a lag between when policy decisions are made and when those decisions are implemented. This includes both more realistic variants of the lost-sales and dual-sourcing models considered so far (e.g. models with distributional

dependencies, parameter uncertainty, complex network structure, and more accurate modeling of costs), as well as fundamentally different models (e.g. inventory systems with remanufacturing when the manufactured and remanufactured lead times differ, cf. [211]; multi-echelon systems with lost sales and positive lead times, cf. [94]; or models with perishable goods). In closing, we note that our approach can more generally be viewed as a methodology to formalize the notion that when there is a high level of uncertainty and randomness in one's supply chain, even simple policies perform nearly as well as very sophisticated policies, since no algorithm can "beat the noise". Exploring this concept from a broader perspective may be fruitful in yielding novel algorithms and insights for a multitude of problems in operations management and operations research.

3.5 Appendix

Proof. [Proof of Lemma 9] By evaluating the policy which never orders, we conclude that for all $\alpha \in (0,1)$, $r, x \in \mathbb{R}$, $\sup_{n \geq 1} V_\alpha^n(r, x)$ is at most

$$\mathbb{E}\left[\sum_{i=1}^{\infty} \alpha^{i-1} G\left(x - \sum_{j=1}^{i}(D_j - r) - \sum_{k=i+1}^{L_0+i}(D_k - r)\right)\right],$$

which by (18) is itself bounded by

$$\max(b, h)(|x|+|r|+\mathbb{E}[D])\sum_{i=1}^{\infty}(i+L_0)\alpha^{i-1} \leq 2(L_0+1)\max(b, h)(|x|+|r|+\mathbb{E}[D])(1-\alpha)^{-2}.$$

Combining with (18) and a straightforward calculation, it follows that for any $r, y \in \mathbb{R}$ such that $|y| \geq \overline{S}_\alpha(r) \triangleq 4(L_0 + 1)\frac{\max(b,h)}{\min(b,h)}(|r| + \mathbb{E}[D])(1 - \alpha)^{-2}$, and all $n \geq 1$, one has that $\mathbb{E}[G(y - \sum_{k=n}^{L_0+n}(D_k - r))] - V_\alpha^n(r, 0) \geq \epsilon_\alpha(r) \triangleq (L_0 + 1)\max(b, h)\mathbb{E}[D]$. Combining the above completes the proof. ∎

Proof. [Proof of Corollary 2] Let $\{i_k, k \geq 1\}$ be a strictly increasing subsequence of positive numbers such that $i_1 \geq L_0 + 1$, and $\lim_{k\to\infty} r_{i_k} = r_\infty$ (existence follows from the definition of $\limsup$). It follows from Lemmas 7, 8, and 9 that for all $k \geq 1$,

$$OPT(i_k) \geq c(\mathbb{E}[D] - r_{i_k}) + (1 - \alpha)V_\alpha^{i_k - L_0}\left(r_{i_k}, -\overline{S}_\alpha(r_{i_k})\right).$$

It follows from the continuity (in r) of $\overline{S}_\alpha(r)$, and monotonicity (in x) of $V_\alpha^n(r, x)$, that there exists $k_0 < \infty$ such that for all $k \geq k_0$,

$$OPT(i_k) \geq c(\mathbb{E}[D] - r_{i_k}) + (1 - \alpha)V_\alpha^{i_k - L_0}\big(r_{i_k}, -2\overline{S}_\alpha(r_\infty)\big).$$

It then follows from the monotonicity (in n) of $V_\alpha^n(r, x)$ that for all $k \geq k_0$ and $L \leq i_k - L_0$,

$$OPT(i_k) \geq c(\mathbb{E}[D] - r_{i_k}) + (1 - \alpha)V_\alpha^L\big(r_{i_k}, -2\overline{S}_\alpha(r_\infty)\big).$$

Fixing $L \geq 1$, letting $k \to \infty$, and applying the continuity (in r) of $V_\alpha^n(r, x)$ completes the proof. ∎

Proof. [Proof of Corollary 3] We first demonstrate that $V_\alpha^\infty(r, x) = \lim_{n \to \infty} V_\alpha^n(r, x)$. The existence of the corresponding limit follows from the monotonicity (in n) guaranteed by Lemma 8. That $V_\alpha^\infty(r, x) \geq \lim_{n \to \infty} V_\alpha^n(r, x)$ for all $\alpha \in (0, 1)$ and $r, x \in \mathbb{R}$ follows immediately from the definitions of the associated optimization problems. To prove the other direction, we note that for any fixed $n \geq 1$, it follows from the convexity ensured by Lemma 8 that there exists an optimal policy π for the problem stated in the r.h.s. of (25) of base-stock form, with order-up-to levels $C_1, \ldots, C_n$ (i.e. order up to level C_i in period i if the pre-order inventory level is below C_i, otherwise order nothing). Furthermore, it follows from Lemma 9 that $\max_{i=1,\ldots,n} |C_i| \leq \overline{S}_\alpha(r)$. Now, consider the policy π' that orders up to level C_i in period i if the pre-order inventory level is below C_i and otherwise orders nothing in periods $i = 1, \ldots, n$; and orders nothing in all remaining periods, irregardless of the inventory level. Note that under policy π', w.p.1 the absolute value of the inventory position at the end of period i is at most $|x| + \overline{S}_\alpha(r) + i|r| + \sum_{k=1}^i D_k$. It then follows from an argument nearly identical to that presented in our proof of Lemma 9 that

$$\mathbb{E}\left[\sum_{i=1}^\infty \alpha^{i-1} C_i^{\pi'}(r, x)\right] - V_\alpha^n(r, x) \leq \max(b, h)(\overline{S}_\alpha(r) + |x| + |r| + \mathbb{E}[D]) \sum_{i=n+1}^\infty (i + L_0)\alpha^{i-1}.$$
$$(31)$$

That $V_\alpha^\infty(r,x) \le \lim_{n\to\infty} V_\alpha^n(r,x)$ then follows from the fact that the r.h.s. of (31) converges to 0 as $n \to \infty$. The remainder of the corollary follows by combining the above with Lemma 9, and applying the fact that convexity and monotonicity are preserved under limits. ∎

Proof. [Proof of Lemma 10] The entirety of the lemma, barring the lower bound involving C^*, follows by combining generally well-known results for generating functions, large deviations, single-server queues, and recurrent random walks (cf. [186, 116, 65, 7, 202]), and we omit the details. We now prove the lower bound involving C^*, which we note would follow from well-known weak-convergence results under additional assumptions on D (e.g. finite variance, cf. [55]). Let $\{A_i^+, i \ge 1\}$ denote an i.i.d. sequence of r.v.s distributed as $\mathbb{E}[D] - D$ conditioned on the event $\{\mathbb{E}[D] > D\}$, and $\{A_i^-, i \ge 1\}$ denote an i.i.d. sequence of r.v.s distributed as $D - \mathbb{E}[D]$ conditioned on the event $\{D \ge \mathbb{E}[D]\}$. Let B_k denote a binomially distributed r.v. with parameters $k, p \overset{\Delta}{=} \mathbb{P}(\{\mathbb{E}[D] > D\})$, independent of $\{A_i^+, i \ge 1\}$ and $\{A_i^-, i \ge 1\}$. Note that we may construct W_k on an appropriate probability space such that $W_k = \sum_{i=1}^{B_k} A_i^+ - \sum_{i=1}^{k-B_k} A_i^-$, in which case (by non-negativity) $\mathbb{E}[\max(0, W_k)]$ is at least

$$\mathbb{E}\left[\sum_{i=1}^{B_k} A_i^+ - \sum_{i=1}^{k-B_k} A_i^- \,\middle|\, \left\{B_k \ge pk + \left(p(1-p)k\right)^{\frac{1}{2}}\right\}\right]\mathbb{P}\left(\frac{B_k - pk}{\left(p(1-p)k\right)^{\frac{1}{2}}} \ge 1\right).$$

Furthermore, since $p\mathbb{E}[A_1^+] = (1-p)\mathbb{E}[A_1^-]$, it follows from non-negativity and independence that

$$\mathbb{E}\left[\sum_{i=1}^{B_k} A_i^+ - \sum_{i=1}^{k-B_k} A_i^- \,\middle|\, \left\{B_k \ge pk + \left(p(1-p)k\right)^{\frac{1}{2}}\right\}\right]$$
$$\ge \quad \left(pk + \left(p(1-p)k\right)^{\frac{1}{2}}\right)\mathbb{E}[A_1^+] - \left((1-p)k - \left(p(1-p)k\right)^{\frac{1}{2}}\right)\mathbb{E}[A_1^-]$$
$$= \quad \left(p(1-p)k\right)^{\frac{1}{2}}\left(\mathbb{E}[A_1^+] + \mathbb{E}[A_1^-]\right).$$

Combining the above with the central limit theorem, we conclude that

$$\liminf_{k\to\infty} \frac{\mathbb{E}[\max(0, W_k)]}{k^{\frac{1}{2}}} \ge \left(p(1-p)\right)^{\frac{1}{2}}\left(\mathbb{E}[A_1^+] + \mathbb{E}[A_1^-]\right)\liminf_{k\to\infty} \mathbb{P}\left(\frac{B_k - pk}{\left(p(1-p)k\right)^{\frac{1}{2}}} \ge 1\right) > 0.$$

64

It then follows from the strict positivity of $E[\max(0, W_k)]$ for all $k \geq 2$ that there exists $C' > 0$ such that $\mathbb{E}[\max(0, W_k)] \geq C'k^{\frac{1}{2}}$ for all $k \geq 2$. Thus for all $i \geq j \geq 2$,

$$
\begin{aligned}
M_i - M_j \;&=\; \sum_{k=j}^{i-1} k^{-1} \mathbb{E}[\max(0, W_k)] \\
&\geq\; C' \sum_{k=j}^{i-1} k^{-\frac{1}{2}} \;\geq\; C' \int_j^i x^{-\frac{1}{2}} dx \;=\; 2C'\big(i^{\frac{1}{2}} - j^{\frac{1}{2}}\big).
\end{aligned}
$$

Combining the above completes the proof. ∎

CHAPTER IV

TIME (IN)CONSISTENCY OF MULTI-STAGE DISTRIBUTIONALLY ROBUST INVENTORY MODELS WITH MOMENT CONSTRAINTS

This chapter is based on [204].

4.1   Introduction and literature review

The news vendor problem, used to analyze the trade-offs associated with stocking an inventory, has its origin in a seminal paper by [53]. In its classical formulation, the problem is stated as a minimization of the expected value of the relevant ordering, backorder, and holding costs. Such a formulation requires a complete specification of the probability distribution of the underlying demand process. 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 [168] 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. [71, 72, 69, 73, 151, 208, 70, 148, 37, 174, 85, 212]). For a more general overview of risk analysis for news vendor and inventory models we can refer, e.g., to [2] and [44]. 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 (cf. [114, 14, 22, 12, 20, 33, 68]) and Bayesian (cf. [169, 171, 127, 125, 117]) paradigms.

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 (see, e.g., [213] and the references therein). Recently, distributionally robust variants of such multistage problems have begun to receive attention in the literature (cf. [73, 2, 43, 174, 180, 117]). 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 [187] - for a more recent overview we refer the reader to the recent survey by [57], and the references therein. Questions of time consistency have also attracted attention in the mathematical finance literature, in the context of assessing the risk and value of investments over time, and have played an important role in the associated theory of coherent risk measures (cf. [196, 6, 161, 42, 165]). These concepts have also been studied from the perspective of robust control across various academic communities (cf. [86, 103, 145, 79, 32, 200]). Recently, these concepts have also begun to receive attention in the setting of inventory control (cf. [38, 39, 205]).

In this chapter, we will consider questions of time (in)consistency in the context of managing an inventory over time. We will give a formal definition of time consistency, which is naturally suited to our framework, in Section 4.4. At this point we would like to give the following intuition. A multistage distributionally robust optimization problem can be viewed in two ways. In one formulation, the policy maker is allowed to recompute his/her policy choice after each stage (we will refer to this as the *multistage-dynamic formulation*), thus taking prior realizations of demand into consideration when performing the relevant minimax calculations at later stages. In that case it follows from known results that there exists a base-stock policy which is optimal. In the second formulation, the policy maker is not allowed to recompute his/her policy

after each stage (we will refer to this as *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 he/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.

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 dynamic programming structure. Outside of this setting, most of the literature focuses on discussing and demonstrating hardness of the underlying optimization problems (cf. [103, 145, 200]). 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. [103]).

4.1.1 Outline of chapter

The structure of the rest of the chapter is organized as follows. In Section 4.2, we review the single-stage classical and distributionally robust formulations and their properties, as well as Scarf's solution to the single-stage distributionally robust formulation and various generalizations. In Section 4.3, we discuss the extension to the multi-stage setting, formally define the multistage-static and multistage-dynamic distributionally robust formulations, and review the notion of rectangularity and its

relation to both our own formulations and robust Markov Decision Processes (MDP). In Section 4.4, we formally define time consistency, 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 4.5, we provide closing remarks and directions for future research. We include a technical appendix in Section 4.6.

4.2   Single-stage formulation

In this section we review both the classical and distributionally robust single-stage formulation, including some relevant results of [168] and [142].

4.2.1   Classical formulation

Consider the following classical formulation of the news vendor problem:

$$\inf_{x \geq 0} \mathbb{E}[\Psi(x, D)], \tag{32}$$

where

$$\Psi(x, d) := cx + b[d - x]_+ + h[x - d]_+, \tag{33}$$

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 this problem has the closed form solution $\bar{x} = F^{-1}\left(\frac{b-c}{b+h}\right)$, where $F(\cdot)$ is the cumulative distribution function (cdf) of the demand D, and F^{-1} is its inverse. Of course, it is assumed here that the probability distribution, i.e. the cdf F, is completely specified.

4.2.2   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 instead

consider the following distributionally robust formulation:

$$\inf_{x \geq 0} \psi(x), \tag{34}$$

where

$$\psi(x) := \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q[\Psi(x, D)], \tag{35}$$

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 $\mathrm{supp}(Q)$ 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 $\mathrm{supp}(Z)$ denote the support of the associated probability measure. For a given closed (and possibly unbounded) subset $\mathcal{I} \subseteq \mathbb{R}$, we let $\mathfrak{P}(\mathcal{I})$ denote the set of probability distributions Q such that $\mathrm{supp}(Q) \subseteq \mathcal{I}$. Although we will be primarily interested in the setting that $\mathcal{I} \subseteq \mathbb{R}_+$ (i.e. demand is nonnegative), it will sometimes be convenient for us to consider more general families of demand distributions. By δ_a we denote the probability measure of mass one at $a \in \mathbb{R}$.

In this chapter, we will study families of distributions satisfying moment constraints of the form

$$\mathfrak{M} := \left\{ Q \in \mathfrak{P}(\mathcal{I}) : \mathbb{E}_Q[D] = \mu, \mathbb{E}_Q[D^2] = \mu^2 + \sigma^2 \right\}. \tag{36}$$

Unless stated otherwise, it will be assumed that $\mathfrak{M}$ is indeed of the form (36), and is nonempty. We let α denote the left-endpoint of $\mathcal{I}$ (or $-\infty$ if $\mathcal{I}$ is unbounded from below), and let β denote the right-endpoint of $\mathcal{I}$ (or $+\infty$ if $\mathcal{I}$ is unbounded from above); i.e., $\mathcal{I} = [\alpha, \beta]$. It may be easily verified that the set $\mathfrak{M}$ is nonempty iff the following conditions hold:

$$\mu \in [\alpha, \beta] \text{ and } \sigma^2 \leq (\beta - \mu)(\mu - \alpha), \tag{37}$$

which will be assumed throughout. (We assume here that $0 \times \infty = 0$, so that if, e.g., $\mu = \alpha$ and $\beta = +\infty$, then the right hand side of (37) is 0.)

Furthermore, one can also identify conditions under which $\mathfrak{M}$ is a singleton.

Observation 3 *If* $-\infty < \alpha < \beta < +\infty$, $\mu \in [\alpha, \beta]$, *and* $\sigma^2 = (\beta - \mu)(\mu - \alpha)$, *then* $\mathfrak{M}$ *consists of the single probability measure which assigns to the point α probability* $p = \frac{\beta - \mu}{\beta - \alpha}$, *and to the point β probability* $1 - p = \frac{\mu - \alpha}{\beta - \alpha}$.

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:

$$
\begin{aligned}
\sup_{Q \in \mathfrak{P}(\mathcal{I})} \quad & \int \zeta(\tau) dQ(\tau) \\
\text{s.t.} \quad & \int \tau dQ(\tau) = \mu, \ \int \tau^2 dQ(\tau) = \mu^2 + \sigma^2.
\end{aligned}
\tag{38}
$$

Our definitions imply that for all $x \in \mathbb{R}$, $\psi(x)$ equals the optimal value of problem (38) for the special case that $\zeta(\tau) = \Psi(x, \tau)$. Problem (38) is a classical problem of moments (cf. [120]). By the Richter-Rogosinski Theorem (e.g., [182, Proposition 6.40]) we have the following.

Observation 4 *If problem* (38) *possesses an optimal solution, then it has an optimal solution with support of at most three points.*

We note that the distributionally robust single-stage news vendor problem considered by [168] is exactly problem (34), when $\mathcal{I} = \mathbb{R}_+$. As it will be useful for later proofs, we briefly review Scarf's explicit solution. We actually state a slight generalization of the results of Scarf, and for completeness we include a proof in the appendix (Section 4.6).

71

Theorem 4 *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 $x \in \mathbb{R}$,*

$$\psi(x) = \begin{cases} c\mu + \frac{b+h}{2}\left((x-\mu)^2 + \sigma^2\right)^{\frac{1}{2}} - \frac{b-h-2c}{2}(x-\mu), & \text{if } x \geq \frac{\mu^2+\sigma^2}{2\mu}, \\[2ex] \frac{(h+c)\sigma^2 - (b-c)\mu^2}{\mu^2+\sigma^2}x + b\mu, & \text{if } x \in [0, \frac{\mu^2+\sigma^2}{2\mu}), \\[2ex] b\mu - (b-c)x, & \text{otherwise.} \end{cases} \quad (39)$$

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 $c\mu + \left((h+c)(b-c)\right)^{\frac{1}{2}}\sigma$.*

(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.*

Furthermore, in all cases $\arg\max_{Q \in \mathfrak{M}} \mathbb{E}_Q[\Psi(x, D)]$ is nonempty for every $x \in \mathbb{R}$. Also, the optimal solution set and value of the problem $\inf_{x \in \mathbb{R}} \psi(x)$ is identical to that of problem (34), i.e. optimizing over $x \in \mathbb{R}$, as opposed to $x \in \mathbb{R}_+$, makes no difference.

For use in later proofs, it will also be useful to demonstrate a particular variant of Theorem 4. Suppose that in problem (34), we are not forced to select a deterministic constant x, but can instead select any distribution D_1 for x. Specifically, let us consider the following minimax problem:

$$\inf_{Q_1 \in \mathfrak{P}(\mathcal{I})} \phi(Q_1), \quad (40)$$

where

$$\phi(Q_1) := \sup_{Q_2 \in \mathfrak{M}} \mathbb{E}_{Q_1 \times Q_2}\left[\Psi(D_1, D_2)\right],$$

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 we defer to the appendix (Section 4.6).

Proposition 2 *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* (40) *has the unique optimal solution* $\bar{Q}_1 = \delta_0$.

We also note that ψ inherits the property of convexity from Ψ.

Observation 5 $\Psi(\cdot, d)$ *is a convex function for every fixed $d \in \mathcal{I}$, ψ is a convex function on $\mathbb{R}$, and problem* (34) *is a convex program.*

As several of our later proofs will be based on duality theory, we now briefly review duality for problem (38).

4.2.3 Duality for Problem (38)

The dual of problem (38) can be constructed as follows (cf. [102]). Consider the Lagrangian

$$L(Q, \lambda) := \int \left[\zeta(\tau) - \sum_{i=0}^{2} \lambda_i \tau^i \right] dQ(\tau) + \lambda_0 + \lambda_1 \mu + \lambda_2 (\mu^2 + \sigma^2).$$

By maximizing $L(Q, \lambda)$ with respect to $Q \in \mathfrak{P}(\mathcal{I})$, and then minimizing with respect to λ, we obtain the following Lagrangian dual for problem (38):

$$
\begin{aligned}
\inf_{\lambda \in \mathbb{R}^3} \quad & \lambda_0 + \lambda_1 \mu + \lambda_2 (\mu^2 + \sigma^2) \\
\text{s.t.} \quad & \lambda_0 + \lambda_1 \tau + \lambda_2 \tau^2 \geq \zeta(\tau), \quad \tau \in \mathcal{I}.
\end{aligned}
\tag{41}
$$

We denote by $\mathrm{val}(P)$ and $\mathrm{val}(D)$ the respective optimal values of the primal problem (38) and its dual problem (41). By convention, if problem (38) is infeasible, we set $\mathrm{val}(P) = -\infty$, and if problem (41) is infeasible, we set $\mathrm{val}(D) = +\infty$. We denote by $\mathrm{Sol}_P(x)$ the set of optimal solutions of the primal problem, and by $\mathrm{Sol}_D(x)$ the set of optimal solutions of the dual problem, and note that these sets may be empty, even when both programs are feasible, e.g. if the respective optimal values are approached but not attained.

Note that $\mathrm{val}(D) \geq \mathrm{val}(P)$. We now give sufficient conditions for there to be no duality gap, i.e. $\mathrm{val}(P) = \mathrm{val}(D)$, as well as conditions for problems (38) and (41) to have optimal solutions. By specifying known general results for duality of such programs, e.g., [28, Theorem 5.97], to the considered setting, we have the following.

Proposition 3 *If $\bar{Q}$ is a probability measure which is feasible for the primal problem (38), $\bar{\lambda} = (\bar{\lambda}_0, \bar{\lambda}_1, \bar{\lambda}_2)$ is a vector which is feasible for the dual problem (41), and*

$$\mathrm{supp}(\bar{Q}) \subseteq \left\{ \tau \in \mathcal{I} : \zeta(\tau) = \bar{\lambda}_0 + \bar{\lambda}_1 \tau + \bar{\lambda}_2 \tau^2 \right\}, \tag{42}$$

then $\bar{Q}$ is an optimal primal solution, $\bar{\lambda}$ is an optimal dual solution, and $\mathrm{val}(P) = \mathrm{val}(D)$. Conversely, if $\mathrm{val}(P) = \mathrm{val}(D)$, and $\bar{Q}$ and $\bar{\lambda}$ are optimal solutions of the respective primal and dual problems, then condition (42) holds.

4.2.4 Explicit solution of Problem (38) for a class of convex, continuous, piecewise affine functions

[168] gave an explicit solution for problems (38) and (41) when $\mathcal{I} = \mathbb{R}_+$, and ζ is a convex, continuous piecewise affine function with exactly two pieces, by explicitly constructing a feasible primal - dual solution pair satisfying the conditions of Proposition 3 (details of this construction can be found in Section 4.6). [142] generalized Scarf's results to a class of convex, continuous, piecewise affine (CCPA) functions with three pieces. We now state the solution to a special case of the problems studied in [142], as we will need the solution to such problems for our later studies of time consistency. For completeness, we provide a proof in the appendix (Section 4.6).

Theorem 5 *[142] 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 $\mathfrak{f}(z) := \left((z - \mu)^2 + \sigma^2\right)^{\frac{1}{2}}$ for all $z \in \mathbb{R}$. Further suppose that $\sigma > 0$, $\mathcal{I} = \mathbb{R}_+$,*

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

and $\eta - \mathfrak{f}(\eta) \geq 0$. *Then the unique optimal solution to the primal problem* (38) *is the probability measure Q having support at two points $h_1 = \eta - \mathfrak{f}(\eta)$ and $h_2 = \eta + \mathfrak{f}(\eta)$, with*

$$Q(h_1) = \sigma^2 \left(\sigma^2 + \left(\eta - \mathfrak{f}(\eta) - \mu \right)^2 \right)^{-1}, \quad Q(h_2) = 1 - Q(h_1). \tag{43}$$

Also, the unique optimal solution to the dual problem (41) *is*

$$\lambda_0 = \tfrac{1}{2}\left(\eta^2 + (\eta - \mu)^2 + \sigma^2 \right)\mathfrak{f}^{-1}(\eta) + \frac{c_1 - c_2}{2}, \quad \lambda_1 = -\eta \mathfrak{f}^{-1}(\eta), \quad \lambda_2 = \tfrac{1}{2}\mathfrak{f}^{-1}(\eta). \tag{44}$$

4.3 Multistage formulation

In this section, we study a multistage extension of the distributionally robust news vendor problem discussed in Section 4.2.2.

4.3.1 Classical formulation

We begin by giving a quick review of the classical (i.e. non-robust) multistage news vendor problem (called inventory problem), and start by introducing some additional notations. 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 $D = (D_1, \ldots, D_T)$. By $d = (d_1, \ldots, d_T)$ 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, $\{y_t, \ t = 1, \ldots, T\}$ and $\{x_t, \ t = 1, \ldots, T\}$. 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 $\pi = \{x_t(d_{[t-1]}), t = 1, \ldots, T\}$, 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 $t = 1, \ldots, T$, where

$$y_{t+1} = x_t(d_{[t-1]}) - d_t, \ t = 1, \ldots, T-1. \tag{45}$$

It follows that any given choice of $\pi \in \Pi$, along with the given y_1, completely determines the associated functions $y_1, \ldots, y_T$. 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 $y_t^\pi(d_{[t-1]})$; other times we will suppress this notation. Combining the above, we can write the classical multistage news vendor problem (inventory problem) as follows:

$$\inf_{\pi \in \Pi} \mathbb{E}\left\{\sum_{t=1}^T \rho^{t-1}\big[c_t\big(x_t^\pi(D_{[t-1]}) - y_t^\pi(D_{[t-1]})\big) + \Psi_t\big(x_t^\pi(D_{[t-1]}), D_t\big)\big]\right\}. \tag{46}$$

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

$$\Psi_t(x_t, d_t) := b_t[d_t - x_t]_+ + h_t[x_t - d_t]_+. \tag{47}$$

Unless stated otherwise, we assume that $b_t > c_t > 0$ and $h_t \geq 0$ for all $t = 1, \ldots, T$.

Problem (46) 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 (46) can be solved using dynamic programming equations, which can be written as

$$V_t(y_t) = \inf_{x_t \geq y_t} \left\{c_t(x_t - y_t) + \mathbb{E}\big[\Psi_t(x_t, D_t) + \rho V_{t+1}(x_t - D_t)\big]\right\}, \tag{48}$$

$t = 1, \ldots, T$, with $V_{T+1}(\cdot) \equiv 0$ (e.g., [213]). 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

$$\inf_{x_t \geq y_t} \left\{ c_t(x_t - y_t) + \mathbb{E}[\Psi_t(x_t, D_t) + \rho V_{t+1}(x_t - D_t)] \right\}, \tag{49}$$

and the optimal value of problem (46) is given by $V_1(y_1)$. Note that $x_t(y_t)$, $t = 1, \ldots, T$, 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., [17] for technical details). Of course, the assumption of stagewise independence is essential for this conclusion.

Under the specified conditions, the objective function of problem (49) 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 (49). In particular, the so-called *base-stock policy* is optimal for the inventory problem (46), where we note that such a result is classical in the inventory literature.

Definition 1 *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*

$$x_t^\pi = \max\left\{y_t^\pi, x_t^*\right\}, \ t = 1, \ldots, T, \tag{50}$$

That is, problem (46) can be solved using the dynamic programming formulation (48) and associated policy (49) in the following sense.

Lemma 11 *The optimal value of problem (46) 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 (46). Conversely, for any optimal policy π for problem (46), and any $t \in \{1, ..., T\}$, there exists a set $A \in \mathbb{R}$ such that $\Pr\big(y_t^\pi(D_{[t-1]}) \in A\big) = 1$, and $x_t^\pi(D_{[t-1]}) \in \mathfrak{X}_t\big(y_t^\pi(D_{[t-1]})\big)$ conditional on the event $\{y_t^\pi(D_{[t-1]}) \in A\}$.*

As we shall see, such an equivalence does not necessarily hold for distributionally robust multistage inventory problems with moment constraints.

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 moments. In this case, it is natural to use the framework previously developed for the single-stage problem (see Section 4.2) 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 intuitively corresponding to the modeling choices of a policy maker who does not recompute his/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 [103] and [145] in the framework of robust MDP, and can also be interpreted through the lens of (non)rectangularity of the associated families of priors (cf. [54, 103, 145]), as we will explore later in this section. We refer to these formulations as *multistage-static* and *multistage-dynamic*, 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 such considerations were one of the original motivations for the study of time consistency in economics (cf. [187]). We further 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 [57], and other recent papers in the optimization community (cf. [103, 27, 32, 97]) for alternative perspectives.

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

We first consider the following formulation, referred to as *multistage-static*, in which the policy maker does not recompute his/her policy choices after each stage. Let us define

$$\mathfrak{M}_t \;:=\; \left\{ Q_t \in \mathfrak{P}(\mathcal{I}_t) : \mathbb{E}_{Q_t}[D_t] = \mu_t, \; \mathbb{E}_{Q_t}[D_t^2] = \mu_t^2 + \sigma_t^2 \right\}, \; t = 1, ..., T; \quad (51)$$

$$\mathfrak{M} \;:=\; \{Q = Q_1 \times \cdots \times Q_T : Q_t \in \mathfrak{M}_t, \; t = 1, ..., T\}. \quad (52)$$

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 4.3.1, to the considered distributionally robust case. In order for the sets $\mathfrak{M}_t$ to be nonempty we assume that (compare with (37))

$$\mu_t \in [\alpha_t, \beta_t] \;\; \text{and} \;\; \sigma_t^2 \leq (\beta_t - \mu_t)(\mu_t - \alpha_t), \; t = 1, ..., T. \quad (53)$$

According to (52), 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.

$$\inf_{\pi \in \Pi} \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q\big[Z^\pi\big], \quad (54)$$

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

$$Z^\pi(D_{[T]}) := \sum_{t=1}^{T} \rho^{t-1}\big[c_t\big(x_t^\pi(D_{[t-1]}) - y_t^\pi(D_{[t-1]})\big) + \Psi_t\big(x_t^\pi(D_{[t-1]}), D_t\big)\big], \quad (55)$$

and Π is the set of admissible policies defined previously in Section 4.3.1. Of course, if the set $\mathfrak{M} = \{Q\}$ is a singleton, then formulation (54) coincides with formulation (46) taken with respect to the distribution $Q = Q_1 \times \cdots \times Q_T$ of the demand vector $D_{[T]}$.

Very little is known about the set of optimal policies for problem (54), as this problem does not enjoy a dynamic programming formulation along the lines of (48).

4.3.2.2 Multistage-dynamic formulation

We now consider the formulation which we refer to as *multistage-dynamic*. In this formulation the policy maker recomputes his/her policy choices after each stage. To build intuition, let us think of what it means for the policy maker to recompute his/her optimal policy at the start of the final stage T. As he/she cannot change past decisions, the only policy decision he/she still has to make is the determination of the function x_T. However, he/she now has knowledge of $D_{[T-1]}$ and y_T, which he/she can incorporate into his/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 $D_{[T-1]}$.

We can approach this from the following point of view. Consider the cost $Z^\pi = Z^\pi(D_{[T]})$ of a policy π, defined in (55). Let $\mathfrak{M}$ be a set of probability distributions of the demand vector $D_{[T]} = (D_1, ..., D_T)$, and let $Q \in \mathfrak{M}$. At the moment we do not assume that Q is of the product form $Q = Q_1 \times \cdots \times Q_T$, we will discuss this later. We can write

$$\mathbb{E}_Q[Z^\pi] = \mathbb{E}_Q\left[\mathbb{E}_{Q|D_1}\left[\cdots \mathbb{E}_{Q|D_{[T-2]}}\left[\mathbb{E}_{Q|D_{[T-1]}}[Z^\pi]\right]\right]\right], \tag{56}$$

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,

$$\sup_{Q\in\mathfrak{M}}\mathbb{E}_Q[Z^\pi] \le \sup_{Q\in\mathfrak{M}}\mathbb{E}_Q\left[\sup_{Q\in\mathfrak{M}}\mathbb{E}_{Q|D_1}\left[\cdots\sup_{Q\in\mathfrak{M}}\mathbb{E}_{Q|D_{[T-1]}}[Z^\pi]\right]\right]. \tag{57}$$

The right hand side of (57) leads to the nested formulation

$$\inf_{\pi\in\Pi}\left\{\sup_{Q\in\mathfrak{M}}\mathbb{E}_Q\left[\sup_{Q\in\mathfrak{M}}\mathbb{E}_{Q|D_1}\left[\cdots\sup_{Q\in\mathfrak{M}}\mathbb{E}_{Q|D_{[T-1]}}[Z^\pi]\right]\right]\right\}. \tag{58}$$

We refer to (58) as the *multistage-dynamic formulation*. It follows from (57) that the optimal value of the multistage-dynamic problem (58) is greater than or equal to the optimal value of the multistage-static problem (54). In particular, if the set $\mathfrak{M}$ is defined in the form (52), i.e., consists of products of probability measures, then the multistage-dynamic formulation (58) simplifies to

$$\inf_{\pi\in\Pi}\left\{\sup_{Q_1\in\mathfrak{M}_1}\mathbb{E}_{Q_1}\left[\sup_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2|D_1}\left[\cdots\sup_{Q_T\in\mathfrak{M}_T}\mathbb{E}_{Q_T|D_{[T-1]}}[Z^\pi]\right]\right]\right\}. \tag{59}$$

For the multistage-dynamic formulation it is possible to write dynamic programming equations (cf. [179]). In particular, for the set $\mathfrak{M}$ of the form (52) and the corresponding multistage-dynamic problem (59) the dynamic programming equations become (compare with (48)):

$$V_t(y_t) = \inf_{x_t\ge y_t}\left\{c_t(x_t - y_t) + \sup_{Q_t\in\mathfrak{M}_t}\mathbb{E}_{Q_t}[\Psi_t(x_t, D_t) + \rho V_{t+1}(x_t - D_t)]\right\}, \tag{60}$$

$t = 1, ..., T$, with $V_{T+1}(\cdot) \equiv 0$. The optimal value of problem (59) is given by $V_1(y_1)$. These dynamic equations 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

$$\mathfrak{Y}_t(y_t) := \arg\min_{x_t\ge y_t}\left\{c_t(x_t - y_t) + \sup_{Q_t\in\mathfrak{M}_t}\mathbb{E}_{Q_t}[\Psi_t(x_t, D_t) + \rho V_{t+1}(x_t - D_t)]\right\}, \ t = 1, ..., T. \tag{61}$$

Let us observe that the nested max-expectation operator in the right hand side of (57) can be represented as a maximum with respect to a certain set of distributions. That is, there exists a set $\widehat{\mathfrak{M}}$ of probability distributions of $D_{[T]}$ such that

$$\sup_{Q\in\widehat{\mathfrak{M}}}\mathbb{E}_Q[\cdot] = \sup_{Q\in\mathfrak{M}}\mathbb{E}_Q\left[\sup_{Q\in\mathfrak{M}}\mathbb{E}_{Q|D_1}\left[\cdots\sup_{Q\in\mathfrak{M}}\mathbb{E}_{Q|D_{[T-1]}}[\cdot]\right]\right]. \tag{62}$$

81

Proof of existence and a construction of such set $\widehat{\mathfrak{M}}$ is similar to the corresponding derivations of conditional risk mappings (cf. [166, section 5]). We refer to [181] for technical details.

If, moreover, $\mathfrak{M}$ is of the product form (52), then (62) simplifies to

$$\sup_{Q \in \widehat{\mathfrak{M}}} \mathbb{E}_Q[\,\cdot\,] = \sup_{Q_1 \in \mathfrak{M}_1} \mathbb{E}_{Q_1} \left[\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2|D_1} \left[\cdots \sup_{Q_T \in \mathfrak{M}_T} \mathbb{E}_{Q_T|D_{[T-1]}}[\,\cdot\,] \right] \right] . \tag{63}$$

We note that the set $\widehat{\mathfrak{M}}$ is not defined uniquely, and that the largest such set will be convex and closed in an appropriate topology. Furthermore, it is always possible to choose $\widehat{\mathfrak{M}}$ in such a way that $\mathfrak{M} \subset \widehat{\mathfrak{M}}$.

Definition 2 *We refer to a set $\widehat{\mathfrak{M}}$, satisfying equation (62) as a* rectangular *set associated with the set $\mathfrak{M}$ of probability measures. In particular, if the set $\mathfrak{M}$ is of the product form (52), then we say that $\widehat{\mathfrak{M}}$ is a rectangular set associated with sets $\mathfrak{M}_t$, $t = 1, ..., T$, if equation (63) holds. Furthermore, we say that a set of measures $\mathfrak{M}$ is* rectangular *if the set $\mathfrak{M}$ itself is a rectangular set associated with the set $\mathfrak{M}$ of probability measures.*

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

$$\inf_{\pi \in \Pi} \sup_{Q \in \widehat{\mathfrak{M}}} \mathbb{E}_Q \left[Z^\pi \right], \tag{64}$$

is equivalent to the corresponding dynamic formulation (59), and the dynamic programming equations (60) can be applied to (64).

We note that the concept of rectangularity has been central to the past literature on time consistency (cf. [54, 79, 97]), especially as it relates to optimization (cf. [103, 145, 200]). 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 [200] 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 [181] for details).

We note that for $\mathfrak{M}$ defined in (51)-(52), we can select one such $\widehat{\mathfrak{M}}$ to be the set of all joint distributions Q for $D_{[T]}$ s.t.

$$D_t \in \mathfrak{P}(\mathcal{I}_t) \ , \quad \mathbb{E}_Q[D_t|D_{[t-1]}] = \mu_t \ , \quad \mathbb{E}_Q[D_t^2|D_{[t-1]}] = \mu_t^2 + \sigma_t^2, \quad t = 1, ..., T. \quad (65)$$

As already mentioned, there may be more than one way to construct such a rectangular set. Furthermore, the question of existence of "minimal" rectangular sets seems to be a delicate issue, beyond the scope of this study, where we note that related questions (under closely related but different definitions and assumptions) have been considered previously in the literature (cf. [97]).

4.3.2.3 Dynamic programming solution to the multistage-dynamic formulation

Consider the dynamic programming equations (60)–(61) with associated optimal value given by $V_1(y_1)$. Again note that $\{x_t(y_t), t = 1, \ldots, T\}$ are (measurable) functions of y_t, i.e., it suffices to consider policies of the form $x_t = \pi_t(y_t)$, $t = 1, \ldots, T$. Then Problem (58) can be solved using the dynamic programming formulation (60) and associated policy (61) in the following sense.

Lemma 12 *The optimal value of Problem (59) 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 (59). Conversely, for any optimal policy π for Problem (59), and any associated rectangular set $\widehat{\mathfrak{M}}$ and 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 note that the same conclusion could also have been drawn by rephrasing our formulation in the language of coherent risk measures, and applying known results for so-called nested risk measures (cf. [166], [182, section 6.7.3]), although we do not pursue such an analysis here.

We now observe that due to certain convexity properties, the set of policies indicated in Lemma 12 has a particularly simple form. We note that such results are generally well-known to hold in this setting (cf. [2]). Recall Definition 1 of a base-stock policy. Let us make the following observation.

Observation 6 *It follows from the convexity of the relevant cost-to-go functions* $V_t(y_t)$ *that both problems* (46) *and* (59) *possess optimal base-stock policies. Furthermore, 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* (46)*, while any such set of base-stock constants such that* $x_t^* \in \mathfrak{Y}_t(0)$ *for all* $t \in [1, T]$ *will yield an optimal policy for problem* (59)*.*

We note that the question of whether or not there exists such an optimal base-stock policy for the multistage-static formulation is considerably more challenging, and will be central to our discussion of time consistency.

Non-rectangular (and intractable) formulations for robust MDP are described in both [103] and [145]. In [103], it is referred to as the static formulation, while in [145], 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 [103], [145], and [200] 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.

4.4 Time consistency

As discussed in Section 4.3, there is no apriori guarantee that the multistage-static formulation is equivalent to the corresponding multistage-dynamic formulation in the distributionally robust setting. Disagreement between these two formulations is undesirable from a policy perspective, as it suggests that a policy which was optimal when performing one's minimax computations before seeing any realized demand may no longer be optimal if one reperforms these computations at a later time. This general problem goes under the heading of time (in)consistency. 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 (cf. [159, 27, 6, 177, 165, 79, 32, 180, 40, 97]).

We note that related issues were addressed even in the seminal work of [11] on dynamic programming, where it is asserted that: *"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."* The same principle has been subsequently reformulated by several authors in a somewhat more precise form, e.g., in the recent work of [32], where it is asserted that *"The decision maker formulates an optimization problem at time t_0 that yields a sequence of optimal decision rules for t_0 and for the following time steps $t_1, ..., t_N = T$. Then, at the next time step t_1, he formulates a new problem starting at t_1 that yields a new sequence of optimal decision rules from time steps t_1 to T. Suppose the process continues until time T is reached. The sequence of optimization problems is said to be dynamically consistent if the optimal strategies obtained when solving the original problem at time t_0 remain optimal for all subsequent problems."*

From a conceptual point of view this is quite natural - an optimal solution obtained by solving the problem at the first stage should remain optimal from the point of view of later stages. The setting in which one re-optimizes at each stage coincides precisely

85

with our multistage-dynamic formulation, while the "problem at time t_0" coincides naturally with our multistage-static formulation. 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 at time t_0, 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) at time t_0, but deemed sub-optimal (i.e. non-implementable) at a later time. This motivates the following definition(s) of time consistency, where we note that similar definitions were presented in [79] in a different context motivated by considerations in decision theory and artificial intelligence.

Definition 3 (Time consistency) *If a policy $\pi \in \Pi$ is optimal for both the multistage-static problem* (54) *and the multistage-dynamic problem* (59), *we say that π is* time consistent. *If there exists at least one optimal policy $\pi \in \Pi$ which is time consistent, we say that problem* (54) *is* weakly time consistent. *If every optimal policy of problem* (54) *is time consistent, we say that problem* (54) *is* strongly time consistent.

Of course the notion of strong time consistency makes sense only if problem (54) possesses at least one optimal solution. Otherwise it is strongly time consistent simply because the set of optimal policies is empty.

We note that our definition 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 4.4.3.3, 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 multistage-dynamic 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. The definition of consistency typically used in the theory of risk measures, i.e. the notion of

dynamic consistency coming from [54] and based on a certain stability of preferences over time, 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. [165, 79]). 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.

Before exploring some of the subtle and interesting features of time (in)consistency for our model, we briefly review some related previously known results for simpler models. Note that if the set $\mathfrak{M}$ is a singleton, then both the multistage-static and multistage-dynamic formulations collapse to the classical formulation, and 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 multistage-dynamic formulations collapse to the so-called adjustable robust formulation (cf. [15], [178]), which is purely deterministic, from which strong time consistency again follows. If one only has information about the support $\mathcal{I}_t$ and first moment $\mu_t \in \mathcal{I}_t$ of demand at each stage, and $\mathcal{I}_t = [\alpha_t, \beta_t]$ is bounded for all t, then it follows from the results of [179, section 4.2.2] that the corresponding problem is again strongly time consistent, as convexity dictates that in every time period, the adversary's choice of demand distribution is independent of all previously realized demands. As we will see, the question of time consistency becomes considerably more interesting in our setting, when one is also given second moment information.

4.4.1 Sufficient conditions for weak time consistency

In this section, we provide simple sufficient conditions for the weak time consistency of Problem (54). Our condition is essentially equal 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 multistage-dynamic 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. [192, 100, 213]).

We begin by providing a different (but equivalent) formulation for Problem (54), 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

$$\hat{\Psi}_t(x_t, d_t) := (c_t - \rho c_{t+1})x_t + b_t[d_t - x_t]_+ + h_t[x_t - d_t]_+, \ t = 1, ..., T. \quad (66)$$

Let us define the problem

$$\inf_{\pi \in \Pi} \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q \left[\sum_{t=1}^T \rho^{t-1} \hat{\Psi}_t \big(x_t(y_t), D_t \big) \right] - c_1 y_1 + \sum_{t=1}^{T-1} \rho^t c_{t+1} \mu_t. \quad (67)$$

Then it follows from a straightforward substitution and calculation that

Observation 7 *Problem* (54) *and Problem* (67) *are equivalent, i.e. each policy $\pi \in \Pi$ has the same value under both formulations.*

We now derive a lower bound for any policy, which intuitively comes from allowing the policy maker to reselect her inventory at the start of each stage, at no cost. As it turns out, this bound is "realizable" when the set of base-stock levels is monotone

increasing. For $x \in \mathbb{R}$, let us define

$$\eta_t(x) := \sup_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(x, D_t)], \quad \Gamma_t^x := \arg \max_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(x, D_t)], \tag{68}$$

and let

$$\hat{\eta}_t := \inf_{x \in \mathbb{R}} \eta_t(x) = \inf_{x \in \mathbb{R}} \sup_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(x, D_t)],$$

$$\hat{\Gamma}_t := \arg \min_{x \in \mathbb{R}} \eta_t(x) = \arg \min_{x \in \mathbb{R}} \sup_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(x, D_t)]. \tag{69}$$

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.

Lemma 13 *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*

$$\mathbb{E}_{\otimes_{t=1}^{T} Q_t}\left[\hat{\Psi}_t\big(x_t(y_t), D_t\big)\right] \geq \hat{\eta}_t \text{ for all } t \geq i + 1. \tag{70}$$

Furthermore, the optimal value of Problem (54) is at least $\sum_{t=1}^{T} \rho^{t-1}\hat{\eta}_t - c_1 y_1 + \sum_{t=1}^{T-1} \rho^t c_{t+1}\mu_t$.

Proof. Suppose $i \in \{0, ..., T\}$ and $Q_1, \ldots, Q_i$ are fixed. We now prove that (70) holds for all $t \geq i + 1$, and proceed by induction. Our particular induction hypothesis will be that there exist $Q_{i+1}, \ldots, Q_{i+n}$ such that

$$\mathbb{E}_{\otimes_{t=1}^{i+n} Q_t}\left[\hat{\Psi}_t\big(x_t(y_t), D_t\big)\right] \geq \hat{\eta}_t \text{ for all } t \in [i+1, i+n]. \tag{71}$$

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 $Q_{i+1} \in \mathfrak{M}_{i+1}$,

$$\mathbb{E}_{\otimes_{t=1}^{i+1} Q_t}\left[\hat{\Psi}_{i+1}\big(x_{i+1}(y_{i+1}), D_{i+1}\big)\right] \geq \mathbb{E}_{Q_{i+1}}\left[\hat{\Psi}_{i+1}\big(\mathbb{E}_{\otimes_{t=1}^{i} Q_t}[x_{i+1}(y_{i+1})], D_{i+1}\big)\right].$$

Taking Q_{i+1} to be any element of $\Gamma_{i+1}^{\mathbb{E}_{\otimes_{t=1}^{i} Q_t}[x_{i+1}(y_{i+1})]}$ ($\Gamma_1^{x_1(y_1)}$ 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 $Q_{i+n+1} \in \mathfrak{M}_{i+n+1}$,

$$\mathbb{E}_{\otimes_{t=1}^{i+n+1} Q_t}\left[\hat{\Psi}_{i+n+1}\big(x_{i+n+1}(y_{i+n+1}), D_{i+n+1}\big)\right]$$

$$\geq \ \mathbb{E}_{Q_{i+n+1}}\left[\hat{\Psi}_{i+n+1}\big(\mathbb{E}_{\otimes_{t=1}^{i+n} Q_t}[x_{i+n+1}(y_{i+n+1})], D_{i+n+1}\big)\right].$$

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

We now show that the bound of Lemma 13 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 multistage-dynamic formulations. In particular, in this setting, the associated base-stock policy is time consistent, and thus the multistage-static problem is weakly time consistent.

Theorem 6 *Suppose there exists nondecreasing sequence x_t^*, $t = 1, ..., T$, such that $y_1 \leq x_1^*$, and $x_t^* \in \hat{\Gamma}_t$, $t = 1, ..., T$, where $\hat{\Gamma}_t$ is defined in (69). Also suppose $\mathcal{I}_t \subset \mathbb{R}_+$ for all $t = 1, ..., T$. Then the base-stock policy π for which $x_t(y_t) = \max\{y_t, x_t^*\}$ for all $y_t \in \mathbb{R}$, is an optimal policy for both the multistage-static and multistage-dynamic formulations, and attains value $\sum_{t=1}^{T} \rho^{t-1}\hat{\eta}_t - c_1 y_1 + \sum_{t=1}^{T-1} \rho^t c_{t+1}\mu_t$. Consequently, this base-stock policy is time consistent, and the multistage-static problem is weakly time consistent.*

Proof. Note that under these assumptions, if policy π is implemented under the multistage-dynamic formulation, then w.p.1 $x_t(y_t) = x_t^*$ for all $t = 1, ..., T$. It then follows from a straightforward induction that π is an optimal policy for the multistage-dynamic formulation, and w.p.1, for all $t = 2, ..., T$,

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

and

$$V_1(y_1) = \sum_{t=1}^{T} \rho^{t-1}\hat{\eta}_t - c_1 y_1 + \sum_{t=1}^{T-1} \rho^t c_{t+1}\mu_t.$$

90

Combining with Lemma 13 and Observation 6 completes the proof. ∎

We note that Theorem 6 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 weakly time consistent, and the multistage-static and multistage-dynamic formulations have the same optimal value.

4.4.2 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". Although such a requirement on the family of optimal policies seems quite stringent, we will later see in Section 4.4.3.2 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 7 *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*

$$\frac{\sigma_t^2}{\mu_t^2} > \frac{b_t'}{h_t'}, \quad t = 1, ..., T. \tag{72}$$

Then the set of optimal policies for the multistage-static problem is exactly the set of policies

$$\Pi^0 := \big\{ \pi = (x_1, \ldots, x_T) \in \Pi : x_1(y_1) = 0, x_t(z) = 0 \text{ for all } z \leq 0 \text{ and } t \in [1, T] \big\},$$

and the multistage-static problem is strongly time consistent.

Proof. Let Π^{opt} denote the set of optimal policies for the multistage-static problem. It follows from Theorem 4.(i) and Theorem 6 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 $\bar{\pi} = (\bar{x}_1, \ldots, \bar{x}_T) \in \Pi^{opt}$ implies $\bar{x}_1(y_1) = 0$. Indeed, it follows from Lemma 13 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 4.

We now show that $\bar{\pi} \in \Pi^{opt}$ implies $\bar{x}_2(z) = 0$ for all $z \leq 0$. Suppose for contradiction that there exists $z' \leq 0$ such that $\bar{x}_2(z') \neq 0$. 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 2 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 13 then yields a contradiction. The proof that $\bar{x}_t(z) = 0$ for all $z \leq 0$ and $t \geq 3$ follows from a nearly identical argument, and we omit the details. ∎

4.4.3 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. 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 dynamic formulations; and hence the rectangularity property does not hold. We also prove that, although the multistage-dynamic formulation

always has an optimal policy of the base-stock form, there may be no such optimal policy for the multistage-static formulation. We note that (i) and (ii) are subtle phenomena which the simpler models discussed in several previous works (e.g. [180]) cannot exhibit. We also note that (iii) emphasizes an interesting and surprising feature of our model and definitions: (strong) time consistency can hold even when the underlying family of measures from which nature can select is non-rectangular. This stands in contrast to much of the related work on time consistency, where rectangularity is essentially taken as a pre-requisite for time consistency. We also note that (iii) stands in contrast to some alternative, less policy-focused definitions of time consistency, e.g. those definitions appearing in the literature on risk measures (cf. [54]), under which time consistency could not hold if an optimal policy took different values under the two formulations. 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 optimal policies for the corresponding multistage-dynamic problem.

4.4.3.1 *Example when the multistage-static problem is not weakly time consistent*

In this section, we explicitly provide an example for which the multistage-static problem is not weakly time consistent, showing that in general, the multistage-static and multistage-dynamic formulations need not have a common optimal policy. Furthermore, for this example, the multistage-static and multistage-dynamic formulations have different optimal values.

Let us define $y_1 = 10$, $\rho = 1$,

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

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

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 the set $\tilde{\Pi}_s$ specifies values of $x_2(y_2)$ only for $y_2 = 9$ and $y_2 = 7$. As we will see in Lemma 14, other values of y_2 are irrelevant for the multistage-static formulation regarding optimality.

Theorem 8 $\Pi_s^{opt} = \tilde{\Pi}_s$, 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 multistage-dynamic problem is $17 + \frac{\sqrt{5}}{2} > 18$. Consequently, the multistage-static problem is not weakly time consistent, and the multistage-static and multistage-dynamic problems have different optimal values.

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

Lemma 14 $\Pi_s^{opt} = \tilde{\Pi}_s$, and the multistage-static problem has optimal value 18.

Proof. It follows from Observation 3 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 $x_1(y_1) = x_1(10) \geq 10$. Consequently, $\Pr(x_1(y_1) \geq D_1) = 1$, 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

$$2x_1(10) - 4 + \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[\tfrac{1}{2}\big(\big|x_2\big(x_1(10) - 1\big) - D_2\big| + \big|x_2\big(x_1(10) - 3\big) - D_2\big|\big)\right]. \quad (73)$$

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 (73) and a straightforward contradiction argument that

$$\bar{x}_1(10) = 10. \quad (74)$$

Combining (73) and (74), we conclude that

$$\big(\bar{x}_2(9), \bar{x}_2(7)\big) \in \arg \min_{(x,y): x \geq 9, y \geq 7} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[\tfrac{1}{2}\big(|x - D_2| + |y - D_2|\big)\right]. \quad (75)$$

Furthermore, it follows from Lemma 13 and Theorem 4 that

$$\inf_{(x,y):x\geq 9, y\geq 7} \sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_2}\left[\tfrac{1}{2}\left(|x - D_2| + |y - D_2|\right)\right] \;\geq\; \sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|8 - D_2|\right] \;=\; 2. \quad (76)$$

Noting that

$$\tfrac{1}{2}\left(|9 - D_2| + |7 - D_2|\right) = 1 + \max(-D_2 + 7, 0, D_2 - 9),$$

it then follows from a straightforward calculation and Theorem 5 that

$$\sup_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_2}\left[\tfrac{1}{2}\left(|9 - D_2| + |7 - D_2|\right)\right] = 2. \quad (77)$$

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 (74) and (75) that $\tfrac{1}{2}\left(\hat{x}_2(9) + \hat{x}_2(7)\right) > 8$. However, it then follows from Jensen's inequality, Theorem 4, and (76) that

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

Combining with (76) and (77) yields a contradiction, completing the proof. ∎

We now characterize the set of optimal policies for the multistage-dynamic problem.

Lemma 15 $\Pi_d^{opt} \subseteq \tilde{\Pi}_d$, and the multistage-dynamic problem has optimal value $17 + \frac{\sqrt{5}}{2}$.

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

$$\bar{x}_1(10) = 10. \quad (78)$$

95

It then follows from (61) 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 4 and a straightforward calculation. ∎

Combining Lemmas 14 and 15 completes the proof of Theorem 8.

4.4.3.2 *Example when the multistage-static problem 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, the multistage-static and multistage-dynamic formulations have a common optimal base-stock policy π^*, with associated base-stock constants x_1^*, x_2^*, satisfying the conditions of Theorem 6, yet the multistage-static problem has other non-trivial optimal policies which are suboptimal for the multistage-dynamic 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. Alternatively, this policy is suboptimal in the multistage-dynamic formulation, as the adversary can first see exactly how the inventory level deviated from that dictated by π^*, and exploit this to achieve a strictly higher cost. 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 6.

Let us define $y_1 = 0$, $\rho = 1$,

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

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

Then we prove the following.

Theorem 9 *The multistage-static problem is weakly time consistent, but not strongly time consistent.*

We first prove that the multistage-static problem is weakly time consistent.

Lemma 16 *The multistage-static problem is weakly time consistent, and both the multistage-static and multistage-dynamic problems have optimal value 2.*

Proof. Note that

$$\hat{\Psi}_1(x_1, d_1) = |x_1 - d_1|, \quad \hat{\Psi}_2(x_2, d_2) = |x_2 - d_2|.$$

It follows from Observation 3 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 4 and a straightforward calculation that

$$\hat{\Gamma}_1 = [1, 3] \quad, \quad \hat{\Gamma}_2 = 10 \quad, \quad \hat{\eta}_2 = 1.$$

Combining the above with Theorem 6, 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 and multistage-dynamic problems, which have common optimal value 2. ∎

We now prove that the multistage-static problem is not strongly time consistent. In particular, consider the policy $\pi' = (x_1', x_2')$ such that

$$x_1'(y) = \max\{3, y\}, \quad \text{and} \quad x_2'(y) = \begin{cases} 9.9, & \text{if } y \leq 0, \\ \max\{10.1, y\}, & \text{otherwise.} \end{cases} \tag{79}$$

Lemma 17 *The policy $\pi' \in \Pi_s^{opt}$, but $\pi' \notin \Pi_d^{opt}$. Consequently, the multistage-static problem is not strongly time consistent.*

Proof. 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

$$\mathbb{E}_{Q_1}|3 - D_1| + 0.1 + \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \max \left\{ 9.9 - D_1, 0, D_1 - 10.1 \right\}. \tag{80}$$

It is easily verified that the conditions of Theorem 5 are met, and we may apply Theorem 5 to conclude that $\arg\max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \max \left\{ 9.9 - D_1, 0, D_1 - 10.1 \right\}$ 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 expression in (80) equals 2, and we conclude that $\pi' \in \Pi_s^{opt}$, completing the proof.

We now show that $\pi' \notin \Pi_d^{opt}$. Suppose, for contradiction, that $\pi' \in \Pi_d^{opt}$. It then follows from a straightforward calculation that

$$9.9 \in \arg\min_{x \geq 0} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}[|x - D_2|] \tag{81}$$

However, it follows from Theorem 4 that the right-hand side of (81) is the singleton $\{10\}$, completing the proof ∎

Combining Lemmas 16 and 17 completes the proof of Theorem 9.

4.4.3.3 Example when the multistage-static problem 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.

Let us define $y_1 = 0$, $\rho = 1$,

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

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

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

Theorem 10 $\Pi_s^{opt} = \tilde{\Pi}$, and the multistage-static problem is strongly time consistent. However, the optimal value of the multistage-static problem equals 5, while the optimal value of the multistage-dynamic problem equals $\sqrt{26} > 5$.

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

Lemma 18 $\Pi_s^{opt} = \tilde{\Pi}$, and the multistage-static problem has optimal value 5.

Proof. It follows from Observation 3 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

$$\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \left[\mathbb{E}_{Q_1} \left[2 \Big(x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big) \Big) + \big| x_2\big(x_1(0) - D_1\big) - D_2 \big| \right] \right]. \quad (82)$$

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

$$\bar{x}_2\big(\bar{x}_1(0) - 1\big) = \bar{x}_1(0) - 1 \quad \text{and} \quad \bar{x}_2\big(\bar{x}_1(0) - 3\big) = \bar{x}_1(0) - 3. \quad (83)$$

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

$$2 \Big(x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big) \Big) + \big| x_2\big(x_1(0) - D_1\big) - D_2 \big|$$

is at least

$$2 \Big(x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big) \Big) + \big| \big(x_1(0) - D_1\big) - D_2 \big| - \big| x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big) \big|,$$

which equals

$$x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big) + \big|x_1(0) - D_1 - D_2\big|. \tag{84}$$

Now, suppose for contradiction that (83) 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 (84), we conclude that (82) is strictly greater than

$$\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\bigg[\mathbb{E}_{Q_1}\Big[\big|x_1(0) - D_1 - D_2\big|\Big]\bigg]. \tag{85}$$

Noting that (85) is the cost incurred by some policy satisfying (83) completes the proof.

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

$$\arg\min_{x_1 \in \mathbb{R}_+} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\big[\tfrac{1}{2}\big(|x_1 - 1 - D_2| + |x_1 - 3 - D_2|\big)\big] = \{102\}. \tag{86}$$

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 5 are met. We may thus apply Theorem 5 to conclude that for all $x_1 \in [100, 104]$,

$$\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\big[\tfrac{1}{2}\big(|x_1 - 1 - D_2| + |x_1 - 3 - D_2|\big)\big] \tag{87}$$

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 $x_1 \in [100, 104]$, (87) 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 (87) is a convex function of x_1 on $\mathbb{R}$. ∎

We now prove that the multistage-static problem is strongly time consistent.

Lemma 19 *The multistage-static problem is strongly time consistent, and the optimal value of the multistage-dynamic 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 (83). 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

$$\arg\min_{x_1 \in \mathbb{R}_+} \left(\tfrac{1}{2} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \big[|x_1 - 1 - D_2| \big] + \tfrac{1}{2} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \big[|x_1 - 3 - D_2| \big] \right) = \{102\}. \tag{88}$$

It is easily verified that for all $x_1 \in [100, 104]$, we may apply Theorem 4 to conclude that

$$\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \big[|x_1 - 1 - D_2| \big] = \big((x_1 - 101)^2 + 25 \big)^{\frac{1}{2}},$$

$$\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \big[|x_1 - 3 - D_2| \big] = \big((x_1 - 103)^2 + 25 \big)^{\frac{1}{2}}.$$

We conclude that for all $x_1 \in [100, 104]$,

$$\tfrac{1}{2} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \big[|x_1 - 1 - D_2| \big] + \tfrac{1}{2} \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \big[|x_1 - 3 - D_2| \big] \tag{89}$$

equals

$$g(x_1) := \tfrac{1}{2} \left(\big((x_1 - 101)^2 + 25 \big)^{\frac{1}{2}} + \big((x_1 - 103)^2 + 25 \big)^{\frac{1}{2}} \right). \tag{90}$$

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 (89) is a convex function of x_1 on $\mathbb{R}$. ∎

Combining Lemmas 18 and 19 completes the proof of Theorem 10.

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 multistage-dynamic problem). Note that this stands in contrast to the multistage-dynamic formulation, which always has an optimal base-stock policy by Observation 6. 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 [119]. Both the results of [119], and our own result, indicate that the structure of the optimal policy 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 $\epsilon \in \left(0, \frac{1}{2}(\sqrt{6} - 2)\right)$ be any sufficiently small strictly positive number. It may be easily verified that for any such ϵ, one has $\epsilon \in (0, \frac{1}{4})$, and

$$\frac{1}{2} - 2\epsilon - \epsilon^2 > 0. \tag{91}$$

Let us define $y_1 = 10 - \epsilon$, $\rho = 1$,

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

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

Then we prove the following.

Theorem 11 *Suppose ϵ satisfies (91). 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}.*

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 20

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

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 us define

$$\phi_{Q_1}(d) \triangleq \mathbb{E}_{Q_1}\Big[|10 - D_1 - D_2| \,\Big|\, \{D_2 = d\} \Big],$$

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 Proposition 3 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 $Q_1 \in \mathfrak{M}_1$, $\sup_{Q_2 \in \mathfrak{M}_2} E_{Q_2}[\phi_{Q_1}(D_2)] = E_{Q_2}[q(D_2)] = 3$. 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 $\phi_{Q_1}(d) \leq \max\big(\phi_{Q_1}(7 - \epsilon), \phi_{Q_1}(9 + \epsilon)\big)$ 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 $\max\big(\phi_{Q_1}(7 - \epsilon), \phi_{Q_1}(9 + \epsilon)\big) \leq \frac{3}{2}$. 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. ∎

Proof. [Proof of Theorem 11] 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 13 and Theorem 4 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

$$\sup_{Q_1 \in \mathfrak{M}_1, \ Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_1 \times Q_2} \Big[2|10 - \epsilon - D_1| + |10 - D_1 - D_2| \Big]$$

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

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

where the final equality follows from Lemma 20.

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 $D_1 \le 3 + \epsilon < 10 - \epsilon$ w.p.1 for all $Q_1 \in \mathfrak{M}_1$, and a straightforward contradiction argument (the details of which we omit), that $\hat{\pi}$ can not be optimal unless $\hat{x}_1 \le 10 - \epsilon$, 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

$$\inf_{\hat{x}_2 \in \mathbb{R}} \sup_{Q_1 \in \mathfrak{M}_1, Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_1 \times Q_2} \Big[\big| \max\{10 - \epsilon - D_1, \hat{x}_2\} - D_2 \big| \Big] > 3. \tag{92}$$

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 $\hat{x}_2 \le 7 + \frac{1}{2}\epsilon$, 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,

$$\sup_{Q_1 \in \mathfrak{M}_1, Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_1 \times Q_2} \Big[\big| \max\{10 - \epsilon - D_1, \hat{x}_2\} - D_2 \big| \Big] \tag{93}$$

is at least

$$\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{\tilde{Q}_1 \times Q_2} \Big[\big| \max\{10 - \epsilon - D_1, \hat{x}_2\} - D_2 \big| \Big]$$
$$= \sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \Big[\frac{1}{2} |\max\{7 - \epsilon, \hat{x}_2\} - D_2| + \frac{1}{2} |9 - \epsilon - D_2| \Big], \tag{94}$$

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 (94) is at least

$$\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \Big[\Big| \frac{1}{2} \max\{7 - \epsilon, \hat{x}_2\} + \frac{1}{2}(9 - \epsilon) - D_2 \Big| \Big]. \tag{95}$$

Note that

$$\frac{1}{2} \max\{7 - \epsilon, \hat{x}_2\} + \frac{1}{2}(9 - \epsilon) \ge \frac{1}{2}(7 - \epsilon) + \frac{1}{2}(9 - \epsilon)$$
$$= 8 - \epsilon. \tag{96}$$

Letting $z \triangleq \frac{1}{2}\max\{7 - \epsilon, \hat{x}_2\} + \frac{1}{2}(9 - \epsilon)$, note that (95) equals $\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\big[(z - D_2)^+ + (D_2 - z)^+\big]$. Applying Theorem 4 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 (95) equals

$$\left((\frac{1}{2}\max\{7 - \epsilon, \hat{x}_2\} + \frac{1}{2}(9 - \epsilon) - 8)^2 + 9 \right)^{\frac{1}{2}}. \tag{97}$$

Combining (96) with the fact that

$$
\begin{aligned}
\frac{1}{2}\max\{7 - \epsilon, \hat{x}_2\} + \frac{1}{2}(9 - \epsilon) &\leq \frac{1}{2}(7 + \frac{1}{2}\epsilon) + \frac{1}{2}(9 - \epsilon) \\
&= 8 - \frac{1}{4}\epsilon,
\end{aligned}
$$

we conclude that (97) is strictly greater than 3, completing the proof of (92) for the case $\hat{x}_2 \leq 7 + \frac{1}{2}\epsilon$.

Alternatively, if $\hat{x}_2 \geq 7 + \frac{1}{2}\epsilon$, 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, (93) is at least

$$\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[\frac{1}{(1+\epsilon)^2 + 1}|\hat{x}_2 - D_2| + \frac{(1+\epsilon)^2}{(1+\epsilon)^2 + 1}\left|\max\left\{10 - \epsilon - \frac{1+2\epsilon}{1+\epsilon}, \hat{x}_2\right\} - D_2\right| \right]. \tag{98}$$

It follows from convexity of the absolute value function that (98) is at least

$$\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[\left| \frac{1}{(1+\epsilon)^2 + 1}\hat{x}_2 + \frac{(1+\epsilon)^2}{(1+\epsilon)^2 + 1}\max\left\{10 - \epsilon - \frac{1+2\epsilon}{1+\epsilon}, \hat{x}_2\right\} - D_2 \right| \right]. \tag{99}$$

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 (99) equals

$$\sup_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[(z - D_2)^+ + (D_2 - z)^+\right].$$

Furthermore,

$$
\begin{aligned}
&\frac{1}{(1+\epsilon)^2 + 1}\hat{x}_2 + \frac{(1+\epsilon)^2}{(1+\epsilon)^2 + 1}\max\left\{10 - \epsilon - \frac{1+2\epsilon}{1+\epsilon}, \hat{x}_2\right\} \\
&\geq \frac{1}{(1+\epsilon)^2 + 1}\left(7 + \frac{1}{2}\epsilon\right) + \frac{(1+\epsilon)^2}{(1+\epsilon)^2 + 1}\left(10 - \epsilon - \frac{1+2\epsilon}{1+\epsilon}\right) \\
&= 8 + \frac{\frac{1}{2} - 2\epsilon - \epsilon^2}{(1+\epsilon)^2 + 1}\epsilon. \tag{100}
\end{aligned}
$$

Applying Theorem 4 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 (91)), we conclude that (99) equals

$$\left(\left(\frac{1}{(1+\epsilon)^2 + 1} \hat{x}_2 + \frac{(1+\epsilon)^2}{(1+\epsilon)^2 + 1} \max\left\{ 10 - \epsilon - \frac{1 + 2\epsilon}{1 + \epsilon}, \hat{x}_2 \right\} - 8 \right)^2 + 9 \right)^{\frac{1}{2}}. \quad (101)$$

Combining with (100) and (91), we conclude that (101) is strictly greater than 3, completing the proof of (92) for the case $\hat{x}_2 \leq 7 + \frac{1}{2}\epsilon$, which completes the proof. ∎

4.5   Conclusion

In this chapter, 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 a novel policy-centric definition for time consistency in this setting, and put our definition in the broad context of prior work on time consistency and rectangularity. More precisely, we defined two natural formulations for the relevant optimization problem. In the multistage-static formulation, the policy-maker cannot recompute his/her policy after observing realized demand. In the multistage-dynamic formulation, he/she is allowed to reperform his/her minimax computations at each stage. If these two formulations have a common optimal policy, we defined the policy to be time consistent, and the multistage-static problem to be weakly time consistent. If all optimal policies of the multistage-static problem are also optimal for the multistage-dynamic problem, we defined the multistage-static problem to be strongly time consistent.

Next, we gave sufficient conditions for weak and strong time consistency. Intuitively, our sufficient condition for weak time consistency coincides with the existence of an optimal base-stock policy in which the base-stock constants are monotone increasing. Our sufficient condition for strong time consistency can be interpreted in two ways. On the one hand, strong time consistency holds if the unique optimal

base-stock policy for the multistage-dynamic formulation is to order-up to 0 at each stage. Alternatively, we saw that this condition also has an interpretation in terms of requiring that the demand variances are sufficiently large relative to their respective means.

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. 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 dynamic formulations. We also proved that, although the multistage-dynamic formulation always has an optimal policy of base-stock form, there may be no such optimal policy for the multistage-static formulation. This stands in contrast to the analogous setting, analyzed in [180], in which only the mean and support of the demand distribution is known at each stage, for which it is known that such time inconsistency cannot occur. Furthermore, 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 multistage-dynamic 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 [54], 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.

Our work leaves many interesting directions for future research. The general question of time consistency remains poorly understood. Furthermore, our work has shown that this question can be quite subtle. For the particular model we consider here, it would be interesting to develop a better understanding of precisely when time consistency holds. It is also an intriguing question to understand how much our two formulations can differ in optimal value and policy, even when time inconsistency occurs, along the lines of [97]. 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 recently carried out in [119]. Of course, it is also an open challenge to understand the question of time consistency more broadly, how precisely the various definitions of time consistency presented throughout the literature relate to one-another, and more generally to understand the relationship between different ways to model multistage optimization under uncertainty.

4.6 Appendix

4.6.1 Proof of Theorem 4

Proof. [Proof of Theorem 4] 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

$$\psi(x) = cx + b(\mu - x). \tag{102}$$

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

$$\psi(x) = cx + \frac{(h-b)(x-\mu)}{2} + \frac{b+h}{2} \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q \left[|x - D| \right]. \tag{103}$$

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 $z \in \mathbb{R}$.

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 Problem (41). 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

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

is feasible for the primal Problem (38). It follows from Proposition 3 that $\bar{Q}$ is an optimal primal solution. Combining the above and simplifying the relevant algebra, we conclude that in this case

$$\psi(x) = \psi_1(x) := c\mu + \frac{b+h}{2} \mathfrak{f}(x) - \frac{b-h-2c}{2}(x - \mu). \tag{104}$$

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

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

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 $\hat{\lambda}$

110

is feasible for the dual Problem (41). 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 (38). It follows from Proposition 3 that $\hat{Q}$ is an optimal primal solution. Combining the above and simplifying the relevant algebra, we conclude that in this case

$$\psi(x) = \psi_2(x) := \frac{(h+c)\sigma^2 - (b-c)\mu^2}{\mu^2 + \sigma^2}x + b\mu. \tag{105}$$

We now use the above to complete the proof of the theorem. Note that since by assumption $b > c$, it follows from (102) 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

$$\frac{\sigma^2}{\mu^2} - \frac{b-c}{h+c} \quad \text{is the same sign as} \quad \frac{\mu^2 + \sigma^2}{2\mu} - \chi. \tag{106}$$

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

111

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 (106) 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 $\arg\min_{x \in \mathbb{R}} \psi(x) = [0, \frac{\mu^2+\sigma^2}{2\mu}]$.

4.6.2 Proof of Proposition 2

Proof. [Proof of Proposition 2] 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 (40) is at least $\mathbb{E}_{Q_1 \times Q_2^*}\left[\Psi(D_1, D_2)\right]$, which itself equals the sum of $c\mu$ and

$$\mathbb{E}_{Q_1}\left[\left(\delta\big((b-c)[0-D_1]_+ + (h+c)[D_1-0]_+\big) + (1-\delta)\big((b-c)[\tau-D_1]_+ + (h+c)[D_1-\tau]_+\big)\right)I(D_1 > 0)\right] \tag{107}$$

$$+\mathbb{E}_{Q_1}\left[\left(\delta\big((b-c)[0-D_1]_+ + (h+c)[D_1-0]_+\big) + (1-\delta)\big((b-c)[\tau-D_1]_+ + (h+c)[D_1-\tau]_+\big)\right)I(D_1 < 0)\right] \tag{108}$$

$$+\mathbb{E}_{Q_1}\left[\left(\delta\big((b-c)[0-D_1]_+ + (h+c)[D_1-0]_+\big) + (1-\delta)\big((b-c)[\tau-D_1]_+ + (h+c)[D_1-\tau]_+\big)\right)I(D_1 = 0)\right] \tag{109}$$

Note that if $P(D_1 > 0) > 0$, then (107) is at least

$$\mathbb{E}\left[\frac{\sigma^2}{\mu^2+\sigma^2}(h+c)D_1 + \frac{\mu^2}{\mu^2+\sigma^2}(b-c)(\frac{\mu^2+\sigma^2}{\mu} - D_1)\Big| D_1 > 0\right]P(D_1 > 0)$$

$$> \mathbb{E}\left[\frac{\mu^2}{\mu^2+\sigma^2}(b-c)D_1 + \frac{\mu^2}{\mu^2+\sigma^2}(b-c)(\frac{\mu^2+\sigma^2}{\mu} - D_1)\Big| D_1 > 0\right]P(D_1 > 0) \text{ by A1}$$

$$= (b-c)\mu P(D_1 > 0). \tag{110}$$

Similarly, if $P(D_1 < 0) > 0$, then (108) is at least

$$\mathbb{E}\left[-\frac{\sigma^2}{\mu^2+\sigma^2}(b-c)D_1 + \frac{\mu^2}{\mu^2+\sigma^2}(b-c)(\frac{\mu^2+\sigma^2}{\mu}-D_1)\Big|D_1 < 0\right]P(D_1 < 0)$$

$$= \mathbb{E}\left[(b-c)(\mu-D_1)\Big|D_1 < 0\right]P(D_1 < 0) > (b-c)\mu P(D_1 < 0). \qquad (111)$$

Furthermore, if $P(D_1 = 0) > 0$, then (109) equals $(b-c)\mu P(D_1 = 0)$. Combining with (110), (111), and the fact that the measure δ_0 attains value $b\mu$ (by Theorem 4), completes the proof. ∎

4.6.3 Proof of Theorem 5

Proof. [Proof of Theorem 5] Recall that $\eta := \frac{1}{2}(c_1+c_2)$, and $\mathfrak{f}(z) := \left((z-\mu)^2+\sigma^2\right)^{\frac{1}{2}}$ 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 Q be the probability measure described in (43), and $\lambda = (\lambda_0, \lambda_1, \lambda_2)$ the vector described in (44). 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 $g(d) \geq \max\left(h_1(d), h_2(d)\right)$ 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 $d \in \mathbb{R}$,

$$g(d) = \frac{1}{2}\mathfrak{f}^{-1}(\eta)(d-\eta)^2 + \frac{1}{2}\left(\mathfrak{f}(\eta) + c_1 - c_2\right). \qquad (112)$$

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 $\frac{1}{2}\left(\mathfrak{f}(\eta) + c_1 - c_2\right) \geq 0$. Combining with (112), 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 Problem (41). Also, it is easily verified that Q is feasible for the primal Problem (38). It follows from Proposition 3 that Q is an optimal primal solution, and λ is an optimal

dual solution. That these optimal solutions are unique then follows from the second part of Proposition 3 and a straightforward contradiction argument. Combining the above and simplifying the relevant algebra completes the proof. ∎

CHAPTER V

DISTRIBUTIONALLY ROBUST INVENTORY CONTROL WHEN DEMAND IS A MARTINGALE

This chapter is based on [201].

5.1 Introduction and literature review

In many practical settings of interest, demands are correlated over time (cf. [98, 169, 171]). As a result, there is a vast literature investigating inventory models with correlated demand, including: studies of the so-called bull-whip effect (cf. [35, 122, 167]); models with Markov-modulated demand (cf. [60, 98, 112]); and models with forecasting, including models in which demand follows an auto-regressive/moving average (ARMA) or exponentially smoothed process (cf. [8, 26, 73, 111, 127, 132, 150]); and models obeying the Martingale Model of Forecast Evolution (MMFE) and its many generalizations (cf. [56, 77, 87, 101, 128, 189]). Although several of these works offer insights into the qualitative impact of correlations on the optimal policy (and associated costs) when managing an inventory over time, these results are typically proven under very particular distributional assumptions, which assume perfect knowledge of all relevant distributions. This is potentially a significant problem, since various authors have previously noted that model misspecification when demand is correlated can lead to very sub-optimal policies (cf. [8]). Indeed, the use of such time series and forecasting models in Operations Research practice is well-documented, and concerns over the practical impact of model mis-specification have been raised repeatedly in the forecasting and Operations Research literature (cf. [62, 63]).

One approach taken in the literature to correcting for such model uncertainty is

so-called distributionally robust optimization as we discussed before. In this frame-work, one assumes that the joint distribution (over time) of the sequence of future demands belongs to some set of joint distributions, and solves the minimax problem of computing the control policy which is optimal against a worst-case distribution belonging to this set. Such a distributionally robust approach is motivated by the reality that perfect knowledge of the exact distribution of a given random process is rarely available (cf. [51, 52, 154, 195]). A typical minimax formulation is as below:

$$\min_{x \in \mathcal{X}} \max_{Q \in \mathfrak{M}} \mathbb{E}_Q \left[f(x, \xi) \right],$$

where $\mathcal{X}$ is a set of decisions and $\mathfrak{M}$ is a set of probability measures. The objective is to pick a decision x that minimizes the average cost of f under a worst-case distribution.

The application of such distributionally robust approaches to the class of inventory control problems was pioneered in [168], where it was assumed that $\mathfrak{M}$ contains all probability measures whose associated distributions satisfy certain moment constraints. Such an approach has been taken to many variants of the single-stage model since then (cf. [69, 71, 72, 85, 148, 208, 212]), and the single-stage distributionally robust model is quite broadly understood.

The analogous questions become more subtle in the multi-stage setting, due to questions regarding the specification of uncertainty in the underlying joint distribution. There have been two approaches taken in the literature, depending on whether the underlying optimization model is *static* or *dynamic* in nature. In a static formulation, one specifies a family of joint distributions for demand over time, typically by fixing various moments and supports (or some generalization thereof), and then solves an associated global minimax optimization problem (cf. [19, 47, 51, 52, 152, 174]). Such static formulations generally cannot be decomposed and solved by dynamic programming, because the distributional constraints do not contain sufficient information about how the distribution behaves under conditioning. Put another way, such models generally do not allow for the incorporation of realized demand information into

model uncertainty/robustness going forwards (i.e. re-optimization in real time), and are generally referred to as time-inconsistent in the literature (cf. [54, 97, 204]). This inability to incorporate realized demand information may make such approaches undesirable for analyzing models which explicitly consider the forecasting of future demand (cf. [117]). We note that although some of these static formulations have indeed been able to model settings in which information is revealed over time, e.g. the excellent work of [174] on factor-models, the fundamental inability to incorporate realized demand (in the sense of time in-consistency) remains.

Alternatively, in a dynamic formulation, the underlying family of potential joint distributions must implicitly satisfy certain conditional independence properties, and thus allow for a resolution by dynamic programming. The existence of such a decomposition is generally referred to as the *rectangularity property* (cf. [54, 97, 103]). For example, the set of all joint distributions for the vector of demands (D_1, D_2) such that: $\mathbb{E}[D_1] = 1, \mathbb{E}[D_2|D_1] = D_1$, and (D_1, D_2) has support on the non-negative integers is rectangular, since the feasible set of joint distributions may be decomposed as follows. To each possible realized value d for D_1 (i.e. each non-negative integer), we may associate a fixed collection $S(d)$ of possible conditional distributions for D_2 (i.e. those distributions with mean d and support on the non-negative integers). Furthermore, every feasible distribution for the vector (D_1, D_2) may be constructed by first selecting a feasible distribution $\mathcal{D}$ for D_1 (i.e. any distribution with mean one and support on the non-negative integers), and for each element d in the support of $\mathcal{D}$, setting the conditional distribution of D_2 (given $\{D_1 = d\}$) to be some fixed element of $S(d)$. Moreover, the set of joint distributions constructible in this manner is precisely the set of feasible distributions for the vector (D_1, D_2). Alternatively, if we had instead required that $\mathbb{E}[D_1] = \mathbb{E}[D_2] = 1$, and $\mathbb{E}[D_1 D_2] = 2$, the corresponding set of feasible joint distributions would *not* be rectangular, as it may be verified that such a decomposition is no longer possible. For a formal definition and

more complete/precise description of the rectangularity property, we refer the reader to [23, 54, 97, 103, 145, 204], and note that since various communities have studied several closely related notions at different times, a complete consensus on a common rigorous definition has not yet been reached.

There are several works which formulate dynamic programming approaches to distributionally robust/risk averse inventory models (cf. [2, 73, 180, 204]). More generally, such dynamic problems can typically be formulated as so-called robust Markov decision processes (MDP) (cf. [103, 145, 200]). However, to our knowledge, none of these works consider applications to correlated demand or forecasting models, with the exception of the very general Bayesian model considered recently in the excellent work of [117].

On a related note, to the best of our knowledge, there seems to have been no systematic study of the *qualitative* impact of positing different joint dependency structures in such multi-stage distributionally robust inventory control problems, i.e. seeing which insights previously derived under specific distributional assumptions extend to the distributionally robust setting, and furthermore what new insights manifest only in the distributionally robust setting. The quest to develop such an understanding in the broader context of stochastic optimization (not specifically inventory control) was recently initiated in [1], where the authors define the so-called *price of correlations* as the ratio between the optimal minimax value when all associated random variables (r.v.) are independent, and the setting where they may take any joint distribution belonging to the allowed family. Although the authors do not look specifically at any inventory problems, they stress the general importance of understanding how positing different joint distribution uncertainty impacts the underlying stochastic optimization.

Combining the above, we are led to the following question.

Question 1 *Can we construct effective dynamic distributionally robust variants of*

118

the time series and forecasting models used in Operations Research? Furthermore, can we develop a theory of how positing different correlation structures qualitatively impacts the optimal policy for such models?

We close this section by briefly reviewing a third major branch of literature on robust inventory control models, namely that of classical (i.e. deterministic) robust optimization, in which the only constraints made are on the supports of the associated random variables (cf. [13, 18]). In this setting, the worst-case distribution is always degenerate, namely a point-mass on a particular worst-case trajectory. Such models often lead to tractable global optimization problems for fairly complex models, and have been successfully applied to several inventory settings (cf. [12, 14, 20, 22, 33, 114]). Indeed, this is to be contrasted with the settings of both static and dynamic formulations for multi-stage distributionally robust optimization, where the question of tractability (under both the static and dynamic formulations) is less clear (cf. [21, 174, 200]). In spite of their potential computational advantages, one drawback of such classical robust approaches is that they may be overly conservative, and unable to capture the stochastic nature of many real-world problems (cf. [174]). We note that the precise relationship between classical robust and distributionally robust approaches remains an intriguing open question. On a related note, questions of robust time series models and their applications to multi-stage inventory models, similar to those we will consider in the present work, were very recently studied in the interesting work of [33]. Although the approach taken there was considerably more conservative than the approach taken in the present work, the authors were able to formulate a quite general notion of robust time series, and solve the associated minimax optimization problems as tractable convex optimization problems. We leave a more precise understanding of the precise relationship between our own work on distributionally robust formulations, and alternative approaches based on classical robust optimization (such as that of [33]), as an interesting direction for future

research.

5.1.1 Outline of chapter

The rest of the chapter proceeds as follows. In Sections 5.2.1 and 5.2.2, we introduce the independent-demand and martingale-demand models respectively. Our main results are stated in Section 5.2.3. In Section 5.3, we present the proof of our main theorem, the explicit expressions of optimal policy and cost. In Section 5.4, we prove our asymptotic results. Section 5.5 provides further insights onto our results and discusses an interesting non-monotonicity of the base-stock level in terms of the backorder cost. In Section 5.6, we summarize our contributions and present directions for future research. We also include a technical appendix in Section 5.7.

5.2 Main results

Before stating our main results, we briefly review the inventory model analyzed in [180], which we refer to as the independent-demand model.

5.2.1 Independent-demand model

Consider the following distributionally robust inventory control problem with backlogging, finite time horizon T, strictly positive linear backorder per-unit cost b, and a per-unit holding cost of 1, where we note that assuming a holding cost equal to 1 is without loss of generality due to a simple scaling argument. Let D_t be the (random) demand in period t, and x_t, be the inventory level at period t after placing an order, $t = 1, \ldots, T$, where we note that the order must be placed in period t before D_t is known. In particular, letting $D_{[t]} \triangleq (D_1, \ldots, D_t)$ and $D_{[0]} \triangleq \emptyset$, we require that x_t is a measurable function of $D_{[t-1]}$, $t = 1, \ldots, T$, and let x_0 denote the initial inventory level, which we assume throughout to lie in $[0, U]$. For $t \in \{1, \ldots, T\}$, the cost incurred in period t equals

$$C_t \triangleq b[D_t - x_t]_+ + [x_t - D_t]_+.$$

We define an admissible policy π to be a T-dimensional vector of measurable functions $\{x_t, t = 1, \ldots, T\}$, s.t. x_t is a measurable map from $\mathbb{R}^{t-1}$ to $\mathbb{R}$, satisfying $x_1 \geq x_0$, and $x_{t+1}(D_{[t]}) \geq x_t(D_{[t-1]}) - D_t$, which is equivalent to requiring that a non-negative amount of inventory is ordered in each period. Let Π denote the family of all admissible policies. Note that once a particular policy $\pi \in \Pi$ is specified, the associated costs $\{C_t, t = 1, \ldots, T\}$ are explicit functions of the vector of demands. Sometimes we will make this dependence of C_t and x_t on the particular policy π explicit, through the notation C_t^π, x_t^π. Suppose $U, b \in \mathbb{R}^+$ are fixed, and for $\mu \in [0, U]$, let $\mathfrak{M}(\mu)$ be the collection of all probability measures with support $[0, U]$ and mean μ, where we say that a probability measure P is supported on a set S if $P(S) = 1$. Furthermore, let **IND** denote the collection of all T-dimensional product measures s.t. all T marginal distributions belong to $\mathfrak{M}(\mu)$. In words, the joint distribution of demand belongs to **IND** iff the demand is independent across time periods, and the demand in each period has support $[0, U]$ and mean μ. Then the particular optimization problem considered in [180] Example 5.1.2 is

$$\inf_{\pi \in \Pi} \sup_{Q \in \mathbf{IND}} \mathbb{E}_Q\big[\sum_{t=1}^{T} C_t^\pi\big]. \tag{113}$$

In Example 5.1.2 of [180], the author proves that due to certain structural properties, Problem (113) can be reformulated as a dynamic program, as we now describe. For a measurable function $f : \mathbb{R} \to \mathbb{R}$, and probability measure Q, let $\mathbb{E}_Q\big[f(D)\big]$ denote the expected value of $f(D)$ when D is a random variable with law Q, assuming the expectation to be well-defined. To formally define the relevant dynamic program, we now define a sequence of functions $\{V^t, t \geq 1\}, \{f^t, t \geq 1\}, \{g^t, t \geq 1\}$, where each V^t, f^t, g^t is a mapping from $\mathbb{R}^2 \to \mathbb{R}$. In general, such a dynamic program would be phrased in terms of the so-called "cost-to-go" functions, with the t-th such function representing the remaining cost incurred by an optimal policy during periods $t, \ldots, T$, subject to the given state at time t. Here and throughout, we use the fact that the backorder and holding costs are the same in every period to simplify the relevant

121

concepts and notations. In particular, due to this symmetry and the associated self-reducibility which it induces, for all $T \geq 1$, it will suffice to define the aforementioned cost-to-go function for the first time period only. Indeed, in the following function definitions, $V^T(y, \mu)$ will coincide with the optimal value of Problem (113) when the initial inventory level is y and the associated mean is μ (we leave the dependence on U implicit), where f^T, g^T have analogous interpretations.

$$f^1(x, D) \triangleq b[D - x]_+ + [x - D]_+, g^1(x, \mu) \triangleq \sup_{Q \in \mathfrak{M}(\mu)} \mathbb{E}_Q f^1(x, D), V^1(y, \mu) \triangleq \inf_{z \geq y} g^1(z, \mu),$$

$$Q^1(x, \mu) \triangleq \arg \max_{Q \in \mathfrak{M}(\mu)} \mathbb{E}_Q f^1(x, D), \quad \chi^1(\mu) \triangleq \arg \min_{z \in [0, U]} g^1(z, \mu);$$

$$(114)$$

and for $T \geq 2$,

$$f^T(x, D) \triangleq b[D - x]_+ + [x - D]_+ + V^{T-1}(x - D, \mu), \quad g^T(x, \mu) \triangleq \sup_{Q \in \mathfrak{M}(\mu)} \mathbb{E}_Q f^T(x, D),$$

$$V^T(y, \mu) \triangleq \inf_{z \geq y} g^T(z, \mu), Q^T(x, \mu) \triangleq \arg \max_{Q \in \mathfrak{M}(\mu)} \mathbb{E}_Q f^T(x, D), \chi^T(\mu) \triangleq \arg \min_{z \in [0, U]} g^T(z, \mu).$$

$$(115)$$

We note that although in principle $\chi^t(\mu)$ is an optimization over $z \in \mathbb{R}$, it follows from a straightforward contradiction that restricting $\chi^1(\mu)$ to the interval $[0, U]$ is without loss of generality.

Let us recall the formal definition of a base-stock policy for such an inventory control problem.

Definition 4 (Base-stock policy) *A policy $\pi \in \Pi$ is said to be a base-stock policy if there exist constants $\{x_t^* \in \mathbb{R}, \ t = 2, \ldots, T\}$, s.t.*

$$x_t^\pi(D_{[t-1]}) = \max \left\{ x_{t-1}^\pi(D_{[t-2]}) - D_{t-1}, \ x_t^* \right\}, \quad t = 2, \ldots, T.$$

We now review the results of Lemma 5.1 and Example 5.1.2 of [180], which characterizes the optimal policy, value, and worst-case distribution for Problem 113, and

relates the solution to dynamic programming formulation (114) - (115). Let us define

$$\chi_{\text{IND}}(\mu, U, b) \triangleq \begin{cases} 0 & \text{if } \mu \leq \frac{U}{b+1}, \\ U & \text{if } \mu > \frac{U}{b+1}; \end{cases}$$

$$\text{Opt}_{\text{IND}}^T(\mu, U, b) \triangleq \begin{cases} Tb\mu & \text{if } \mu \leq \frac{U}{b+1}, \\ T(U - \mu) & \text{if } \mu > \frac{U}{b+1}. \end{cases}$$

Then the following is proven in [180]. Let $\mathcal{D}_\mu$ be the probability measure s.t. $\mathcal{D}_\mu(0) = 1 - \frac{\mu}{U}, \mathcal{D}_\mu(U) = \frac{\mu}{U}$.

Theorem 12 *For all $U, b \in \mathbb{R}^+, T \geq 1$, and $\mu, x_0 \in [0, U]$, Problem 113 always has an optimal base-stock policy π^*, in which the associated base-stock constants $\{x_t^*, t = 1, \ldots, T\}$ satisfy $x_t^* = \chi_{IND}(\mu, U, b)$ for all t. If $x_0 \leq \chi_{IND}(\mu, U, b)$, the optimal value of Problem 113 equals $\text{Opt}_{IND}^T(\mu, U, b)$; and for all $x_0 \in [0, U]$, the product measure s.t. all marginals are distributed according to law $\mathcal{D}$ belongs to $\arg\max_{Q \in \textbf{IND}} \mathbb{E}_Q[\sum_{t=1}^T C_t^{\pi^*}]$. Furthermore, the dynamic programming formulation (114) - (115) can be used to compute these optimal policies and values. In particular, for all $x_0 \in [0, U]$, $V^T(x_0, \mu)$ is the optimal value of Problem 113, $\mathcal{D}_\mu \in Q^T(x, \mu)$ for all $x \in [0, U]$, and $\chi_{IND}(\mu, U, b) \in \chi^T(\mu)$.*

5.2.2 Martingale-demand model

In this subsection, we formally define our distributionally robust martingale-demand model, and state some preliminary results. Let $T, b, h, \Pi, C_t^\pi, \mathfrak{M}(\mu)$ be exactly as defined for the independent-demand model in Subsection 5.2.1, i.e. the time horizon, backorder and holding costs, set of admissible policies, cost incurred in period t under policy π, and collection of all probability measures with support $[0, U]$ and mean μ, respectively. Furthermore, let **MAR** denote the collection of all discrete-time martingale sequences $(D_1, \ldots, D_T)$ s.t. all T marginal distributions belong to $\mathfrak{M}$. In words, the joint distribution of demand belongs to **MAR** iff the demand is a

martingale, and the demand in each period has support $[0, U]$ and mean μ. Then in analogy with Problem 113, the optimization problem of interest is

$$\inf_{\pi \in \Pi} \sup_{Q \in \mathbf{MAR}} \mathbb{E}_Q[\sum_{t=1}^{T} C_t^{\pi}]. \tag{116}$$

In analogy with the dynamic programming formulation (114) - (115) given for the independent-demand model, we now define an analogous formulation for the martingale-demand setting. We define the following functions.

$$\hat{f}^1(x, D) \triangleq b[D-x]_+ + [x-D]_+, \hat{g}^1(x, \mu) \triangleq \sup_{Q \in \mathfrak{M}(\mu)} \mathbb{E}_Q \hat{f}^1(x, D), \hat{V}^1(y, \mu) \triangleq \inf_{z \geq y} \hat{g}^1(z, \mu),$$

$$\hat{Q}^1(x, \mu) \triangleq \arg\max_{Q \in \mathfrak{M}(\mu)} \mathbb{E}_Q \hat{f}^1(x, D), \quad \hat{\chi}^1(\mu) \triangleq \arg\min_{z \in [0,U]} \hat{g}^1(z, \mu); \tag{117}$$

and for $T \geq 2$,

$$\hat{f}^T(x, D) \triangleq b[D-x]_+ + [x-D]_+ + \hat{V}^{T-1}(x-D, D), \quad \hat{g}^T(x, \mu) \triangleq \sup_{Q \in \mathfrak{M}(\mu)} \mathbb{E}_Q \hat{f}^T(x, D),$$

$$\hat{V}^T(y, \mu) \triangleq \inf_{z \geq y} \hat{g}^T(z, \mu), \hat{Q}^T(x, \mu) \triangleq \arg\max_{Q \in \mathfrak{M}(\mu)} \mathbb{E}_Q \hat{f}^T(x, D), \hat{\chi}^T(\mu) \triangleq \arg\min_{z \in [0,U]} \hat{g}^T(z, \mu). \tag{118}$$

We note that as in the independent-demand model, it follows from a straightforward contradiction that restricting $\hat{\chi}^t(\mu)$ to the interval $[0, U]$ is without loss of generality.

Let us recall the formal definition of a state-dependent base-stock policy for such inventory control problems.

Definition 5 (State-dependent base-stock policy) *A policy $\pi \in \Pi$ is said to be a state-dependent base-stock policy if there exist measurable functions $\{x_t^*(D_{[t-1]}), t = 2, \ldots, T\}$, s.t. $x_t^\pi(D_{[t-1]}) = \max\{x_{t-1}^\pi(D_{[t-2]}) - D_{t-1}, \ x_t^*(D_{[t-1]})\}$, $t = 2, \ldots, T$.*

Then the following theorem may be proven by a straightforward induction argument, in which one demonstrates that $\hat{V}^T(y, \mu)$ is a convex function of y for each fixed μ, and that the set **MAR** enjoys a suitable variant of the well-known rectangularity property, as defined and applied in, e.g [54], [103], [145]. The demonstration

of convexity follows nearly identically to the proof of Theorems 3 and 4 given in [2], and the remainder of the proof follows from a straightforward but tedious argument combining the definition of a martingale with the structure of our inventory control problem and certain associated measurability properties, and we omit the details.

Theorem 13 *For all $U, b \in \mathbb{R}^+, T \geq 1$, and $\mu, x_0 \in [0, U]$, Problem 116 always has an optimal state-dependent base-stock policy. Furthermore, the dynamic programming formulation (117) - (118) can be used to compute these optimal policies and values. In particular, $\hat{V}^T(x_0, \mu)$ is the optimal value of Problem 113. An optimal policy π^* may be constructed by selecting the associated state-dependent base-stock levels $\{x_t^*(D_{[t-1]}), t = 1, \ldots, T\}$ s.t. $x_1^* \in \hat{\chi}^T(\mu)$, and $x_t^*(D_{[t-1]}) \in \hat{\chi}^{T+1-t}(D_{t-1}), t = 2, \ldots, T$. For such an optimal policy, any random vector $D_{[T]}$ s.t. $D_1 \in \hat{Q}^T(x_1^{\pi^*}, \mu)$, and $D_t \in \hat{Q}^t\left(x_t^{\pi^*}(D_{[t-1]}), D_{t-1}\right)$ for $t = 2, \ldots, T$ belongs to $\arg\max_{Q \in \mathbf{MAR}} \mathbb{E}_Q[\sum_{t=1}^T C_t^{\pi^*}]$.*

Although the structure of the optimal policy for various inventory problems in which the demand is assumed to have some kind of martingale-like structure, e.g. MMFE, has been previously studied, in general it may be computationally intensive to calculate $x_t^*(D_{[t-1]})$, not to mention to find closed-form solutions. We do note that there are several known structural results about the associated optimal state-dependent base-stock levels, and refer the reader to [101], [127], [184], [197] for details.

5.2.3 Main results

We now state our main results. We begin by introducing some additional definitions and notations. Whenever possible, we will supress dependence on the parameters μ, U, b for simplicity. As a notational convenience, let us define all empty products to equal unity, and all empty sums to equal zero. For $j \in [-1, T]$, let

$$A_j^T \triangleq \begin{cases} U \prod_{k=j+1}^{T-1} \frac{k}{b+k} & \text{if } j \leq T - 1, \\ \frac{b+T}{T} U & \text{if } j = T; \end{cases}$$

and

$$B_j^T \triangleq \frac{j}{b+T}A_j^T.$$

Note that $A_{-1}^T = B_{-1}^T = B_0^T = 0$, $A_{T-1}^T = B_T^T = U$, and both $\{A_j^T, j = -1, \ldots, T\}$ and $\{B_j^T, j = -1, \ldots, T\}$ are increasing in j, and decreasing in T. For $x, \mu \in [0, U]$, let $\mathfrak{q}_{x,\mu}^T$ be the unique (up to sets of measure zero) probability measure s.t.

$$
\begin{cases}
\mathfrak{q}_{x,\mu}^T(A_j^T) = \frac{A_{j+1}^T - \mu}{A_{j+1}^T - A_j^T} \ , \ \mathfrak{q}_{x,\mu}^T(A_{j+1}^T) = \frac{\mu - A_j^T}{A_{j+1}^T - A_j^T} & \text{if } \mu \in (A_j^T, A_{j+1}^T], \ x \in [0, B_{j+1}^T); \\[2mm]
\mathfrak{q}_{x,\mu}^T(0) = 1 - \frac{\mu}{A_k^T} \ , \ \mathfrak{q}_{x,\mu}^T(A_k^T) = \frac{\mu}{A_k^T} & \text{if } \mu \in (A_j^T, A_{j+1}^T], \ x \in [B_k^T, B_{k+1}^T), \\[2mm]
& \qquad\qquad\qquad\qquad k \geq j+1; \\[2mm]
\mathfrak{q}_{x,\mu}^T(0) = 1 - \frac{\mu}{U} \ , \ \mathfrak{q}_{x,\mu}^T(U) = \frac{\mu}{U} & \text{if } \mu = 0 \text{ or } x = U.
\end{cases}
$$

Also, for $j \in [0, T]$, let

$$G_j^T(x, \mu) \triangleq (T - \frac{b+T}{A_j^T}\mu)x + (T - j)b\mu;$$

in which case we define

$$
\Gamma_\mu^T \triangleq
\begin{cases}
0 & \text{if } \mu = 0, \\[2mm]
j+1 & \text{if } \mu \in (A_j^{T+1}, A_{j+1}^{T+1}] \ , \ j \in [-1, T-1];
\end{cases}
$$

and note that $\mu \in (A_{\Gamma_\mu^{T-1}-1}^T, A_{\Gamma_\mu^{T-1}}^T]$ for all $T \geq 2$ and $\mu \in (0, U]$, while $0 = A_{\Gamma_0^{T-1}-1}^T$. We also define

$$\chi_{\text{MAR}}^T(\mu, U, b) \triangleq \beta_\mu^T \triangleq B_{\Gamma_\mu^T}^T \quad , \quad \text{Opt}_{\text{MAR}}^T(\mu, U, b) \triangleq G_{\Gamma_\mu^T}^T(\beta_\mu^T, \mu).$$

Then the explicit solution to Problem 116 is as follows.

Theorem 14 *For all $U, b \in \mathbb{R}^+, T \geq 1$, and $\mu, x_0 \in [0, U]$, an optimal policy π^* for Problem 116 may be constructed by selecting the associated state-dependent base-stock levels s.t. $x_1^* = \chi_{MAR}^T(\mu, U, b)$, and $x_t^*(D_{[t-1]}) = \chi_{MAR}^{T+1-t}(D_{t-1}, U, b)$. For such an optimal policy, the random vector $D_{[T]}^*$ s.t. $D_1^* = \mathfrak{q}_{x_1^*,\mu}^T$, and $D_t^* = \mathfrak{q}_{x_t^*(D_{[t-1]}^*), D_{t-1}^*}^{T+1-t}$ for $t \in [2, T]$ belongs to $\arg\max_{Q \in MAR} \mathbb{E}_Q[\sum_{t=1}^T C_t^{\pi^*}]$. If in addition $x_0 \leq \chi_{MAR}^T(\mu, U, b)$, the optimal value of Problem 116 equals $Opt_{MAR}^T(\mu, U, b)$.*

Interestingly, if the initial inventory level is sufficiently small, and the inventory manager uses the (optimal) policy π^* described in Theorem 14, and the adversary selects the (optimal) sequence of demands $(D_1^*, \ldots, D_T^*)$ described in Theorem 14, the resulting stochastic inventory process is quite intuitive. Let $\Lambda^T \triangleq T - \Gamma_\mu^T$, and for $t \in [1, \Lambda^T]$, let us also define

$$D_t^T \triangleq A_{\Gamma_\mu^T}^{T+1-t} \quad , \quad X_t^T \triangleq B_{\Gamma_\mu^T}^{T+1-t}. \tag{119}$$

Note that $\Lambda^T \in [1, T]$, and represents the first time that D_t^T reaches U. Then the corresponding inventory dynamics under an optimal inventory manager and adversary simplify as follows. Consider the following discrete time Markov chain $\{M^T(t) \triangleq (\mathcal{X}_t^T, \mathcal{D}_t^T), t \geq 1\}$, with randomized initial conditions, which we will later prove evolves identically to the optimal sequence of demands and post-ordering inventory levels described in Theorem 14.

$$\left(\mathcal{X}_1^T, \mathcal{D}_1^T\right) = \begin{cases} \left(X_1^T, D_1^T\right) & \text{w.p. } \frac{\mu}{D_1^T}, \\ \left(X_1^T, 0\right) & \text{w.p. } 1 - \frac{\mu}{D_1^T}; \end{cases}$$

for $t \in [2, \Lambda^T - 1]$,

$$\left(\mathcal{X}_t^T, \mathcal{D}_t^T\right) = \begin{cases} \left(\mathcal{X}_{t-1}^T, 0\right) & \text{w.p. } 1, \text{ if } \mathcal{D}_{t-1}^T = 0, \\ \left(X_t^T, D_t^T\right) & \text{w.p. } \frac{D_{t-1}^T}{D_t^T}, \text{ if } \mathcal{D}_{t-1}^T \neq 0, \\ \left(X_t^T, 0\right) & \text{w.p. } 1 - \frac{D_{t-1}^T}{D_t^T}, \text{ if } \mathcal{D}_{t-1}^T \neq 0; \end{cases}$$

for $t = \Lambda^T$,

$$\left(\mathcal{X}_t^T, \mathcal{D}_t^T\right) = \begin{cases} \left(\mathcal{X}_{t-1}^T, 0\right) & \text{w.p. } 1, \text{ if } \mathcal{D}_{t-1}^T = 0, \\ \left(X_t^T, U\right) & \text{w.p. } \frac{D_{t-1}^T}{U}, \text{ if } \mathcal{D}_{t-1}^T \neq 0, \\ \left(X_t^T, 0\right) & \text{w.p. } 1 - \frac{D_{t-1}^T}{U}, \text{ if } \mathcal{D}_{t-1}^T \neq 0; \end{cases}$$

for $t \geq \min\left(\Lambda^T + 1, T\right)$,

$$\left(\mathcal{X}_t^T, \mathcal{D}_t^T\right) = \begin{cases} \left(\mathcal{X}_{\Lambda^T}^T, 0\right) & \text{w.p. 1, if } \mathcal{D}_{\Lambda^T}^T = 0; \\ \left(U, U\right) & \text{w.p. 1, if } \mathcal{D}_{\Lambda^T}^T = U. \end{cases}$$

Corollary 5 *Suppose $U, b \in \mathbb{R}^+, T \geq 1, \mu \in [0, U]$, and $x_0 \in [0, \chi_{MAR}^T(\mu, U, b)]$. Then one can construct the optimal sequence of post-ordering inventory levels $\{x_t^{\pi^*}(D_{[t-1]}^*), t = 1, \ldots, T\}$, and optimal vector of demands $D_{[T]}^*$, as described in Theorem 14, on a common probability space with M^T s.t. $\{(x_t^{\pi^*}(D_{[t-1]}^*), D_t^*), t = 1, \ldots, T\}$ equals $\{(\mathcal{X}_t^T, \mathcal{D}_t^T), t = 1, \ldots, T\}$ w.p.1.*

We now give an alternate description of the dynamics described in Corollary 5, explicitly describing the randon amount of time until the corresponding optimal adversary demands either 0 or U. Let $D_0^T \triangleq \mu$, $X_0^T \triangleq \frac{\Gamma_\mu^T}{b+T+1} D_0^T$. Let Z^T denote the r.v., with support on the integers belonging to $[1, \Lambda^T]$, whose corresponding probability measure $\mathcal{Z}^T$ satisfies

$$\mathcal{Z}^T(t) = \begin{cases} \left(1 - \frac{D_{t-1}^T}{D_t^T}\right) \frac{\mu}{D_{t-1}^T} & \text{if } t \in [1, \Lambda^T - 1]; \\ \frac{\mu}{D_{t-1}^T} & \text{if } t = \Lambda^T. \end{cases}$$

Also, let Y^T denote a r.v., with support on $\{0, U\}$, independent of Z^T, whose corresponding probability measure $\mathcal{Y}^T$ satisfies

$$\mathcal{Y}^T(x) = \begin{cases} \frac{D_{\Lambda^T-1}^T}{U} & \text{if } x = U; \\ 1 - \frac{D_{\Lambda^T-1}^T}{U} & \text{if } x = 0. \end{cases}$$

Corollary 6 *Under the same assumptions, and on the same probability space, as described in Corollary 5, one can also construct Z^T, Y^T s.t. all of the following hold w.p.1. $\{(\mathcal{X}_t^T, \mathcal{D}_t^T), t = 1, \ldots, Z^T - 1\}$ equals $\{(X_t^T, D_t^T), t = 1, \ldots, Z^T - 1\}$. If $Z^T \leq \Lambda^T - 1$, then $(\mathcal{X}_t^T, \mathcal{D}_t^T) = (X_{Z^T}^T, 0)$ for all $t \in [Z^T, T]$. If $Z^T = \Lambda^T$, and $Y^T = U$, then $(\mathcal{X}_{\Lambda^T}^T, \mathcal{D}_{\Lambda^T}^T) = (X_{\Lambda^T}^T, U)$, and $(\mathcal{X}_t^T, \mathcal{D}_t^T) = (U, U)$ for all $t \in [\Lambda^T + 1, T]$. If $Z^T = \Lambda^T$, and $Y^T = 0$, then $(\mathcal{X}_t^T, \mathcal{D}_t^T) = (X_{\Lambda^T}^T, 0)$ for all $t \in [\Lambda^T, T]$.*

In words, the dynamics at optimality described in Corollaries 5 - 6 have the following interpretation. The adversary will always select a demand distribution with (at most) two-point support, with one of these points equal to zero. The adversary reasons that, if her demand in some period t happens to be zero, the martingale property ensures that she will order zero in all subsequent periods. This ensures that the inventory manager is "stuck" holding all inventory she held at the start of period t (after ordering) for the entire remainder of the time horizon. As the martingale property ensures that the probability mass at zero is maximized when all other probability mass is put at U, one might guess that the optimal adversary will always put all probability on 0 and U, as in the independent-demand model. However, as Corollaries 5 - 6 show, this is not the case. The adversary does not put any probability mass at U until time Λ^T, if ever. The reason is that in the martingale-demand model, there is an additional "hidden cost" for the adversary associated with putting support on U. In particular, if a demand of U ever occurs, then the martingale property ensures that all future demands must be U as well. However, it is "easy" for an inventory manager to handle an adversary that always orders U - in particular, she can simply order up to U in every period, incurring zero cost. In other words, the aforementioned hidden cost to the adversary comes in the form of a "loss of randomness", making the adversary perfectly predictable going forwards. Corollaries 5 - 6 indicate that at optimality, this tradeoff manifests (at optimality) by having the adversary always put some probability at 0, and some probability on a different quantity D_t^T, which increases monotonically to U as $t \uparrow \Lambda^T$. This "ramping up" can be explained by observing that although higher values for this second point of the support result in a greater "loss of randomness", the cost of such a loss becomes smaller over time, as the associated window of time during which the adversary is perfectly predictable shrinks accordingly. The inventory manager will similarly "ramp up" there post-ordering inventory levels, to levels which similarly

address the aforementioned trade-offs.

To better understand these dynamics, we now prove that the Markov chain $M(t), t = 1, \ldots, T$ converges weakly to a simple limiting process. Let us define

$$\gamma \triangleq \frac{\mu}{U}, \qquad \Lambda^\infty \triangleq 1 - \gamma^{b-1}.$$

Let Z^∞ denote the mixed (i.e. both continuous and discrete components) random variable, with continuous support on $[0, \Lambda^\infty)$, and discrete support on the singleton Λ^∞, whose corresponding probability measure $\mathcal{Z}^\infty$ satisfies

$$\mathcal{Z}^\infty(\alpha) = \begin{cases} b(1-\alpha)^{b-1}d\alpha & \text{if } \alpha \in [0, \Lambda^\infty); \\ \gamma & \text{if } \alpha = \Lambda^\infty. \end{cases}$$

For $\alpha \in [0, \Lambda^\infty]$, where α is again a continuous parameter, let us also define

$$D_\alpha^\infty \triangleq \mu(1-\alpha)^{-b} \quad , \quad X_\alpha^\infty \triangleq \mu\gamma^{b-1}(1-\alpha)^{-(b+1)}.$$

Note that $D_{\Lambda^\infty}^\infty = X_{\Lambda^\infty}^\infty = U$. We now define an appropriate limiting process. Let $\mathcal{M}^\infty(\alpha)_{0 \leq \alpha \leq 1}$ denote the following two-dimensional process, constructed on the same probability space as Z^∞.

$$\mathcal{M}^\infty(\alpha) = \begin{cases} \left(X_\alpha^\infty, D_\alpha^\infty\right) & \text{if } \alpha \in [0, Z^\infty); \\ \left(X_{Z^\infty}^\infty, 0\right) & \text{if } \alpha \geq Z^\infty \text{ and } Z^\infty < \Lambda^\infty; \\ \left(U, U\right) & \text{if } \alpha \geq Z^\infty \text{ and } Z^\infty = \Lambda^\infty. \end{cases}$$

In the definition of $\mathcal{M}^\infty(\alpha)$, Z^∞ can be regarded as a "stopping" time that the limiting demand in the process first reaches 0 or U. The limiting inventory level and limiting demand both increase before the demand hits 0 or U. When the demand hits 0, then the inventory manager is forever stuck holding all inventory she held at that time, which equals $X_{Z^\infty}^\infty$. When the demand hits U, then the inventory manager also orders up to U to handle the adversary.

Then we have the following weak convergence result. For $\alpha \in [0,1]$, let $\mathcal{M}^T(\alpha) \triangleq M(\lfloor \alpha T \rfloor)$. For an excellent review of the formal definition of weak convergence, and the relevant topological spaces and metrics, we refer the reader to [25] or [198].

Theorem 15 *For all $U, b > 0$ and $\mu \in (0, U)$, the sequence of stochastic processes $\{\mathcal{M}^T(\alpha)_{0 \le \alpha \le 1}, T \ge 1\}$ converges weakly to the process $\mathcal{M}^\infty(\alpha)_{0 \le \alpha \le 1}$ on the space $D\left([0,1], \mathbb{R}^2\right)$ under the J_1 topology.*

We conclude by noting interesting comparative results between the independent-demand and martingale-demand models. First, we prove that, all other parameters being equal, the expected cost incurred under the independent-demand model is the highest one among all models with given support and mean constraints. In particular, the expected cost incurred under the independent-demand model is *at least* the expected cost incurred under the martingale-demand model. In other words, the "martingale adversary" is, in a sense, "weaker" than the "independent adversary".

Theorem 16 *Let $\mathcal{M}$ be a family of T-dimensional measures s.t. all T marginal distributions belong to $\mathfrak{M}(\mu)$, and $Opt_{\mathcal{M}}^T(\mu, U, b) \triangleq \inf_{\pi \in \Pi} \sup_{Q \in \mathcal{M}} \mathbb{E}_Q[\sum_{t=1}^T C_t^\pi]$. Assume $x_0 = 0$. Then for all $U, b > 0, T \ge 1$, and $\mu \in [0, U]$, $Opt_{\mathcal{M}}^T(\mu, U, b) \le Opt_{IND}^T(\mu, U, b)$. In particular, $Opt_{MAR}^T(\mu, U, b) \le Opt_{IND}^T(\mu, U, b)$.*

Note that $\mathrm{Opt}_{MAR}^T(\mu, U, b) = \mathrm{Opt}_{IND}^T(\mu, U, b)$ when $\mu = 0$ or $\mu = U$. To gain further insight into Theorem 16, we analyze this ratio as $T \to \infty$.

Theorem 17 *For all $U, b \in \mathbb{R}^+$, and $\mu \in (0, U)$,*

$$\lim_{T \to \infty} \frac{Opt_{MAR}^T(\mu, U, b)}{Opt_{IND}^T(\mu, U, b)} = \begin{cases} 1 - \gamma^{b-1} & \text{if } \mu \le \frac{U}{b+1}, \\ \frac{(1 - \gamma^{b-1})b\mu}{U - \mu} & \text{if } \mu > \frac{U}{b+1}. \end{cases}$$

We also note the following corollary, which arises from Theorem 17 in the perfectly symmetric case.

Corollary 7 *Suppose $\mu = \frac{U}{2}$ and $b = 1$. Then*

$$\lim_{T \to \infty} \frac{Opt^T_{MAR}(\mu, U, b)}{Opt^T_{IND}(\mu, U, b)} = \frac{1}{2}.$$

5.3 Proof of Theorem 14 and Corollaries 5 - 6

In this section, we complete the proof of our main result, Theorem 14, which yields an explicit solution to Problem 116, as well as Corollaries 5 - 6. We proceed to prove Theorem 14 by combining ideas from convex analysis with the fundamental properties of martingales to explicitly compute all desired functions in an inductive manner. We actually prove a more general theorem, explicitly computing all quantities relevant to the dynamic programming formulation 117 - 118. For $x, \mu \in [0, U]$, and $j \in [-1, T-1]$, let

$$F^T_j(x, \mu) \triangleq -bx + (b + T)B^T_{j+1} + \big(Tb - (b+1)(j+1)\big)\mu.$$

Also, let us define

$$\mathfrak{g}^T(x, \mu) \triangleq \begin{cases} F^T_j(x, \mu) & \text{if } \mu \in (A^T_j, A^T_{j+1}] \, , \ x \in [0, B^T_{j+1}); \\[2mm] G^T_k(x, \mu) & \text{if } \mu \in (A^T_j, A^T_{j+1}] \, , \ x \in [B^T_k, B^T_{k+1}) \, , \ k \ge j+1; \\[2mm] G^T_k(x, 0) & \text{if } \mu = 0, \ x \in [B^T_k, B^T_{k+1}); \\[2mm] G^T_T(U, \mu) & \text{if } x = U. \end{cases}$$

For later proofs, it will be convenient to note that $\mathfrak{g}$ may be equivalently expressed as follows. Let

$$\Upsilon^T_x \triangleq \begin{cases} T & \text{if } x = U, \\[2mm] j & \text{if } x \in [B^{T+1}_j, B^{T+1}_{j+1}); \end{cases}$$

and note that $x \in [B^T_{\Upsilon^{T-1}_x}, B^T_{\Upsilon^{T-1}_x+1})$ for all $T \ge 2$ and $x \in [0, U)$, while $U = B^T_{\Upsilon^{T-1}_U+1}$. Noting that $G^T_T(U, \mu) = G^T_{T-1}(U, \mu)$ implies

$$\mathfrak{g}^T(x, \mu) = \begin{cases} F^T_{\Gamma^{T-1}_\mu - 1}(x, \mu) & \text{if } x < B^T_{\Gamma^{T-1}_\mu}; \\[2mm] G^T_{\Upsilon^{T-1}_x}(x, \mu) & \text{if } x \ge B^T_{\Gamma^{T-1}_\mu}. \end{cases}$$

In light of Theorem 13, note that to prove Theorem 14, it suffices to prove the following.

Theorem 18 *Under the same assumptions as Theorem 14, $\hat{g}^T(x, \mu) = \mathfrak{g}^T(x, \mu)$, $\mathfrak{q}^T_{x,\mu} \in \hat{Q}^T(x, \mu)$, and $\chi^T_{MAR}(\mu, U, b) \in \hat{\chi}^T(\mu)$. Furthermore,*

$$\hat{V}^T(x_0, \mu) = \mathfrak{g}^T\left(\max\left(x_0, \chi^T_{MAR}(\mu, U, b)\right), \mu\right);$$

and if $x_0 \le \chi^T_{MAR}(\mu, U, b)$, then $\hat{V}^T(x_0, \mu) = Opt^T_{MAR}(\mu, U, b)$. Moreover, the above follows from the fact that:

(I) *$\mathfrak{g}^T(x, d)$ is a continuous and convex function of x on $(0, U)$, and a right (left) continuous function of x at 0 (U);*

(II) *$\beta^T_d \in \arg\min_{z \in [0, U]} \mathfrak{g}^T(z, d)$;*

(III) *$\mathfrak{g}^T(x, d) = \max_{Q \in \mathfrak{M}(d)} \mathbb{E}_Q\left[b(D-x)_+ + (x-D)_+ + \mathfrak{g}^{T-1}\left(\max\left(x-D, \beta^{T-1}_D\right), D\right)\right]$;*

(IV) *$\mathfrak{q}^T_{x,d} \in \arg\max_{Q \in \mathfrak{M}(d)} \mathbb{E}_Q\left[b(D-x)_+ + (x-D)_+ + \mathfrak{g}^{T-1}\left(\max\left(x-D, \beta^{T-1}_D\right), D\right)\right]$;*

(V) *$\mathfrak{g}^T\left(\beta^T_d, d\right) = Opt^T_{MAR}(d, U, b)$.*

We proceed to prove Theorem 18 by proving (I) - (V), in order. First, we briefly provide some intuition behind the form of $\mathfrak{g}^T$. Note that in the statement of our main results, namely Theorem 14, the case corresponding to the setting in which $\mathfrak{g}^T = F^T$ does not arise. Indeed, that case corresponds to the setting in which the post-ordering inventory level is sufficiently small (relative to the mean) that the adversary is not sufficiently incentivized to put probability mass at 0, as being "stuck" with so little inventory would not incur significant costs. As can be seen by examining the form of $\mathfrak{q}^T$ in this setting, the adversary instead puts support on a demand distribution with two non-zero points. Our results imply that an optimal inventory manager will always

order up to a level which avoids this setting, i.e. the optimal post-ordering inventory levels always sufficiently incentivize the adversary to put positive probability at 0. However, to prove our results, all these cases must be considered.

We begin our analysis by reviewing a well-known sufficient condition for convexity of non-differentiable functions. Recall that for a one-dimensional function $f(x)$, the right-derivative of f evaluated at x_0, which we denote by $\partial_x^+ f(x_0)$, equals $\lim_{h \downarrow 0} \frac{f(x_0+h)-f(x_0)}{h}$. When this limit exists, we say that f is right-differentiable at x_0. Then the following sufficient condition for convexity is stated in [164] Section 5, Proposition 18.

Lemma 21 ([164]) *A one-dimensional function $f(x)$, which is continuous and right-differentiable on an open interval (a,b) with non-decreasing right-derivative on (a,b), is convex on (a,b).*

We now use Lemma 21 to complete the proof of Theorem 18.(I)

Proof. [Proof of Theorem 18.(I)] We first prove continuity. That $\mathfrak{g}^T(x,d)$ is a continuous function of x on $(0,U) \setminus \bigcup_{i=\Gamma_d^{T-1}}^{T-1}\{B_i^T\}$, and a right-continuous function of x on $[0,U] \setminus \{U\}$, follows from definitions, and the fact that $F_i^T(x,d)$ and $G_i^T(x,d)$ are continuous functions of x on $[0,U]$ for all i. It similarly follows that $\lim_{x \uparrow B_i^T} \mathfrak{g}^T(x,d)$ exists for all $i \in [\Gamma_d^{T-1}, T] \setminus \{0\}$. It thus suffices to demonstrate that $\lim_{x \uparrow B_i^T} \mathfrak{g}^T(x,d)$ equals $\mathfrak{g}^T(B_i^T, d)$ for all $i \in [\Gamma_d^{T-1}, T] \setminus \{0\}$. We treat two cases: $i = \Gamma_d^{T-1}$, and $i \in [\Gamma_d^{T-1}+1, T]$, and begin with the case $i = \Gamma_d^{T-1}$. By assumption we preclude the case $i = 0$. Thus suppose $i = \Gamma_d^{T-1} \in [1, T]$. In this case,

$$
\begin{aligned}
\lim_{x \uparrow B_i^T} \mathfrak{g}^T(x,d) &= F_{i-1}^T(B_i^T, d) \\
&= -bB_i^T + (b+T)B_i^T + \big(Tb - (b+1)i\big)d. \\
&= TB_i^T + \big(Tb - (b+1)i\big)d.
\end{aligned}
$$

134

Alternatively,

$$
\begin{aligned}
\mathfrak{g}^T(B_i^T, d) &= G_i^T(B_i^T, d) \\
&= (T - \frac{b+T}{A_i^T} d) B_i^T + (T - i) b d \\
&= T B_i^T + \big(Tb - (b+1)i\big) d.
\end{aligned}
$$

Combining the above completes the proof for this case. Next, suppose $i \in [\Gamma_d^{T-1} + 1, T]$. In this case,

$$
\begin{aligned}
\lim_{x \uparrow B_i^T} \mathfrak{g}^T(x, d) &= G_{i-1}^T(B_i^T, d) \\
&= (T - \frac{b+T}{A_{i-1}^T} d) B_i^T + \big(T - (i-1)\big) b d \\
&= T B_i^T + \big(Tb - (b+1)i\big) d.
\end{aligned}
$$

Alternatively,

$$
\begin{aligned}
\mathfrak{g}^T(B_i^T, d) &= G_i^T(B_i^T, d) \\
&= (T - \frac{b+T}{A_i^T} d) B_i^T + (T - i) b d \quad = \quad T B_i^T + \big(Tb - (b+1)i\big) d,
\end{aligned}
$$

completing the proof.

We now prove convexity. As in our proof of continuity, that $\mathfrak{g}^T(x, d)$ is a right-differentiable function of x on $(0, U) \backslash \bigcup_{i=\Gamma_d^{T-1}}^{T} \{B_i^T\}$, with non-decreasing right-derivative on the same set, follows from definitions and the fact that $F_i^T(x, d)$ and $G_i^T(x, d)$ are linear functions of x on $[0, U]$ for all i. It similarly follows that $\mathfrak{g}^T(x, d)$ is a right-differentiable function of x on $[0, U] \setminus \{U\}$, and that $\lim_{x \uparrow B_i^T} \partial_x^+ \mathfrak{g}^T(x, d)$ exists for all $i \in [\Gamma_d^{T-1}, T-1] \setminus \{0\}$. It thus suffices to demonstrate that $\lim_{x \uparrow B_i^T} \partial_x^+ \mathfrak{g}^T(x, d) \leq \partial_x^+ \mathfrak{g}^T(B_i^T, d)$ for all $i \in [\Gamma_d^{T-1}, T-1] \setminus \{0\}$. We treat two cases: $i = \Gamma_d^{T-1}$, and $i \in [\Gamma_d^{T-1} + 1, T-1]$, and begin with the case $i = \Gamma_d^{T-1}$. By assumption we preclude the cases $i = 0, T$. Thus suppose $i = \Gamma_d^{T-1} \in [1, T-1]$. Then

$$
\begin{aligned}
\lim_{x \uparrow B_i^T} \partial_x^+ \mathfrak{g}^T(x, d) &= \lim_{x \uparrow B_i^T} \partial_x^+ F_{i-1}^T(x, d) \\
&= -b.
\end{aligned}
$$

135

Alternatively,

$$
\begin{aligned}
\partial_x^+ \mathfrak{g}^T(B_i^T, d) &= \partial_x^+ G_i^T(B_i^T, d) \\
&= T - \frac{b+T}{A_i^T} d. \\
&\geq T - \frac{b+T}{A_i^T} A_i^T \quad = \quad -b.
\end{aligned}
$$

Combining the above completes the proof for this case. Next, suppose $i \in [\Gamma_d^{T-1} + 1, T-1]$. Then

$$
\begin{aligned}
\lim_{x \uparrow B_i^T} \partial_x^+ \mathfrak{g}^T(x, d) &= \lim_{x \uparrow B_i^T} \partial_x^+ G_{i-1}^T(x, d) \\
&= T - \frac{b+T}{A_{i-1}^T} d.
\end{aligned}
$$

Alternatively,

$$
\begin{aligned}
\partial_x^+ \mathfrak{g}^T(B_i^T, d) &= \partial_x^+ G_i^T(B_i^T, d) \\
&= T - \frac{b+T}{A_i^T} d.
\end{aligned}
$$

The desired result then follows from the fact that A_i^T is increasing in i and d is non-negative. Combining the above completes the proof. ∎

We now use Theorem 18.(I) to complete the proof of Theorem 18.(II). Before proceeding, it will be useful to prove that the sets $\{A_j^T, j = 0, \dots, T-2\}$ and $\{A_j^{T+1}, j = 0, \dots, T-2\}$ interlace, in an appropriate sense. Indeed, since A_{j+1}^{T+1} is trivially strictly less than A_{j+1}^T for all $j \in [0, T-2]$, and $\frac{A_j^T}{A_{j+1}^{T+1}} = \frac{\frac{b+j+1}{j+1}}{\frac{T}{b+T}} < 1$ for all $j \in [0, T-2]$, we conclude the following.

Lemma 22 $A_j^T < A_{j+1}^{T+1} < A_{j+1}^T$ for $j \in [0, T-2]$. It follows that for all $T \geq 2$ and $d \in [0, U]$, $\Gamma_d^T \in \{\Gamma_d^{T-1}, \Gamma_d^{T-1} + 1\}$.

We now complete the proof of Theorem 18.(II)

Proof. [Proof of Theorem 18.(II)] Let $i = \Gamma_d^T, j = \Gamma_d^{T-1}$. From Theorem 18.(I), it suffices to prove that $\partial_x^+ \mathfrak{g}^T(x, d) \leq 0$ for all $x < B_i^T$, and $\partial_x^+ \mathfrak{g}^T(B_i^T, d) \geq 0$; or

136

that $\partial_x^+ \mathfrak{g}^T(x, d) \leq 0$ for all $x < U$ and $B_i^T = U$. It follows from Lemma 22 that $i \in \{j, j+1\}$. We now proceed by a case analysis. First, suppose $i = j$. In that case, $B_i^T = B_j^T$, $j \leq T - 1$, and $B_i^T < U$. We conclude that for all $x < B_i^T$,

$$\partial_x^+ \mathfrak{g}^T(x, d) \quad = \quad \partial_x^+ F_{j-1}^T(x, \mu) \quad = \quad -b.$$

Noting that

$$
\begin{aligned}
\partial_x^+ \mathfrak{g}^T\left(B_i^T, d\right) \quad &= \quad \partial_x^+ G_i^T(B_i^T, d) \\
&= \quad T - \frac{b+T}{A_i^T} d \quad \geq \quad T - \frac{b+T}{A_i^T} A_i^{T+1} \quad = \quad 0
\end{aligned}
$$

completes the proof in this setting.

Alternatively, suppose that $i = j + 1$. In this case, $B_i^T > B_j^T$, and

$$
\begin{aligned}
\lim_{x \uparrow B_i^T} \partial_x^+ \mathfrak{g}^T\left(x, d\right) \quad &= \quad \lim_{x \uparrow B_i^T} \partial_x^+ G_{i-1}^T(x, d) \\
&= \quad T - \frac{b+T}{A_{i-1}^T} d \quad \leq \quad T - \frac{b+T}{A_{i-1}^T} A_{i-1}^{T+1} \quad = \quad 0.
\end{aligned}
$$

If $B_i^T = U$, the lemma follows from Theorem 18.(I). Otherwise,

$$
\begin{aligned}
\partial_x^+ \mathfrak{g}^T\left(B_i^T, d\right) \quad &= \quad \partial_x^+ G_i^T(B_i^T, d) \\
&= \quad T - \frac{b+T}{A_i^T} d \quad \geq \quad T - \frac{b+T}{A_i^T} A_i^{T+1} \quad = \quad 0.
\end{aligned}
$$

Combining the above completes the proof. $\blacksquare$

We now embark on the proof of Theorem 18.(III) - (IV). This is where the primary difficulty lies in the proof of our main results. First, to simplify matters, we will prove an auxiliary lemma simplifying the expression $\mathfrak{g}^{T-1}\left(\max\left(\beta_D^{T-1}, x - D\right), D \right)$, and begin with some further definitions. For $x, d \in [0, U]$, and $j \in [0, T]$, let

$$\overline{F}_j^T(x, d) \triangleq Tx + \left((b-1)T - bj - \frac{b+T}{A_j^T} x\right) d + \frac{b+T}{A_j^T} d^2;$$

and for $j \in [-1, T-1]$, let

$$\overline{G}_j^T(d) \triangleq TB_{j+1}^T + \left(Tb - (b+1)(j+1)\right) d.$$

Note that β_d^T is monotone increasing in d, with $\beta_0^T = 0$, and $\beta_U^T = U$. It follows that for all $x \in [0, U]$,

$$z_x^T \triangleq \inf\{d \geq 0 \quad \text{s.t.} \quad \beta_d^T \geq x - d\},$$

is well defined,

$$z_x^T \leq x \quad \text{for} \quad x \in [0, U], \tag{120}$$

and $z_x^T = 0$ iff $x = 0$. For $x, d \in [0, U]$, let us define

$$\overline{\mathfrak{g}}^T(x, d) \triangleq \begin{cases} \overline{F}_j^T(x, d) & \text{if } d \in [0, z_x^T) \bigcap (x - B_{j+1}^T, x - B_j^T] \ , \ j \in [0, T-1]; \\ \overline{G}_j^T(d) & \text{if } d \in [z_x^T, U] \bigcap (A_j^{T+1}, A_{j+1}^{T+1}] \ , \ j \in [-1, T-1]; \\ \overline{G}_{-1}^T(0) & \text{if } d = 0 \ , x = 0; \\ \overline{F}_{T-1}^T(U, 0) & \text{if } d = 0 \ , x = U. \end{cases}$$

For later proofs, it will be convenient to note that $\overline{\mathfrak{g}}$ may be equivalently expressed as follows.

$$\overline{\mathfrak{g}}^T(x, d) = \begin{cases} \overline{F}_{\Upsilon_{x-d}^{T-1}}^T(x, d) & \text{if } d < z_x^T; \\ \overline{G}_{\Gamma_d^T - 1}^T(d) & \text{if } d \geq z_x^T. \end{cases}$$

We now prove that $\overline{\mathfrak{g}}^T(x, d) = \mathfrak{g}^T\left(\max\left(\beta_d^T, x - d\right), d\right)$.

Lemma 23 *For all $T \geq 1$, $x, d \in [0, U]$, $\overline{\mathfrak{g}}^T(x, d) = \mathfrak{g}^T\left(\max\left(\beta_d^T, x - d\right), d\right)$.*

Proof. Let us proceed by showing that for each fixed $x \in [0, U]$, $\overline{\mathfrak{g}}^T(x, d) = \mathfrak{g}^T\left(\max\left(\beta_d^T, x - d\right), d\right)$ for all $d \in [0, U]$. As the equivalence is easily verified for the case $d = 0$, in which case $\overline{\mathfrak{g}}^T(x, d) = \mathfrak{g}^T\left(\max\left(\beta_d^T, x - d\right), d\right) = Tx$, suppose $d > 0$. We proceed by a case analysis, beginning with the setting $d \in (0, z_x^T)$. In this case, $\max\left(\beta_d^T, x - d\right) = x - d$, and

$$x \quad > \quad d + B_{\Gamma_d^T}^T \quad \geq \quad d + B_{\Gamma_d^{T-1}}^T,$$

138

where the second inequality follows from Lemma 22 and the monotonicity (in i) of B_i^T. Combining with the easily verified fact that $\overline{F}_j^T(x,d) = G_j^T(x-d,d)$ for all j completes the proof in this case.

Alternatively, suppose $d \in [z_x^T, U]$, which implies that $\max(\beta_d^T, x-d) = \beta_d^T$, and $d \geq x - \beta_d^T$. We proceed by a case analysis. First, suppose $d \in (A_j^T, A_{j+1}^{T+1}]$ for some $j \in [-1, T-2]$. In this case, Lemma 22 implies that $d \in (A_j^T, A_{j+1}^T] \bigcap (A_j^{T+1}, A_{j+1}^{T+1}]$, and

$$\overline{\mathfrak{g}}^T(x,d) \;=\; \overline{G}_j^T(d).$$
$$=\; G_{j+1}^T(B_{j+1}^T, d) \;=\; \mathfrak{g}^T(\beta_d^T, d).$$

Alternatively, suppose $d \in (A_{j+1}^{T+1}, A_{j+1}^T]$ for some $j \in [-1, T-2]$. In this case, Lemma 22 implies that $d \in (A_j^T, A_{j+1}^T] \bigcap (A_{j+1}^{T+1}, A_{j+2}^{T+1}]$, and

$$\overline{\mathfrak{g}}^T(x,d) \;=\; \overline{G}_{j+1}^T(d).$$
$$=\; G_{j+2}^T(B_{j+2}^T, d) \;=\; \mathfrak{g}^T(\beta_d^T, d).$$

Lemma 22 implies that this treats all cases. Combining the above completes the proof. ∎

We now make one further definition, to further simplify notations. Let us define

$$\mathfrak{f}^T(x,d) \overset{\Delta}{=} b(d-x)_+ + (x-d)_+ + \overline{\mathfrak{g}}^{T-1}(x,d).$$

Then in follows from Lemma 23 that to prove Theorem 18.(III) - (IV), it suffices to demonstrate the following.

$$\mathfrak{g}^T(x,d) = \sup_{Q \in \mathfrak{M}(d)} \mathbb{E}_Q[\mathfrak{f}^T(x,D)] \quad, \quad \text{and} \quad \mathfrak{q}_{x,d}^T \in \arg\max_{Q \in \mathfrak{M}(d)} \mathbb{E}_Q[\mathfrak{f}^T(x,D)]. \quad (121)$$

To accomplish this, we will rely on the following easily verified sufficient condition for optimality in certain distributionally robust optimization problems.

Lemma 24 *Suppose $f : [0, U] \to \mathbb{R}$ is any bounded real-valued function with domain $[0, U]$. Suppose $0 \leq L \leq \mu < R \leq U$, and the linear function $\eta(d) \overset{\Delta}{=} \frac{f(R)-f(L)}{R-L}d +$*

$\frac{Rf(L)-Lf(R)}{R-L}$, *i.e. the line intersecting f at the points L and R, satisfies $\eta(d) \geq f(d)$ for all $d \in [0, U]$, i.e. lies above f on $[0, U]$. Then the measure q s.t. $q(L) = \frac{R-\mu}{R-L}, q(R) = \frac{\mu-L}{R-L}$, belongs to $\arg\max_{Q \in \mathfrak{M}(\mu)} \mathbb{E}_Q[f(D)]$.*

Proof.   Note that for any measure $Q \in \mathfrak{M}(\mu)$, it is true that

$$\mathbb{E}_Q[f(D)] \leq \mathbb{E}_Q[\eta(D)] = \frac{f(R)-f(L)}{R-L}\mu + \frac{Rf(L)-Lf(R)}{R-L}.$$

But since q has support only on points $d \in [0, U]$ s.t. $f(d) = \eta(d)$, it follows that

$$\mathbb{E}_q[f(D)] = \mathbb{E}_q[\eta(D)] = \frac{f(R)-f(L)}{R-L}\mu + \frac{Rf(L)-Lf(R)}{R-L}.$$

Combining the above completes the proof. ∎

In order to apply Lemma 24 to $\mathfrak{f}^T$, thus solving the distributionally robust optimization problem associated with (121), we will prove that $\mathfrak{f}^T$ has a very special structure. In particular, we will prove that it is a continuous "gluing together" of convex and concave functions, in an appropriate sense. We begin by proving certain relevant structural properties for $\overline{\mathfrak{g}}^T$, which we will use to prove corresponding structural properties for $\mathfrak{f}^T$.

Lemma 25 *For each fixed $x \in [0, U]$, $\overline{\mathfrak{g}}^T(x, d)$ is a continuous function of d on $(0, U)$, and a right (left) - continuous function of d at 0 (U). Also, $\overline{\mathfrak{g}}^T(x, d)$ is a convex function of d on $(0, z_x^T)$, and a concave function of d on (z_x^T, U).*

Proof.   First, let us treat the case $d \in [0, z_x^T)$, and begin by proving continuity. Right-continuity at 0 when $x \neq 0$ follows from the fact that $\lim_{d\downarrow 0} \overline{F}_j^T(x, d) = Tx$ for all j, and right-continuity at 0 when $x = 0$ follows from definitions. That $\overline{\mathfrak{g}}^T(x, d)$ is a continuous function of d on $(0, z_x^T) \setminus \bigcup_{j=1}^{T-1}\{x - B_j^T\}$, and a left-continuous function of d on $(0, z_x^T)$, follows from the continuity (in d) of $\overline{F}_j^T(x, d)$ for all j. It similarly follows that $\lim_{d\downarrow x-B_j^T} \overline{\mathfrak{g}}^T(x, d)$ exists for all j s.t. $x - B_j^T \in (0, z_x^T)$, and it thus suffices to demonstrate that $\lim_{d\downarrow x-B_j^T} \overline{\mathfrak{g}}^T(x, d)$ equals $\overline{\mathfrak{g}}^T(x, x - B_j^T)$ for all $j \in [1, T-1]$ s.t.

$x - B_j^T \in (0, z_x^T)$. Note that for all such j,

$$
\begin{aligned}
\lim_{d \downarrow x - B_j^T} \bar{\mathfrak{g}}^T(x, d) &= \overline{F}_{j-1}^T(x, x - B_j^T) \\
&= Tx + \left((b-1)T - b(j-1) - \frac{b+T}{A_{j-1}^T}x\right)(x - B_j^T) + \frac{b+T}{A_{j-1}^T}(x - B_j^T)^2.
\end{aligned}
$$

Alternatively,

$$
\begin{aligned}
\bar{\mathfrak{g}}^T(x, x - B_j^T) &= \overline{F}_j^T(x, x - B_j^T) \\
&= Tx + \left((b-1)T - bj - \frac{b+T}{A_j^T}x\right)(x - B_j^T) + \frac{b+T}{A_j^T}(x - B_j^T)^2.
\end{aligned}
$$

It follows that $\bar{\mathfrak{g}}^T(x, x - B_j^T) - \lim_{d \downarrow x - B_j^T} \bar{\mathfrak{g}}^T(x, d)$ equals

$$
\begin{aligned}
&\left(-b + \left(\frac{b+T}{A_{j-1}^T} - \frac{b+T}{A_j^T}\right)x\right)(x - B_j^T) - \left(\frac{b+T}{A_{j-1}^T} - \frac{b+T}{A_j^T}\right)(x - B_j^T)^2 \\
&= (x - B_j^T)\left(-b + \left(\frac{b+T}{A_{j-1}^T} - \frac{b+T}{A_j^T}\right)x - \left(\frac{b+T}{A_{j-1}^T} - \frac{b+T}{A_j^T}\right)(x - B_j^T)\right) \\
&= (x - B_j^T)\left(-b + (b+T)B_j^T\left(\frac{1}{A_{j-1}^T} - \frac{1}{A_j^T}\right)\right) = 0,
\end{aligned}
$$

completing the proof of continuity.

We now prove convexity. Again applying Lemma 21, it suffices to demonstrate that $\partial_d^+ \bar{\mathfrak{g}}^T(x, d)$ exists and is non-decreasing on $(0, z_x^T)$. Since $\overline{F}_j^T(x, d)$ is a convex quadratic function of d for all j, we conclude that: $\partial_d^+ \bar{\mathfrak{g}}^T(x, d)$ exists on $(0, z_x^T)$; $\partial_d^+ \bar{\mathfrak{g}}^T(x, d)$ is non-decreasing on $(0, z_x^T) \setminus \bigcup_{j=1}^{T-1}\{x - B_j^T\}$; and $\lim_{d \uparrow x - B_j^T} \partial_d^+ \bar{\mathfrak{g}}^T(x, d)$ exists for all $j \in [1, T-1]$ s.t. $x - B_j^T \in (0, z_x^T)$. It thus suffices to demonstrate that $\lim_{d \uparrow x - B_j^T} \partial_d^+ \bar{\mathfrak{g}}^T(x, d) \le \partial_d^+ \bar{\mathfrak{g}}^T(x, x - B_j^T)$ for all $j \in [1, T-1]$ s.t. $x - B_j^T \in (0, z_x^T)$. Note that for any such j,

$$
\begin{aligned}
\lim_{d \uparrow x - B_j^T} \partial_d^+ \bar{\mathfrak{g}}^T(x, d) &= \partial_d^+ \overline{F}_j^T(x, x - B_j^T) \\
&= (b-1)T - bj - \frac{b+T}{A_j^T}x + 2\frac{b+T}{A_j^T}(x - B_j^T) \\
&= (b-1)T - (b+2)j + x\frac{b+T}{A_j^T}.
\end{aligned}
$$

141

Alternatively, it follows from continuity that

$$\partial_d^+ \overline{\mathfrak{g}}^T(x, x - B_j^T) = \partial_d^+ \overline{F}_{j-1}^T(x, x - B_j^T)$$

$$= (b-1)T - b(j-1) - \frac{b+T}{A_{j-1}^T}x + 2\frac{b+T}{A_{j-1}^T}(x - B_j^T)$$

$$= (b-1)T - (b+2)j - b + x\frac{b+T}{A_{j-1}^T}.$$

It follows that

$$\partial_d^+ \overline{\mathfrak{g}}^T(x, x - B_j^T) - \lim_{d \uparrow x - B_j^T} \partial_d^+ \overline{\mathfrak{g}}^T(x, d) \tag{122}$$

equals

$$-b + x\frac{b+T}{A_{j-1}^T} - x\frac{b+T}{A_j^T},$$

which will be the same sign as

$$-bA_{j-1}^T + x(b+T) - x(b+T)\frac{j}{b+j} = -bA_{j-1}^T + x(b+T)\frac{b}{b+j}. \tag{123}$$

Noting that $\frac{b+j}{b+T}A_{j-1}^T = B_j^T$, and multiplying through the right-hand side of (123) by $\frac{b+j}{b(b+T)}$, we further conclude that (122) will be the same sign as $x - B_j^T$. Since by assumption $x - B_j^T \in (0, z_x^T)$, we conclude that $x - B_j^T \geq 0$, completing the proof of continuity and convexity for $d \in (0, z_x^T)$ and right-continuity at 0.

Next, let us treat the case $d \in (z_x^T, U]$, and begin by proving continuity. That $\overline{\mathfrak{g}}^T(x, d)$ is a continuous function of d on $(z_x^T, U) \setminus \bigcup_{j=0}^{T-1}\{A_j^{T+1}\}$, and a left-continuous function of d on $(z_x^T, U]$, follows from the continuity (in d) of $\overline{G}_j^T(d)$ for all j. It similarly follows that $\lim_{d \downarrow A_j^{T+1}} \overline{\mathfrak{g}}^T(x, d)$ exist for all $j \in [0, T-1]$ s.t. $A_j^{T+1} > z_x^T$. It thus suffices to demonstrate that $\lim_{d \downarrow A_j^{T+1}} \overline{\mathfrak{g}}^T(x, d) = \overline{\mathfrak{g}}^T(x, A_j^{T+1})$ for all $j \in [0, T-1]$ s.t. $A_j^{T+1} > z_x^T$. Note that for any such j,

$$\lim_{d \downarrow A_j^{T+1}} \overline{\mathfrak{g}}^T(x, d) = \overline{G}_j^T(A_j^{T+1})$$

$$= TB_{j+1}^T + \big(Tb - (b+1)(j+1)\big)A_j^{T+1}.$$

142

Alternatively,

$$
\begin{aligned}
\bar{\mathfrak{g}}^T(x, A_j^{T+1}) &= \overline{G}_{j-1}^T(A_j^{T+1}) \\
&= TB_j^T + \big(Tb - (b+1)j\big)A_j^{T+1}.
\end{aligned}
$$

Thus

$$
\begin{aligned}
\bar{\mathfrak{g}}^T(x, A_j^{T+1}) - \lim_{d \downarrow A_j^{T+1}} \bar{\mathfrak{g}}^T(x, d) &= T(B_j^T - B_{j+1}^T) + (b+1)A_j^{T+1} \\
&= \frac{T}{b+T}(jA_j^T - (j+1)A_{j+1}^T) + (b+1)\frac{T}{b+T}A_j^T \\
&= \frac{T}{b+T}\big((b+j+1)A_j^T - (j+1)A_{j+1}^T\big) = 0,
\end{aligned}
$$

completing the proof of continuity.

We now prove concavity. Again applying Lemma 21, it suffices to demonstrate that $\partial_d^+ \bar{\mathfrak{g}}^T(x, d)$ exists and is non-increasing on (z_x^T, U). Since $\overline{G}_j^T(d)$ is a linear function of d for all j, it follows from the piece-wise definition of $\bar{\mathfrak{g}}^T(x, d)$ that demonstrating the desired concavity is equivalent to showing that $\partial_d^+ \overline{G}_j^T(0)$ is non-increasing in j. Noting that $\partial_d^+ \overline{G}_j^T(0) = Tb - (b+1)(j+1)$, which is trivially decreasing in j, completes the proof.

Finally, let us prove continuity at z_x^T. We consider two cases, depending on how $\eta_d \overset{\Delta}{=} x - d$ comes to go from lying above β_d^T to lying below β_d^T. This "crossing" can occur in two ways. In particular, either β_d^T and η_d actually intersect, or z_x^T occurs at a jump discontinuity of β_d^T and the two functions never truly intersect. We proceed by a case analysis. Let $i = \Gamma_{z_x^T}^T$.

First, suppose that β_d^T and η_d actually intersect at z_x^T, namely $B_i^T = x - z_x^T$. If $z_x^T \in \bigcup_{j=-1}^T \{A_j^{T+1}\}$, the proof of right-continuity follows identically to our previous proof of right-continuity at A_j^{T+1} for all $j \in [0, T-1]$ s.t. $A_j^{T+1} > z_x^T$, and we omit the details. Otherwise, right-continuity at z_x^T follows from definitions. Either way,

we need only demonstrate left-continuity. Note that $z_x^T \in (x - B_{i+1}^T, x - B_i^T]$, and

$$
\begin{aligned}
\lim_{d \uparrow z_x^T} \overline{\mathfrak{g}}^T(x, d) &= \overline{F}_i^T(x, z_x^T) \\
&= Tx + \left((b-1)T - bi - \frac{b+T}{A_i^T} x \right)(x - B_i^T) + \frac{b+T}{A_i^T}(x - B_i^T)^2 \\
&= TB_i^T + \left(Tb - (b+1)i \right)(x - B_i^T).
\end{aligned}
$$

Alternatively,

$$
\begin{aligned}
\overline{\mathfrak{g}}^T(x, z_x^T) &= \overline{G}_{i-1}^T(x - B_i^T) \\
&= TB_i^T + \left(Tb - (b+1)i \right)(x - B_i^T),
\end{aligned}
$$

completing the proof of continuity in this case.

Alternatively, suppose that β_d^T and η_d do not truly intersect at z_x^T. In this case, $z_x^T = A_i^{T+1} \in (x - B_{i+1}^T, x - B_i^T]$, and

$$
\begin{aligned}
\lim_{d \uparrow z_x^T} \overline{\mathfrak{g}}^T(x, d) &= \overline{F}_i^T(x, A_i^{T+1}) \\
&= Tx + \left((b-1)T - bi - \frac{b+T}{A_i^T} x \right) A_i^{T+1} + \frac{b+T}{A_i^T}(A_i^{T+1})^2 \\
&= b(T - i) A_i^{T+1},
\end{aligned}
$$

where the final equality follows from straightforward algebraic manipulations, the details of which we omit. Alternatively,

$$
\begin{aligned}
\overline{\mathfrak{g}}^T(x, z_x^T) &= \overline{G}_{i-1}^T(A_i^{T+1}) \\
&= TB_i^T + \left(Tb - (b+1)i \right) A_i^{T+1} \\
&= b(T - i) A_i^{T+1},
\end{aligned}
$$

where the final equality again follows from straightforward algebraic manipulations. This completes the proof of left-continuity at z_x^T. The proof of right-continuity in this case follows identically to our previous proof of right-continuity at A_j^{T+1} for all $j \in [0, T-1]$ s.t. $A_j^{T+1} > z_x^T$, and we omit the details. Combining the above completes the proof of the lemma. ∎

We now use Lemma 25 to prove the relevant structural properties for $\mathfrak{f}^T$, and begin by introducing some additional notations to simplify our arguments. For $x \in [0, U]$, let us define $\alpha_x^T \triangleq A_{\Gamma_x^{T-1}}^T, \zeta_x^T \triangleq \Gamma_{z_x^T}^T, \mathcal{A}_x^T \triangleq A_{\zeta_x^{T-1}}^T$, and $\aleph_x^T \triangleq A_{\Upsilon_x^{T-1}}^T$. Note that for all $j \in \{-1, \ldots, T-1\}$,

$$
\begin{aligned}
B_{j+1}^T &= \frac{j+1}{b+T} A_{j+1}^T \\
&\leq \frac{j+1}{b+j+1} A_{j+1}^T = A_j^T.
\end{aligned}
\tag{124}
$$

Combining with (120) and a straightforward contradiction argument, we conclude that for all $x \in [0, U]$ and $T \geq 2$,

$$
\Upsilon_x^T \geq \Gamma_x^T \geq \zeta_x^T \quad, \quad \text{and} \quad \aleph_x^T \geq \alpha_x^T \geq \mathcal{A}_x^T.
\tag{125}
$$

Then the following lemma demonstrates the desired structural properties of $\mathfrak{f}^T$.

Lemma 26 *For each fixed $x \in [0, U]$, $\mathfrak{f}^T(x, d)$ is a continuous function of d on $(0, U)$, and a right (left) - continuous function of d at 0 (U). Also, for all $j \in [-1, T-2]$, $\mathfrak{f}^T(x, d)$ is a convex function of d on (A_j^T, A_{j+1}^T). Furthermore, $\mathfrak{f}^T(x, d)$ is a convex function of d on $(0, \mathcal{A}_x^T)$, and a concave function of d on (α_x^T, U).*

Proof. The statements regarding continuity follow from Lemma 25. Noting that $d > \alpha_x^T$ implies $d > x$, concavity on (α_x^T, U) follows from (120) and Lemma 25. Let $i = \zeta_x^{T-1}$. Convexity on $(0, z_x^{T-1})$ follows from (120) and Lemma 25. Supposing $z_x^{T-1} \notin \{0, U\}$, it follows from definitions that $z_x^{T-1} \in (A_{i-1}^T, A_i^T]$, where $\mathcal{A}_x^T = A_i^T$. Combining with the convexity of $\overline{F}_j^{T-1}(x, d)$ and $\overline{G}_j^{T-1}(d)$ for all j, to prove the lemma, it suffices to demonstrate that: 1. $z_x^{T-1} \notin \bigcup_{j=-1}^{T-1} \{A_j^T\}$ implies that

$$
\lim_{d \uparrow z_x^{T-1}} \partial_d^+ \mathfrak{f}^T(x, d) \leq \partial_d^+ \mathfrak{f}^T(x, z_x^{T-1});
\tag{126}
$$

and 2. $x \notin \bigcup_{j=-1}^{T-1} \{A_j^T\}$ implies

$$
\lim_{d \uparrow x} \partial_d^+ \mathfrak{f}^T(x, d) \leq \partial_d^+ \mathfrak{f}^T(x, x).
\tag{127}
$$

145

We treat several cases, and assume throughout that $z_x^{T-1}, x \notin \bigcup_{j=-1}^{T-1}\{A_j^T\}$. First, suppose $z_x^{T-1} = x$. In this case, it follows from a straightforward contradiction argument that $\Gamma_x^T = i = 0$, $x \in (0, A_0^T)$, and

$$
\begin{aligned}
\lim_{d\uparrow z_x^{T-1}} \partial_d^+ \mathfrak{f}^T(x, d) &= -1 + \lim_{d\uparrow z_x^{T-1}} \partial_d^+ \overline{F}_0^{T-1}(x, d) \\
&= -1 + (b-1)(T-1) - \frac{b+T-1}{A_0^{T-1}}x + 2\frac{b+T-1}{A_0^{T-1}}z_x^{T-1} \\
&= -1 + (b-1)(T-1) + \frac{b+T-1}{A_0^{T-1}}x \\
&\leq -1 + (b-1)(T-1) + \frac{b+T-1}{A_0^{T-1}}A_0^T \quad = \quad bT - (b+1).
\end{aligned}
$$

Alternatively,

$$
\begin{aligned}
\partial_d^+ \mathfrak{f}^T(x, z_x^{T-1}) &= b + \partial_d^+ \overline{G}_{-1}^{T-1}(x) \\
&= b + (T-1)b \quad = \quad bT.
\end{aligned}
$$

Combining the above completes the proof of both (126) and (127) in this case.

Next, suppose $z_x^{T-1} < x$. We again consider two cases, depending on how $\eta_d \overset{\Delta}{=} x - d$ comes to go from lying above β_d^{T-1} to lying below β_d^{T-1}. In the case that the two actually intersect, it may be easily verified that $z_x^{T-1} = x - B_i^{T-1}$, and

$$
\begin{aligned}
\lim_{d\uparrow z_x^{T-1}} \partial_d^+ \mathfrak{f}^T(x, d) &= -1 + \lim_{d\uparrow z_x^{T-1}} \partial_d^+ \overline{F}_i^{T-1}(x, d) \\
&= -1 + (b-1)(T-1) - bi - \frac{b+T-1}{A_i^{T-1}}(z_x^{T-1} + B_i^{T-1}) \\
&\quad + 2\frac{b+T-1}{A_i^{T-1}}z_x^{T-1} \\
&= -1 + (b-1)(T-1) - (b+1)i + \frac{b+T-1}{A_i^{T-1}}z_x^{T-1}.
\end{aligned}
$$

Alternatively,

$$
\begin{aligned}
\partial_d^+ \mathfrak{f}^T(x, z_x^{T-1}) &= -1 + \partial_d^+ \overline{G}_{i-1}^{T-1}(z_x^{T-1}) \\
&= -1 + (T-1)b - (b+1)i.
\end{aligned}
$$

146

It follows that

$$\partial_d^+ \mathfrak{f}^T(x, z_x^{T-1}) - \lim_{d\uparrow z_x^{T-1}} \partial_d^+ \mathfrak{f}^T(x, d) \quad = \quad T - 1 - \frac{b+T-1}{A_i^{T-1}} z_x^{T-1},$$

which will be the same sign as

$$A_i^T - z_x^{T-1} \quad \geq \quad A_i^T - A_i^T \quad = \quad 0.$$

Combining the above completes the proof of (126) in this case. Furthermore, the proof of (127) follows from the fact that $x \notin \bigcup_{j=-1}^{T-1}\{A_j^T\}$, from which it follows that

$$\partial_d^+ \mathfrak{f}^T(x, x) - \lim_{d\uparrow x} \partial_d^+ \mathfrak{f}^T(x, d) = b + 1.$$

Finally, suppose that $z_x^{T-1} < x$ and $z_x^{T-1} \notin \bigcup_{j=1}^{T-2}\{x - B_j^{T-1}\}$. As $x \notin \bigcup_{j=-1}^{T-1}\{A_j^T\}$, this final case follows nearly identically to the proof of (127) in the previous case, and we omit the details. Combining all of the above cases completes the proof of the lemma. ∎

We now use Lemma 26 to explicitly construct (for each fixed value of x) a family $\mathcal{F}$ of lines $\{L_i\}$, s.t. each line L_i lies above $\mathfrak{f}^T(x, d)$ for all $d \in [0, U]$, each line L_i intersects $\mathfrak{f}^T(x, d)$ at exactly two points p_i^1, p_i^2, and for each $\mu \in [0, U]$ there exists $L_i \in \mathcal{F}$ s.t. $\mu \in [p_i^1, p_i^2]$. Our construction will ultimately allow us to apply Lemma 24, explicitly solve the distributionally robust optimization problem $\max_{Q\in\mathfrak{M}(d)} \mathbb{E}_Q[\mathfrak{f}^T(x, D)]$, and complete the proof. We begin by explicitly constructing the family of lines $\mathcal{F}$. For $d \in \mathbb{R}$, let us define

$$\mathcal{K}^T(x, d) \triangleq \frac{\mathfrak{f}^T(x, \aleph_x^T) - Tx}{\aleph_x^T} d + Tx;$$

and for $j \in [-1, \ldots, T-2]$,

$$\mathcal{L}_j^T(x, d) \triangleq \frac{\mathfrak{f}^T(x, A_{j+1}^T) - \mathfrak{f}^T(x, A_j^T)}{A_{j+1}^T - A_j^T} d + \frac{A_{j+1}^T \mathfrak{f}^T(x, A_j^T) - A_j^T \mathfrak{f}^T(x, A_{j+1}^T)}{A_{j+1}^T - A_j^T}.$$

Noting that $\overline{F}_j^{T-1}(x, 0) = (T-1)x$ for all j, it follows from Lemma 23 that

$$\mathfrak{f}^T(x, 0) = Tx. \tag{128}$$

147

It may be easily verified, using (128), that $\mathcal{K}^T$ defines the unique line passing through the (x, y) co-ordinates $\left(0, \mathfrak{f}^T(x, 0)\right)$ and $\left(\aleph_x^T, \mathfrak{f}^T(x, \aleph_x^T)\right)$; and $\mathcal{L}_j^T$ defines the unique line passing through the (x, y) co-ordinates $\left(A_j^T, \mathfrak{f}^T(x, A_j^T)\right)$ and $\left(A_{j+1}^T, \mathfrak{f}^T(x, A_{j+1}^T)\right)$.

Lemma 27 *For each fixed $x \in [0, U]$, $\mathcal{K}^T(x, d) \geq \mathfrak{f}^T(x, d)$ for all $d \in [0, U]$. Also, for all $\ell \in [\Upsilon_x^{T-1}, T-2]$, $\mathcal{L}_\ell^T(x, d) \geq \mathfrak{f}^T(x, d)$ for all $d \in [0, U]$.*

Proof. We first prove that $\mathcal{K}^T(x, d) \geq \mathfrak{f}^T(x, d)$ for all $d \in [0, U]$. Let $i = \zeta_x^{T-1}, j = \Gamma_x^{T-1}$, and $k = \Upsilon_x^{T-1}$. That $\mathcal{K}^T(x, 0) \geq \mathfrak{f}^T(x, 0)$ follows from definitions. We now prove that $\mathcal{K}^T(x, A_l^T) \geq \mathfrak{f}^T(x, A_l^T)$ for all $l \in [i, j-1]$. First, it will be useful to rewrite $\mathcal{K}$ in a more convenient form. Noting that $\aleph_x^T = A_k^T$,

$$
\begin{aligned}
\mathfrak{f}^T(x, A_k^T) &= b(A_k^T - x) + \overline{G}_{k-1}^{T-1}(A_k^T) \\
&= b(A_k^T - x) + (T-1)B_k^{T-1} + \left((T-1)b - (b+1)k\right)A_k^T,
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{K}^T(x, d) &= \frac{\mathfrak{f}^T(x, A_k^T) - Tx}{A_k^T} d + Tx \\
&= \left(Tb - (b+1)k + \frac{(T-1)B_k^{T-1} - (b+T)x}{A_k^T}\right) d + Tx \\
&= \left(b(T-k) - \frac{(b+T)x}{A_k^T}\right) d + Tx. \tag{129}
\end{aligned}
$$

Combining with the fact that for all $l \in [i, j-1]$ one has $A_l^T \in [z_x^{T-1}, x]$, proving the desired statement is equivalent to proving that

$$
\left(b(T-k) - \frac{(b+T)x}{A_k^T}\right) A_l^T + Tx \geq x - A_l^T + (T-1)B_l^{T-1} + \left((T-1)b - (b+1)l\right)A_l^T,
$$

which is itself equivalent to demonstrating that

$$
\left(b(l+1-k) + 1 - \frac{(b+T)x}{A_k^T}\right) A_l^T + (T-1)x \geq 0. \tag{130}
$$

First, it will be useful to prove that the left-hand side of (130),

$$
\eta(l) \triangleq \left(b(l+1-k) + 1 - \frac{(b+T)x}{A_k^T}\right) A_l^T + (T-1)x,
$$

148

is decreasing in l, for $l \in [0, j-1]$. Indeed, after simplifying, we find that

$$\eta(l+1) - \eta(l) = \left(b + \frac{b}{b+l+1}\left(b(l+1-k)+1 - \frac{(b+T)x}{A_k^T}\right)\right)A_{l+1}^T,$$

which will be the same sign as

$$\left(b(l+2-k)+l+2\right)A_k^T - (b+T)x$$

$$\leq \left(b(l+2-k)+l+2\right)A_k^T - (b+T)B_k^T$$

$$= (b+1)(l+2-k)A_k^T.$$

Noting that $\ell + 1 \leq j - 1$ and (125) implies $j \leq k$ completes the proof of monotonicity, which we now use to complete the proof of (130). In particular, the above monotonicity implies that to prove (130), it suffices to prove that $\eta(k-1) \geq 0$. Note that

$$\eta(k-1) = A_{k-1}^T + \frac{b(T-1-k)-k}{b+k}x.$$

If $b(T-1-k)-k \geq 0$, then trivially $\eta(k-1) \geq 0$. Thus suppose $b(T-1-k)-k < 0$. In this case, $\eta(k-1) \geq 0$ iff

$$x \leq \frac{(b+k)A_{k-1}^T}{k - b(T-1-k)}.$$

As $x \leq B_{k+1}^T$, it thus suffices to prove that

$$B_{k+1}^T \leq \frac{b+k}{k - b(T-1-k)}A_{k-1}^T,$$

which, dividing both sides by $A_{k-1}^T(b+k)$ and simplifying, is itself equivalent to proving that

$$(b+T)k \geq \left(k - b(T-1-k)\right)(b+k+1).$$

Noting that $k \leq T - 1$, and thus $k - b(T-1-k) \leq k$, thus completes the desired proof that $\mathcal{K}^T(x, A_l^T) \geq \mathfrak{f}^T(x, A_l^T)$ for all $l \in [i, j-1]$.

We now prove that $\mathcal{K}^T(x, A_l^T) \geq \mathfrak{f}^T(x, A_l^T)$ for all $l \in [j, k]$. By construction,

$$\mathcal{K}^T(x, A_k^T) = \mathfrak{f}^T(x, A_k^T),$$

and thus it suffices to prove the desired claim for $l \in [j, k-1]$. Note that the degenerate case for which $x = A_k^T$ can be ignored, as in that case $j = k$. Thus suppose $x < A_k^T$. In this case, Lemma 26 implies that $\mathfrak{f}^T(x, d)$ is a continuous, concave, piecewise linear function of d on $[A_j^T, A_k^T]$. As $\mathcal{K}^T(x, d)$ is a linear function of d, it follows from the basic properties of concave functions that to prove the desired claim, it suffices to demonstrate that

$$\partial_d^+ \mathcal{K}(x, A_k^T) \leq \lim_{d \uparrow A_k^T} \partial_d^+ \mathfrak{f}^T(x, d), \tag{131}$$

which is equivalent to proving that

$$b(T - k) - \frac{(b + T)x}{A_k^T} \geq b + \lim_{d \uparrow A_k^T} \partial_d^+ \overline{G}_{k-1}^{T-1}(d). \tag{132}$$

It follows from definitions that

$$A_k^T \quad = \quad \frac{b+T}{k} B_k^T \quad \leq \quad \frac{b+T}{k} x.$$

Combining with (132), we find that to prove the desired claim, it suffices to demonstrate that

$$b(T - k) - k \geq b + (T - 1)b - (b + 1)k. \tag{133}$$

Noting that both sides of (133) are equivalent completes the proof.

Finally, let us prove that $\mathcal{K}^T(x, A_l^T) \geq \mathfrak{f}^T(x, A_l^T)$ for all $l \in [k + 1, T - 1]$, which will complete the proof. It again follows from Lemma 26 and the basic properties of concave functions that in this case it suffices to demonstrate that $\partial_d^+ \mathcal{K}^T(x, A_k^T) \geq \partial_d^+ \mathfrak{f}^T(x, A_k^T)$, which is equivalent to proving that

$$b(T - k) - \frac{(b + T)x}{A_k^T} \geq b + \partial_d^+ \overline{G}_k^{T-1}(A_k^T). \tag{134}$$

It follows from definitions that

$$
\begin{aligned}
A_k^T \quad &= \quad \frac{k+1}{b+k+1} A_{k+1}^T \\
&= \quad \frac{b+T}{b+k+1} B_{k+1}^T \quad \geq \quad \frac{b+T}{b+k+1} x.
\end{aligned}
$$

Combining with (134), we find that in this case it suffices to demonstrate that

$$b(T - k) - (b + k + 1) \geq b + (T - 1)b - (b + 1)(k + 1). \tag{135}$$

Noting that both sides of (135) are equivalent completes the proof. Combining all of the above with the piece-wise convexity guaranteed by Lemma 26 completes the proof that $\mathcal{K}^T(x, d) \geq \mathfrak{f}^T(x, d)$ for all $d \in [0, U]$.

We now prove that for all $l \in [k, T - 2]$, $\mathcal{L}_l^T(x, d) \geq \mathfrak{f}^T(x, d)$ for all $d \in [0, U]$. Note that $\mathfrak{f}^T(x, d)$ is a concave function on $[A_k^T, U]$, and by construction $\mathcal{L}_l^T(x, d)$ is a line tangent to $\mathfrak{f}^T(x, d)$ at A_l^T. It follows from the basic properties of concave functions that

$$\mathcal{L}_l^T(x, d) \geq \mathfrak{f}^T(x, d) \quad \text{for all} \quad d \in [A_k^T, U].$$

Combining those same properties with (131) and the basic properties of linear functions, it follows that $\mathcal{L}_l^T(x, d) \geq \mathcal{K}^T(x, d)$ for all $d \in [0, A_k^T]$. Combining all of the above completes the proof. ■

We now prove that by combining Lemmas 24 and 27, we can complete the proof of Theorem 18.(III) - (IV).

Proof. [Proof of Theorem 18.(III) - (IV)] Theorem 18.(IV) follows from Lemmas 24 and 27, definitions, and a straightforward case analysis, the details of which we omit. We now prove Theorem 18.(III), and proceed by a case analysis. Let $j = \Gamma_d^{T-1}$ and $k = \Upsilon_x^{T-1}$. First, suppose $d \in [0, A_k^T]$. In light of Lemmas 24 and 27, in this case it suffices to demonstrate that

$$\mathfrak{g}^T(x, d) = \mathcal{K}^T(x, d). \tag{136}$$

In this case, the left-hand side of (136) equals

$$G_k^T(x, d) = (T - \frac{b + T}{A_k^T} d)x + (T - k)bd. \tag{137}$$

Alternatively, from (129), the right-hand side of (136) equals

$$\left(b(T - k) - \frac{(b + T)x}{A_k^T}\right)d + Tx. \tag{138}$$

151

Noting that (137) equals (138) completes the proof in this case.

Alternatively, suppose $d > A_k^T$. In this case it suffices to demonstrate that for all $\ell \in [k, T-2]$,

$$\mathfrak{g}^T(x, d) = \mathcal{L}_{j-1}^T(x, d). \tag{139}$$

Note that the left-hand side of (139) equals

$$F_{j-1}^T(x, d) = -bx + (b+T)B_j^T + \big(Tb - (b+1)j\big)d. \tag{140}$$

Alternatively, it follows from (120) and (124) that the right-hand side of (139) equals

$$
\begin{aligned}
&\frac{\mathfrak{f}^T(x, A_j^T) - \mathfrak{f}^T(x, A_{j-1}^T)}{A_j^T - A_{j-1}^T} d + \frac{A_j^T \mathfrak{f}^T(x, A_{j-1}^T) - A_{j-1}^T \mathfrak{f}^T(x, A_j^T)}{A_j^T - A_{j-1}^T} \\
=& b(d-x) + \frac{\overline{G}_{j-1}^{T-1}(A_j^T) - \overline{G}_{j-2}^{T-1}(A_{j-1}^T)}{A_j^T - A_{j-1}^T} d + \frac{A_j^T \overline{G}_{j-2}^{T-1}(A_{j-1}^T) - A_{j-1}^T \overline{G}_{j-1}^{T-1}(A_j^T)}{A_j^T - A_{j-1}^T}.
\end{aligned}
\tag{141}
$$

We now simplify (141) separately. Note that $\overline{G}_{j-1}^{T-1}(A_j^T) - \overline{G}_{j-2}^{T-1}(A_{j-1}^T)$ equals

$$
\begin{aligned}
& (T-1)(B_j^{T-1} - B_{j-1}^{T-1}) + \left((T-1)b - (b+1)j\right)(A_j^T - A_{j-1}^T) - (b+1)A_{j-1}^T \\
=& \left((T-1)b - (b+1)j\right)(A_j^T - A_{j-1}^T) + (T-1)\left(\frac{j}{T-1}A_j^T - \frac{j-1}{T-1}A_{j-1}^T\right) \\
& -(b+1)A_{j-1}^T \\
=& \left((T-1)b - (b+1)j\right)(A_j^T - A_{j-1}^T), \tag{142}
\end{aligned}
$$

and $A_j^T \overline{G}_{j-2}^{T-1}(A_{j-1}^T) - A_{j-1}^T \overline{G}_{j-1}^{T-1}(A_j^T)$ equals

$$
\begin{aligned}
& (T-1)\left(A_j^T B_{j-1}^{T-1} - A_{j-1}^T B_j^{T-1}\right) + (b+1)A_{j-1}^T A_j^T \\
=& (T-1)\left(A_j^T A_{j-1}^T \frac{j-1}{T-1} - A_{j-1}^T A_j^T \frac{j}{T-1}\right) + (b+1)A_{j-1}^T A_j^T \\
=& bA_{j-1}^T A_j^T \quad = \quad jA_j^T\left(A_j^T - A_{j-1}^T\right). \tag{143}
\end{aligned}
$$

Plugging (142) and (143) into (141), and comparing to (140), completes the proof. ∎

Finally, as it is easily verified that Theorem 18.(V) follows from definitions, combining the above completes the proof of Theorem 18.

We now complete the proof of Corollaries 5 - 6. In light of Theorems 14 - 18, and the fundamental inventory dynamics underlying our model, note that it suffices to prove the following.

Lemma 28 *For all $U, b \in \mathbb{R}^+, T \geq 1, \mu \in [0, U]$, and $t \in \{1, \ldots, \Lambda^T\}$,*

(I) $\chi_{MAR}^{T+1-t}\left(D_{t-1}^T, U, b\right) = X_t^T = B_{\Gamma_\mu^T}^{T+1-t}$;

(II) $X_t^T \geq B_{\Gamma_{D_{t-1}^T}^{T-t}}^{T+1-t}$;

(III) $D_t^T = A_{\Gamma_\mu^T}^{T+1-t}$;

(IV) $D_t^T \geq X_t^T$.

Proof. The results clearly hold true when $\mu = 0$ from definitions. Suppose $\mu > 0$. We proceed by induction on t, beginning with the case $t = 1$, and proving (I) - (IV) in order. (I) follows from definitions, which guarantee that $\mu \in \left(A_{\Gamma_\mu^T - 1}^{T+1}, A_{\Gamma_\mu^T}^{T+1}\right]$, and thus $\chi_{MAR}^T(\mu, U, b) = B_{\Gamma_\mu^T}^T$. (II) follows from (I), and the fact that $\Gamma_\mu^T \geq \Gamma_\mu^{T-1}$, which itself follows from the fact that A_j^τ is decreasing in τ. (III) follows from definitions. Furthermore, (IV) follows from the fact that $A_j^\tau \geq B_j^\tau$ for all τ, j.

Now, suppose the induction is true for all $k = 1, \ldots, t$ for some t, and let us prove the induction holds for $k = t + 1$, beginning with the proof of (I). Noting that $D_t^T = A_{\Gamma_\mu^T}^{T+1-t}$, it follows from definitions that $\chi_{MAR}^{T-t}\left(D_t^T, U, b\right) = B_{\Gamma_\mu^T}^{T-t} = X_{t+1}^T$. We next prove (II). From definitions, $\Gamma_{D_t^T}^{T-t} = \Gamma_\mu^T$. The fact that A_j^τ is decreasing in τ then ensures that $\Gamma_{D_{t-1}^T}^{T-t} \leq \Gamma_\mu^T$, and the desired result then follows from the monotonicity of B_j^τ in j. (III) again follows from definitions, and (IV) again follows from the fact that $A_j^\tau \geq B_j^\tau$ for all τ, j. Combining the above completes the proof. ∎

5.4 Asymptotic analysis

In this section, we complete the proof of remaining results, Theorems 15, 16, 17, and Corollary 7.

5.4.1 Proof of Theorem 15

We prove Theorem 15 in this section. We begin by proving the following asymptotic results.

Lemma 29 *For any $\mu \in [0, U]$, $\alpha \geq 0$, $\beta \in \{0, 1\}$, define $m(T) \triangleq \min\{\lfloor \alpha T \rfloor, \Lambda^T - 1\} + \beta$. Then*

$$\lim_{T \to \infty} \left(\Gamma_\mu^T / T \right) = \gamma^{b^{-1}}, \qquad \lim_{T \to \infty} A_{\Gamma_\mu^T}^{T+1} = \mu, \qquad \lim_{T \to \infty} \left(\Lambda^T / T \right) = \Lambda^\infty; \qquad (144)$$

$$\lim_{T \to \infty} D_{m(T)}^T = D_{\min\{\alpha, \Lambda^\infty\}}^\infty, \lim_{T \to \infty} D_{\lfloor \alpha T \rfloor + \beta}^T = D_\alpha^\infty, \lim_{T \to \infty} D_{\Lambda^T - 1 + \beta}^T = D_{\Lambda^\infty}^\infty; \qquad (145)$$

$$\lim_{T \to \infty} X_{m(T)}^T = X_{\min\{\alpha, \Lambda^\infty\}}^\infty, \lim_{T \to \infty} X_{\lfloor \alpha T \rfloor + \beta}^T = X_\alpha^\infty, \lim_{T \to \infty} X_{\Lambda^T - 1 + \beta}^T = X_{\Lambda^\infty}^\infty. \qquad (146)$$

Proof. Let us first prove the three equalities in (144). They clearly hold if $\mu = 0$. Suppose $\mu > 0$ and let $i = \Gamma_\mu^T$. From the definition, $A_{i-1}^{T+1} < \mu \leq A_i^{T+1}$. Noting that $1 - x \leq \exp(-x)$ for all $x \geq 0$, we conclude that $U^{-1} A_i^{T+1}$ equals

$$\prod_{k=i+1}^{T} \frac{k}{b+k} \leq \prod_{k=i+1}^{T} \exp\left(-\frac{b}{b+k} \right)$$

$$= \exp\left(-b \sum_{k=i+1}^{T} \frac{1}{b+k} \right) \qquad (147)$$

$$\leq \exp\left(-b \int_{i+1}^{T+1} \frac{1}{b+x} dx \right) = \left(\frac{b+i+1}{b+T+1} \right)^b.$$

In addition, from the definition, Γ_μ^T is non-decreasing in T and $\lim_{T \to \infty} \Gamma_\mu^T = \infty$ if $\mu > 0$. Noting that there exists $\epsilon > 0$ such that for all $x \in (0, \epsilon)$, $1 - x \geq \exp\left(-x - x^{\frac{3}{2}} \right)$, we further conclude that for large T, $U^{-1} A_{i-1}^{T+1}$ equals

$$\prod_{k=i}^{T} \frac{k}{b+k} \geq \prod_{k=i}^{T} \exp\left(-\frac{b}{b+k} - \left(\frac{b}{b+k} \right)^{\frac{3}{2}} \right)$$

$$= \exp\left(-b \sum_{k=i}^{T} \frac{1}{b+k} \right) \exp\left(-\sum_{k=i}^{T} \left(\frac{b}{b+k} \right)^{\frac{3}{2}} \right)$$

$$\geq \exp\left(-b \int_{i-1}^{T} \frac{1}{b+x} dx \right) \exp\left(-\sum_{k=i}^{T} \left(\frac{b}{b+k} \right)^{\frac{3}{2}} \right) \qquad (148)$$

$$= \left(\frac{b+i-1}{b+T} \right)^b \exp\left(-\sum_{k=i}^{T} \left(\frac{b}{b+k} \right)^{\frac{3}{2}} \right).$$

154

(147) and (148) together imply that

$$\limsup_{T\to\infty}\left[\left(\frac{b+i-1}{b+T}\right)^b\exp\left(-\sum_{k=i}^{T}\left(\frac{b}{b+k}\right)^{\frac{3}{2}}\right)\right]\leq\ \gamma\ \leq\ \liminf_{T\to\infty}\left(\frac{b+i+1}{b+T+1}\right)^b.$$

(149)

It follows from the fact that the series $\sum_{k=1}^{\infty}\left(\frac{b}{b+k}\right)^{\frac{3}{2}}$ converges,

$$\lim_{T\to\infty}\exp\left(-\sum_{k=i}^{T}\left(\frac{b}{b+k}\right)^{\frac{3}{2}}\right)=1.$$

(150)

(149) and (150) together prove the three equalities in (144). Let us prove the first equality in (145). If $\alpha=\beta=0$, then $m(T)=0$ for all T such that the equality clearly holds from the definition. Otherwise, note that

$$D_{m(T)}^{T}=\ A_i^{T+1-m(T)}=\ A_i^{T+1}\prod_{k=T+1-m(T)}^{T}\frac{b+k}{k}.$$

(151)

By using a similar argument as we did in proving upper and lower bounds in (147) and (148), one can show that

$$\lim_{T\to\infty}\prod_{k=T+1-m(T)}^{T}\frac{b+k}{k}=\ \lim_{T\to\infty}\left(\frac{T}{T-m(T)}\right)^b=\ \left(\frac{1}{1-\min\{\alpha,\Lambda^\infty\}}\right)^b.$$

(152)

Combining (151), (152) with (144) completes the proof of the first equality in (145). The second and third ones are special cases of the first one when $\alpha<\Lambda^\infty$ and $\alpha>\Lambda^\infty$ respectively. Finally, combining (145) with the following fact

$$X_{m(T)}^{T}=\ \frac{i}{b+T+1-m(T)}D_{m(T)}^{T},$$

we complete the proofs of the three equalities in (146). ∎

Let $\mathtt{f}(z)\triangleq b\left(1-z\right)^{b-1}$ and it is not difficult to verify that for any $\alpha\in[0,\Lambda^\infty)$,

$$\mathbb{P}\left(Z^\infty\leq\alpha\right)=\ \int_0^\alpha\mathtt{f}(z)dz=\ 1-(1-\alpha)^b.$$

(153)

For any $\alpha_1,\alpha_2\geq 0$ such that $\alpha_1<\min\{\alpha_2,\Lambda^\infty\}$, define

$$X_{\alpha_1,\alpha_2}^{T}\triangleq\ X_{Z^T}^{T}\mathbb{I}\left(\lfloor\alpha_1 T\rfloor+1\leq Z^T\leq\min\{\lfloor\alpha_2 T\rfloor,\Lambda^T-1\}\right),$$

$$X_{\alpha_1,\alpha_2}^{\infty}\triangleq\ X_{Z^\infty}^{\infty}\mathbb{I}\left(\alpha_1\leq Z^\infty<\min\{\alpha_2,\Lambda^\infty\}\right).$$

155

To prove the weak convergence of processes, let us first prove the following weak convergence of $X^T_{\alpha_1,\alpha_2}$ to $X^\infty_{\alpha_1,\alpha_2}$.

Lemma 30 *Assume $\mu \in (0,U)$. For any $\alpha_1, \alpha_2 \geq 0$ such that $\alpha_1 < \min\{\alpha_2, \Lambda^\infty\}$, $X^T_{\alpha_1,\alpha_2}$ converges weakly to $X^\infty_{\alpha_1,\alpha_2}$. As a consequence, for any bounded uniformly continuous function $H : [0,U] \to \mathbb{R}$,*

$$\lim_{T\to\infty} \sum_{k=\lfloor \alpha_1 T \rfloor + 1}^{\min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\}} H\left(X^T_k\right) \mathbb{P}(Z^T = k) = \int_{\alpha_1}^{\min\{\alpha_2,\Lambda^\infty\}} H\left(X^\infty_z\right) \mathbf{f}(z) dz. \qquad (154)$$

Proof. Let us first prove the weak convergence of $X^T_{\alpha_1,\alpha_2}$. From the definition of weak convergence, it suffices to show that for any $x \in \mathbb{R}$,

$$\lim_{T\to\infty} \mathbb{P}\left(X^T_{\alpha_1,\alpha_2} \leq x\right) = \mathbb{P}\left(X^\infty_{\alpha_1,\alpha_2} \leq x\right). \qquad (155)$$

Note that if $x \in \left[X^\infty_{\alpha_1}, \ X^\infty_{\min\{\alpha_2,\Lambda^\infty\}}\right]$, there exists a unique $\alpha_x \in [\alpha_1, \ \min\{\alpha_2, \Lambda^\infty\}]$ such that $X^\infty_{\alpha_x} = x$ due to the strict monotonicity of X^∞_α in α on $[0, \Lambda^\infty]$. Combining above with (153) implies that

$$\mathbb{P}\left(X^\infty_{\alpha_1,\alpha_2} \leq x\right) = \begin{cases} 1 - (1-\alpha_1)^b + (1 - \min\{\alpha_2, \Lambda^\infty\})^b & \text{if } 0 \leq x < X^\infty_{\alpha_1}; \\ 1 - (1-\alpha_x)^b + (1 - \min\{\alpha_2, \Lambda^\infty\})^b & \text{if } x \in \left[X^\infty_{\alpha_1}, \ X^\infty_{\min\{\alpha_2,\Lambda^\infty\}}\right]; \\ 1 & \text{if } x > X^\infty_{\min\{\alpha_2,\Lambda^\infty\}}. \end{cases}$$

$$\qquad (156)$$

Let us proceed to prove (155) by a case analysis. Due to the nonnegativity of $X^T_{\alpha_1,\alpha_2}$ and $X^\infty_{\alpha_1,\alpha_2}$, (155) clearly holds if $x < 0$. If $x > X^\infty_{\min\{\alpha_2,\Lambda^\infty\}}$, from Lemma 29, $X^T_{\min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\}}$ converges to $X^\infty_{\min\{\alpha_2,\Lambda^\infty\}}$ as $T \to \infty$ such that $\mathbb{P}\left(X^T_{\alpha_1,\alpha_2} \leq x\right) = 1$ for all sufficiently large T. Combining above with (156) proves (155) in this case. If $0 \leq x < X^\infty_{\alpha_1}$, from Lemma 29, $X^T_{\lfloor \alpha_1 T \rfloor}$ converges to $X^\infty_{\alpha_1}$ as T grows such that for all sufficiently large T,

$$\mathbb{P}\left(X^T_{\alpha_1,\alpha_2} \leq x\right) = \mathbb{P}\left(Z^T < \lfloor \alpha_1 T \rfloor + 1 \text{ or } Z^T > \min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\}\right)$$

$$= 1 - \frac{\mu}{D^T_{\lfloor \alpha_1 T \rfloor}} + \frac{\mu}{D^T_{\min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\}}}.$$

156

Combining above with Lemma 29 and (156) proves (155) in this case. If $x \in \left(X_{\alpha_1}^\infty, X_{\min\{\alpha_2, \Lambda^\infty\}}^\infty \right)$, for any sufficiently small $\epsilon > 0$ s.t. $\alpha_x \pm \epsilon \in \left(\alpha_1, \min\{\alpha_2, \Lambda^\infty\} \right)$, let $\eta_1 \triangleq X_{\alpha_x + \epsilon}^\infty - X_{\alpha_x}^\infty$, $\eta_2 \triangleq X_{\alpha_x}^\infty - X_{\alpha_x - \epsilon}^\infty$. Note that η_1, η_2 are strictly positive due to the fact that X_α^∞ is strictly increasing in α on $[0, \Lambda^\infty]$. Claim that there exist $T_1, T_2 > 0$ such that for all $T \geq \max\{T_1, T_2\}$,

$$\mathbb{P}\left(X_{\alpha_1, \alpha_2}^T \leq x \right) \leq \mathbb{P}\left(Z^T \leq \lfloor (\alpha_x + \epsilon)T \rfloor \text{ or } Z^T > \min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\} \right), \quad (157)$$

$$\mathbb{P}\left(X_{\alpha_1, \alpha_2}^T \leq x \right) \geq \mathbb{P}\left(Z^T \leq \lfloor (\alpha_x - \epsilon)T \rfloor \text{ or } Z^T > \min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\} \right). \quad (158)$$

Indeed, if $\lfloor (\alpha_x + \epsilon)T \rfloor + 1 \leq Z^T \leq \min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\}$, then $X_{\alpha_1, \alpha_2}^T = X_{Z^T}^T \geq X_{\lfloor (\alpha_x + \epsilon)T \rfloor + 1}^T$ from the monotonicity of X_j^T in j. From Lemma 29, $\lim_{T \to \infty} X_{\lfloor (\alpha_x + \epsilon)T \rfloor + 1}^T = X_{\alpha_x + \epsilon}^\infty$, which implies that there exists $T_1 > 0$ such that $\left| X_{\lfloor (\alpha_x + \epsilon)T \rfloor + 1}^T - X_{\alpha_x + \epsilon}^\infty \right| < \eta_1$ for all $T \geq T_1$. It follows that

$$
\begin{aligned}
X_{\alpha_1, \alpha_2}^T &\geq X_{\lfloor (\alpha_x + \epsilon)T \rfloor + 1}^T \\
&> X_{\alpha_x + \epsilon}^\infty - \eta_1 = X_{\alpha_x}^\infty = x,
\end{aligned}
$$

which proves (157). Similarly, one may prove that there exists $T_2 > 0$ such that (158) holds for all $T \geq T_2$. As a consequence, from Lemma 29,

$$
\begin{aligned}
\limsup_{T \to \infty} \mathbb{P}\left(X_{\alpha_1, \alpha_2}^T \leq x \right) &\leq \limsup_{T \to \infty} \mathbb{P}\left(Z^T \leq \lfloor (\alpha_x + \epsilon)T \rfloor \text{ or } Z^T > \min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\} \right) \\
&= 1 - \frac{\mu}{D_{\alpha_x + \epsilon}^\infty} + \frac{\mu}{D_{\min\{\alpha_2, \Lambda^\infty\}}^\infty}, \quad (159)
\end{aligned}
$$

$$
\begin{aligned}
\liminf_{T \to \infty} \mathbb{P}\left(X_{\alpha_1, \alpha_2}^T \leq x \right) &\geq \liminf_{T \to \infty} \mathbb{P}\left(Z^T \leq \lfloor (\alpha_x - \epsilon)T \rfloor \text{ or } Z^T > \min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\} \right) \\
&= 1 - \frac{\mu}{D_{\alpha_x - \epsilon}^\infty} + \frac{\mu}{D_{\min\{\alpha_2, \Lambda^\infty\}}^\infty}. \quad (160)
\end{aligned}
$$

Since ϵ is arbitrary and D_α^∞ is continuous in α, we conclude that

$$\lim_{T \to \infty} \mathbb{P}\left(X_{\alpha_1, \alpha_2}^T \leq x \right) = 1 - \frac{\mu}{D_{\alpha_x}^\infty} + \frac{\mu}{D_{\min\{\alpha_2, \Lambda^\infty\}}^\infty}.$$

157

Combining above with (156) completes the proof of (155) in the case. If $x = X_{\alpha_1}^\infty$, we note that (157) and (159) still hold. Combining with the fact that

$$\liminf_{T\to\infty} \mathbb{P}\left(X_{\alpha_1,\alpha_2}^T \le x\right) \ge \liminf_{T\to\infty} \mathbb{P}\left(Z^T \le \lfloor \alpha_1 T \rfloor \text{ or } Z^T > \min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\}\right)$$

$$= 1 - \frac{\mu}{D_{\alpha_1}^\infty} + \frac{\mu}{D_{\min\{\alpha_2,\Lambda^\infty\}}^\infty},$$

and letting $\epsilon \to 0$ in (159) complete the proof of (155) in this case. Finally, if $x = X_{\min\{\alpha_2,\Lambda^\infty\}}^\infty$, similarly, (158) and (160) still hold. Combining with the fact that $\mathbb{P}\left(X_{\alpha_1,\alpha_2}^T \le x\right) \le 1$ and letting $\epsilon \to 0$ in (160) complete the proof of (155) in this case, which completes the proof of the weak convergence of X_{α_1,α_2}^T. For any bounded uniformly continuous function H, from Portmanteau theorem (cf. Theorem 2.1 in [25]), $\mathbb{E}\left[H(X_{\alpha_1,\alpha_2}^T)\right]$ converges to $\mathbb{E}\left[H(X_{\alpha_1,\alpha_2}^\infty)\right]$ as T grows. Note that $\mathbb{E}\left[H(X_{\alpha_1,\alpha_2}^T)\right]$ equals

$$\sum_{k=\lfloor \alpha_1 T \rfloor +1}^{\min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\}} H\left(X_k^T\right) \mathbb{P}(Z^T = k) + \mathbb{P}\left(Z^T \le \lfloor \alpha_1 T \rfloor \text{ or } Z^T > \min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\}\right) H(0),$$

and

$$\mathbb{E}\left[H(X_{\alpha_1,\alpha_2}^\infty)\right] = \int_{\alpha_1}^{\min\{\alpha_2,\Lambda^\infty\}} H\left(X_z^\infty\right) \mathtt{f}(z)dz + \mathbb{P}\left(Z^\infty < \alpha_1 \text{ or } Z^\infty \ge \min\{\alpha_2,\Lambda^\infty\}\right) H(0).$$

In addition, from Lemma 29,

$$\lim_{T\to\infty} \mathbb{P}\left(Z^T \le \lfloor \alpha_1 T \rfloor \text{ or } Z^T > \min\{\lfloor \alpha_2 T \rfloor, \Lambda^T - 1\}\right) = \mathbb{P}\left(Z^\infty < \alpha_1 \text{ or } Z^\infty \ge \min\{\alpha_2,\Lambda^\infty\}\right).$$

Combining all of the above completes the proof of (154). ∎

We now complete the proof of Theorem 15, and begin by providing some additional background about the space $D\left([0,1],\ \mathbb{R}^k\right)$. Recall that the space $D\left([0,1],\ \mathbb{R}^k\right)$ contains all functions $x : [0,1] \to \mathbb{R}^k$ that are right-continuous in $[0,1)$ and have left-limits in $(0,1]$. Let Λ denote the class of strictly increasing, continuous mappings λ of $[0,1]$ onto itself such that $\lambda(0) = 0$, $\lambda(1) = 1$. The following metric defines the Skorohod J_1 topology:

$$d(x,y) \triangleq \sum_{j=1}^{k} \left(\inf_{\lambda \in \Lambda} \{||\lambda - I|| \vee ||x_j - y_j \circ \lambda||\} \right),$$

158

where $x \triangleq (x_1, \ldots, x_k)$, $||x_j|| \triangleq \sup_{t \in [0,1]} |x_j(t)|$, I is the identity map, " $\circ$ " denotes the composition of functions, and $a \vee b \triangleq \max\{a, b\}$. Furthermore, the norm in $D\left([0,1], \mathbb{R}^k\right)$ can be defined as $||x|| \triangleq \sum_{j=1}^k ||x_j||$. More details can be found in Section 3.3, [198].

In general, in order to prove the weak convergence of processes, it is sufficient to prove the tightness of the processes and the finite-dimensional weak convergence (cf. Theorem 13.1 in [25]). Assume a set $\{\alpha_i\}_{i=0}^m$ satisfies $0 = \alpha_0 < \alpha_1 < \ldots < \alpha_m = 1$. We call $\{\alpha_i\}_{i=0}^m$ δ-sparse if it satisfies $\min_{1 \leq i \leq m}(\alpha_i - \alpha_{i-1}) \geq \delta$. Define

$$w_x(\delta) \triangleq \inf_{\{\alpha_i\}} \max_{1 \leq i \leq m} \left(\sum_{j=1}^k \sup_{\alpha, \alpha' \in [\alpha_{i-1}, \alpha_i)} |x_j(\alpha) - x_j(\alpha')| \right),$$

where the infimum extends over all δ-sparse sets $\{\alpha_i\}$. We revisit the following necessary and sufficient conditions for tightness in the space $D\left([0,1], \mathbb{R}^k\right)$.

Lemma 31 (Theorem 13.2, [25]) *A sequence of probability measures $\{P_T\}_{T \geq 1}$ on the space $D\left([0,1], \mathbb{R}^k\right)$ under the J_1 topology is tight if and only if the following two conditions hold:*

1.

$$\lim_{a \to \infty} \limsup_T P_T \left(x : ||x|| \geq a \right) = 0; \tag{161}$$

2. for each ϵ,

$$\lim_{\delta \to 0} \limsup_T P_T \left(x : w_x(\delta) \geq \epsilon \right) = 0. \tag{162}$$

We show the tightness of the processes $\{\mathcal{M}^T(\alpha)_{0 \leq \alpha \leq 1}, T \geq 1\}$ in the following lemma by proving (161) and (162).

Lemma 32 *Assume $\mu \in (0, U)$. $\{\mathcal{M}^T(\alpha)_{0 \leq \alpha \leq 1}, T \geq 1\}$ is tight in $D\left([0,1], \mathbb{R}^2\right)$.*

Proof. (161) follows from the fact that $\mathcal{M}^T(\alpha)$ is supported on $[0, U]^2$ w.p.1 for all $\alpha \in [0,1]$, $T \geq 1$. Let us prove (162). Note that $w_x(\delta)$ is increasing in δ for any x

such that $\limsup_T P_T (x : w_x(\delta) \geq \epsilon)$ is increasing in δ for each ϵ. It suffices to prove that there exists a sequence $\{\delta_n\}_{n=1}^\infty$ that converges to zero such that

$$\lim_{n \to \infty} \limsup_T P_T (x : w_x(\delta_n) \geq \epsilon) = 0.$$

Define $\delta_n \triangleq \frac{\Lambda^\infty}{n+\frac{1}{2}}$. It follows that $n\delta_n < \Lambda^\infty < (n+1)\delta_n$ for each n. From Lemma 29, there exists $T_1(\delta_n)$ such that for all $T \geq T_1(\delta_n)$,

$$\lfloor n\delta_n T \rfloor < \Lambda^T - 1 < \Lambda^T < \lfloor (n+1)\delta_n T \rfloor. \tag{163}$$

There exist $m \triangleq m(\delta_n) \in \mathbb{Z}_+$, $r \triangleq r(\delta_n) \in [0, \delta_n)$ such that $1 = m\delta_n + r$. Note that $\mu \in (0, U)$ implies $\Lambda^\infty \in (0, 1)$. It follows that there exists n_1 such that for all $n \geq n_1$, $(n+2)\delta_n < \Lambda^\infty + 2\delta_n < 1$, which implies $n + 2 \leq m$. Combining above with (163) implies that for all $n \geq n_1$, $T \geq T_1(\delta_n)$, $\Lambda^T/T \leq (m-1)\delta_n$. If $Z^T/T \in \{j\delta_n\}_{j=0}^{m-1}$, then choose the following δ_n-sparse set: $\{1\} \bigcup \{j\delta_n\}_{j=0}^{m-1}$. Otherwise, there exists k such that $k\delta_n < Z^T/T < (k+1)\delta_n$. In that case, we choose the following δ_n-sparse set: $\{Z^T/T, 1\} \bigcup \{j\delta_n\}_{j=0}^{k-1} \bigcup \{j\delta_n\}_{j=k+2}^{m-1}$. Such selections above ensure that Z^T/T is in the δ_n-sparse set. In addition, for any $\alpha, \alpha' \in [\alpha_{i-1}, \alpha_i)$, where α_{i-1}, α_i are in the δ_n-sparse set, $|\alpha - \alpha'| \leq 2\delta$. We treat two cases. If $Z^T \leq \Lambda^T - 1$, from Corollary 6,

$$\mathcal{M}^T(\alpha) = \begin{cases} \left(X^T_{\lfloor \alpha T \rfloor}, D^T_{\lfloor \alpha T \rfloor}\right) & \text{if } \lfloor \alpha T \rfloor \leq Z^T - 1; \\ \left(X^T_{Z^T}, 0\right) & \text{if } \lfloor \alpha T \rfloor \geq Z^T. \end{cases}$$

Combining (163) with the fact that $\mathcal{M}^T(\alpha)$ is a constant if $\lfloor \alpha T \rfloor \geq Z^T$, we conclude that for all $n \geq n_1$, $T \geq T_1(\delta_n)$, $w_{\mathcal{M}^T}(\delta_n) \leq \overline{M}$, where

$$\overline{M} \triangleq \max \left\{ \max_{j=0,\ldots,n-2} \left(\left| X^T_{\lfloor (j+2)\delta_n T \rfloor} - X^T_{\lfloor j\delta_n T \rfloor} \right| + \left| D^T_{\lfloor (j+2)\delta_n T \rfloor} - D^T_{\lfloor j\delta_n T \rfloor} \right| \right), \right.$$

$$\left. \left(\left| U - X^T_{\lfloor (n-1)\delta_n T \rfloor} \right| + \left| U - D^T_{\lfloor (n-1)\delta_n T \rfloor} \right| \right) \right\}.$$

If $Z^T = \Lambda^T$ and $Y^T = U$, from Corollary 6,

$$
\mathcal{M}^T(\alpha) = \begin{cases} \left(X^T_{\lfloor \alpha T \rfloor}, D^T_{\lfloor \alpha T \rfloor} \right) & \text{if } \lfloor \alpha T \rfloor \leq \Lambda^T - 1; \\[2mm] \left(X^T_{\Lambda^T}, U \right) & \text{if } \lfloor \alpha T \rfloor = \Lambda^T; \\[2mm] \left(U, U \right) & \text{if } \lfloor \alpha T \rfloor \geq \Lambda^T + 1. \end{cases}
$$

Following from a similar argument, we conclude that for all $n \geq n_1$, $T \geq T_1(\delta_n)$, $w_{\mathcal{M}^T}(\delta_n) \leq \max\left\{ \overline{M}, \ \left| U - X^T_{\Lambda^T} \right| \right\}$. Combining Lemma 29 with the uniform continuity of X^∞_α and D^∞_α on $[0, \Lambda^\infty]$, we conclude that for any $\epsilon > 0$, there exists $n_2 > 0$ such that for all $n \geq n_2$, there exists $T_2(\delta_n) > 0$ such that for all $T \geq T_2(\delta_n)$, $w_{\mathcal{M}^T}(\delta_n) < \epsilon$ in both cases. As a consequence, for all $n \geq n_2$,

$$
\limsup_{T \to \infty} P_T \left(x : w_x(\delta_n) \geq \epsilon \right) \leq \limsup_{T \to \infty} \mathbb{P} \left(Y^T = 0 \right)
$$
$$
= \limsup_{T \to \infty} \left(1 - \frac{D^T_{\Lambda^T - 1}}{U} \right) = 0,
$$

where the last equality is from Lemma 29. We complete the proof. ∎

We prove the following finite-dimensional weak convergence.

Lemma 33 *Assume $\mu \in (0, U)$. For any $0 \leq \alpha_1 < \ldots < \alpha_n \leq 1$, $\left(\mathcal{M}^T(\alpha_1), \ldots, \mathcal{M}^T(\alpha_n) \right)$ converges weakly to $(\mathcal{M}^\infty(\alpha_1), \ldots, \mathcal{M}^\infty(\alpha_n))$.*

Proof. Since $\mu \in (0, U)$, we have $0 < \Lambda^\infty < 1$. Define $\alpha_0 \overset{\Delta}{=} 0$ if $\alpha_1 > 0$ and $\alpha_{n+1} \overset{\Delta}{=} 1$ if $\alpha_n < 1$. Then there exists $k_0 \in \{0, \ldots, n\}$ such that $\alpha_{k_0} \leq \Lambda^\infty < \alpha_{k_0+1}$. From Portmanteau theorem (cf. Theorem 2.1 in [25]), it suffices to show that for any bounded uniformly continuous function $H : [0, U]^{2n} \to \mathbb{R}$,

$$
\lim_{T \to \infty} \mathbb{E} \left[H \left(\mathcal{M}^T(\alpha_1), \ldots, \mathcal{M}^T(\alpha_n) \right) \right] = \mathbb{E} \left[H \left(\mathcal{M}^\infty(\alpha_1), \ldots, \mathcal{M}^\infty(\alpha_n) \right) \right]. \tag{164}
$$

As a notational convenience, we define for $i \geq 0$, $j \geq 1$,

$$
\begin{aligned}
H^{\infty}\left(i, X_z^{\infty}, 0\right) &\triangleq H\left(X_{\alpha_1}^{\infty}, D_{\alpha_1}^{\infty}, \ldots, X_{\alpha_i}^{\infty}, D_{\alpha_i}^{\infty}, X_z^{\infty}, 0, \ldots, X_z^{\infty}, 0\right), \\
H^{\infty}\left(i, U, U\right) &\triangleq H\left(X_{\alpha_1}^{\infty}, D_{\alpha_1}^{\infty}, \ldots, X_{\alpha_i}^{\infty}, D_{\alpha_i}^{\infty}, U, U, \ldots, U, U\right), \\
H^T\left(i, X_j^T, 0\right) &\triangleq H\left(X_{\lfloor \alpha_1 T \rfloor}^T, D_{\lfloor \alpha_1 T \rfloor}^T, \ldots, X_{\lfloor \alpha_i T \rfloor}^T, D_{\lfloor \alpha_i T \rfloor}^T, X_j^T, 0, \ldots, X_j^T, 0\right), \\
H^T\left(i, U, U\right) &\triangleq H\left(X_{\lfloor \alpha_1 T \rfloor}^T, D_{\lfloor \alpha_1 T \rfloor}^T, \ldots, X_{\lfloor \alpha_i T \rfloor}^T, D_{\lfloor \alpha_i T \rfloor}^T, U, U, \ldots, U, U\right), \\
H^T\left(i, X_{\Lambda^T}^T, U\right) &\triangleq H\left(X_{\lfloor \alpha_1 T \rfloor}^T, D_{\lfloor \alpha_1 T \rfloor}^T, \ldots, X_{\lfloor \alpha_i T \rfloor}^T, D_{\lfloor \alpha_i T \rfloor}^T, X_{\Lambda^T}^T, U, U, \ldots, U, U\right).
\end{aligned}
$$

From the definition of $\mathcal{M}^{\infty}(\alpha)$,

$$
\mathbb{E}\left[H\left(\mathcal{M}^{\infty}(\alpha_1), \ldots, \mathcal{M}^{\infty}(\alpha_n)\right)\right] = W_1^{\infty} + W_2^{\infty} + W_3^{\infty},
$$

where

$$
\begin{aligned}
W_1^{\infty} &\triangleq \sum_{i=0}^{k_0-1} \int_{\alpha_i}^{\alpha_{i+1}} H^{\infty}\left(i, X_z^{\infty}, 0\right) \mathtt{f}(z) dz, \\
W_2^{\infty} &\triangleq \int_{\alpha_{k_0}}^{\Lambda^{\infty}} H^{\infty}\left(k_0, X_z^{\infty}, 0\right) \mathtt{f}(z) dz, \\
W_3^{\infty} &\triangleq \gamma H^{\infty}\left(k_0, U, U\right).
\end{aligned}
$$

In addition, from Lemma 29 and the fact that $\alpha_{k_0} \leq \Lambda^{\infty} < \alpha_{k_0+1}$, we have $\lfloor \alpha_{k_0-1} T \rfloor < \Lambda^T - 1$ and $\Lambda^T < \lfloor \alpha_{k_0+1} T \rfloor$ when T is sufficiently large. From Corollary 6,

$$
\mathbb{E}\left[H\left(\mathcal{M}^T(\alpha_1), \ldots, \mathcal{M}^T(\alpha_n)\right)\right] = W_1^T + W_2^T + W_3^T + W_4^T,
$$

where

$$W_1^T \triangleq \sum_{i=0}^{k_0-2} \sum_{j=\lfloor \alpha_i T \rfloor + 1}^{\lfloor \alpha_{i+1} T \rfloor} H^T\left(i, X_j^T, 0\right) \mathbb{P}(Z^T = j)$$

$$+ \sum_{j=\lfloor \alpha_{k_0-1} T \rfloor + 1}^{\min\{\lfloor \alpha_{k_0} T \rfloor, \Lambda^T - 1\}} H^T\left(k_0 - 1, X_j^T, 0\right) \mathbb{P}(Z^T = j),$$

$$W_2^T \triangleq \mathbb{I}\left(\lfloor \alpha_{k_0} T \rfloor < \Lambda^T - 1\right) \sum_{j=\lfloor \alpha_{k_0} T \rfloor + 1}^{\Lambda^T - 1} H^T\left(k_0, X_j^T, 0\right) \mathbb{P}(Z^T = j),$$

$$W_3^T \triangleq \Bigg[\mathbb{I}\left(\lfloor \alpha_{k_0} T \rfloor = \Lambda^T\right) H^T\left(k_0 - 1, X_{\Lambda^T}^T, U\right) + \mathbb{I}\left(\lfloor \alpha_{k_0} T \rfloor > \Lambda^T\right) H^T\left(k_0 - 1, U, U\right)$$

$$+ \mathbb{I}\left(\lfloor \alpha_{k_0} T \rfloor < \Lambda^T\right) H^T\left(k_0, U, U\right) \Bigg] \mathbb{P}\left(Z^T = \Lambda^T\right) \mathbb{P}\left(Y^T = U\right),$$

$$W_4^T \triangleq \Bigg[\mathbb{I}\left(\lfloor \alpha_{k_0} T \rfloor \geq \Lambda^T\right) H^T\left(k_0 - 1, X_{\Lambda^T}^T, 0\right)$$

$$+ \mathbb{I}\left(\lfloor \alpha_{k_0} T \rfloor < \Lambda^T\right) H^T\left(k_0, X_{\Lambda^T}^T, 0\right) \Bigg] \mathbb{P}\left(Z^T = \Lambda^T\right) \mathbb{P}\left(Y^T = 0\right),$$

and $\mathbb{I}(.)$ denotes the indicator function. Note that $H^T\left(i, X_j^T, 0\right)$ can be decomposed as the sum of

$$H^T\left(i, X_j^T, 0\right) - H\left(X_{\alpha_1}^\infty, D_{\alpha_1}^\infty, \ldots, X_{\alpha_i}^\infty, D_{\alpha_i}^\infty, X_j^T, 0, \ldots, X_j^T, 0\right) \tag{165}$$

and

$$H\left(X_{\alpha_1}^\infty, D_{\alpha_1}^\infty, \ldots, X_{\alpha_i}^\infty, D_{\alpha_i}^\infty, X_j^T, 0, \ldots, X_j^T, 0\right).$$

From Lemma 29 and the uniform continuity of H, for each i, (165) converges to zero uniformly in j as $T \to \infty$. Combining above with the boundedness of H, Lemmas 29 and 30, we conclude that as $T \to \infty$,

$$W_1^T \to W_1^\infty, \quad W_2^T \to W_2^\infty, \quad W_3^T \to W_3^\infty, \quad W_4^T \to 0,$$

which completes the proof of (164). ∎

Lemmas 32 and 33 together imply Theorem 15.

5.4.2 Proofs of Theorems 16, 17 and Corollary 7

In this section, we prove Theorems 16, 17 and Corollary 7. Note that Corollary 7 immediately follows from Theorem 17.

Proof. [Proof of Theorem 16] Define two admissible policies as below:

$$\pi_1: \quad x_1^{\pi_1} = 0, \ x_t^{\pi_1}(D_{[t-1]}) = 0 \text{ for all } D_{[t-1]}, \ t = 2, \ldots, T,$$

$$\pi_2: \quad x_1^{\pi_2} = U, \ x_t^{\pi_2}(D_{[t-1]}) = U \text{ for all } D_{[t-1]}, \ t = 2, \ldots, T,$$

i.e., π_1 is the policy always ordering up to 0 and π_2 is the one always ordering up to U. Note that

$$\max_{Q \in \mathcal{M}} \mathbb{E}_Q\left[\sum_{t=1}^{T} C_t^{\pi_1}\right] = \max_{Q \in \mathcal{M}} \mathbb{E}_Q\left[\sum_{t=1}^{T} b D_t\right] = Tb\mu,$$

$$\max_{Q \in \mathcal{M}} \mathbb{E}_Q\left[\sum_{t=1}^{T} C_t^{\pi_2}\right] = \max_{Q \in \mathcal{M}} \mathbb{E}_Q\left[\sum_{t=1}^{T} (U - D_t)\right] = T(U - \mu),$$

which implies

$$\mathrm{Opt}_{\mathcal{M}}^T(\mu, U, b) \leq \min\left\{ \max_{Q \in \mathcal{M}} \mathbb{E}_Q\left[\sum_{t=1}^{T} C_t^{\pi_1}\right], \ \max_{Q \in \mathcal{M}} \mathbb{E}_Q\left[\sum_{t=1}^{T} C_t^{\pi_2}\right] \right\}$$

$$= \min\left\{ Tb\mu, \ T(U - \mu) \right\}.$$

Noting that $\mathrm{Opt}_{\mathrm{IND}}^T(\mu, U, b) = \min\left\{ Tb\mu, \ T(U - \mu) \right\}$, completes the proof. ■

Proof. [Proof of Theorem 17] From Theorems 12 and 14,

$$\frac{\mathrm{Opt}_{\mathrm{MAR}}^T(\mu, U, b)}{\mathrm{Opt}_{\mathrm{IND}}^T(\mu, U, b)} = \frac{G_{\Gamma_\mu^T}^T(\beta_\mu^T, \mu)}{\min\left\{ Tb\mu, \ T(U - \mu) \right\}} = \frac{\Gamma_\mu^T A_{\Gamma_\mu^T}^{T+1} + \left(Tb - (b+1)\Gamma_\mu^T \right) \mu}{\min\left\{ Tb\mu, \ T(U - \mu) \right\}}.$$

Combining above with Lemma 29, we complete the proof of Theorem 17. ■

5.5 *Discussion*

In this section, we provide further insights on our closed form solutions, and discuss one interesting non-monotonicity of $\chi_{\mathrm{MAR}}^T(\mu, U, b)$ in terms of b.

5.5.1 Further insights into X_t^T and D_t^T

In this section, we provide further insights into the expressions which describe the dynamics between the inventory manager and adversary. Although the complete proof of Theorem 14 follows from a non-trivial induction involving ideas from convex analysis and probability, there is an intriguing intuition based on a certain "Indifference principle", which we now sketch.

First, note that the two-point support property of D_t conditioned on D_{t-1} $(\neq 0)$ follows from the Richter-Rogosinski Theorem (cf. [182]). Now, let us suppose *apriori* that one point in this support is always 0, and let d_t^T denote the other support point of D_t. In addition, let x_t^T denote the order-up-to level in period t supposing no previous demand equalled zero and assume $x_t^T \leq d_t^T$. In this case, the cost incurred in period 1 equals $\left(1 - \frac{\mu}{d_1^T}\right) x_1^T + \frac{\mu}{d_1^T} b \left(d_1^T - x_1^T\right)$. Applying the definition of martingale, the amount of inventory x_1^T will be carried-on for the remaining periods conditioned on $D_1 = 0$, and the cost conditioned on $D_1 = d_1^T$ equals $\left(1 - \frac{d_1^T}{d_2^T}\right) x_2^T + \frac{d_1^T}{d_2^T} b \left(d_2^T - x_2^T\right)$. Repeating this procedure recursively, we can rewrite the iterated min-max problem as below:

$$
\min_{x_1^T} \max_{d_1^T} \left\{ \left(1 - \frac{\mu}{d_1^T}\right) T x_1^T + \frac{\mu}{d_1^T} \left[b \left(d_1^T - x_1^T\right) \right. \right.
$$
$$
\left. \left. + \min_{x_2^T} \max_{d_2^T} \left\{ \left(1 - \frac{d_1^T}{d_2^T}\right) (T-1) x_2^T + \frac{d_1^T}{d_2^T} \left[b \left(d_2^T - x_2^T\right) + \ldots \right] \right\} \right] \right\}.
$$

Furthermore, for any fixed set of choices $\{x_t^T, d_t^T, t = 1, \ldots, T\}$, it follows from the above and a straightforward induction that the associated overall expected cost equals

$$
\left\{ b\mu \sum_{t=1}^{T} \left(1 - \frac{x_t^T}{d_t^T}\right) + \mu \sum_{t=1}^{T} x_t^T (T+1-t) \left(\frac{1}{d_{t-1}^T} - \frac{1}{d_t^T}\right) \right\}, \tag{166}
$$

where $d_0^T = \mu$. In this expansion, one may compute that the coefficient of x_{t+1}^T equals $\mu \left(\frac{T-t}{d_t^T} - \frac{b+T-t}{d_{t+1}^T}\right)$, and the coefficient of $\frac{1}{d_t^T}$ equals $\mu \left((T-t) x_{t+1}^T - (b+T+1-t) x_t^T\right)$. Setting these coefficients equal to 0, so the inventory manager and the adversary are indifferent to their choices of x_{t+1}^T and d_t^T, yields $x_{t+1}^T = \frac{b+T+1-t}{T-t} x_t^T$, $d_{t+1}^T = \frac{b+T-t}{T-t} d_t^T$.

Noting that these are precisely the recursions satisfied by X_t^T and D_t^T in (119) explains the particular form of X_t^T and D_t^T.

Now let us explain the particular form of X_α^∞ and D_α^∞. Assume $\alpha = \frac{t}{T}$ and then (166) equals

$$b\mu \sum_{\alpha = \frac{1}{T}, \frac{2}{T}, \ldots, 1} \left(1 - \frac{x_{\alpha T}^T}{d_{\alpha T}^T}\right) + \mu \sum_{\alpha = \frac{1}{T}, \frac{2}{T}, \ldots, 1} x_{\alpha T}^T (T + 1 - \alpha T) \left(\frac{1}{d_{(\alpha - \frac{1}{T})T}^T} - \frac{1}{d_{\alpha T}^T}\right).$$

Assume that $x_{\alpha T}^T, d_{\alpha T}^T$ have limits $x_\alpha^\infty, d_\alpha^\infty$ as $T \to \infty$ for each α. Then the associated average cost (i.e., the overall expected cost in (166) divided by T) converges to

$$b\mu \int_0^1 \left(1 - \frac{x_\alpha^\infty}{d_\alpha^\infty}\right) d\alpha - \mu \int_0^1 x_\alpha^\infty (1 - \alpha) \frac{d}{d\alpha}\left(\frac{1}{d_\alpha^\infty}\right) d\alpha. \tag{167}$$

Similarly, one may compute that the coefficient of x_α^∞ equals $\mu \left(-\frac{b}{d_\alpha^\infty} - (1 - \alpha) \frac{d}{d\alpha}\left(\frac{1}{d_\alpha^\infty}\right)\right)$, and the coefficient of $\frac{1}{d_\alpha^\infty}$ equals $\mu \left(-bx_\alpha^\infty + \frac{d}{d\alpha}(x_\alpha^\infty(1 - \alpha))\right)$ after integration by parts. Similar to the pre-limit case, we set these coefficients equal to 0 and obtain

$$\frac{b}{d_\alpha^\infty} + (1 - \alpha) \frac{d}{d\alpha}\left(\frac{1}{d_\alpha^\infty}\right) = 0, \quad -bx_\alpha^\infty + \frac{d}{d\alpha}(x_\alpha^\infty(1 - \alpha)) = 0.$$

Solving the differential equations above and plugging in the boundary conditions $x_{\Lambda^\infty}^\infty = d_{\Lambda^\infty}^\infty = U$, we obtain the exact expressions of X_α^∞ and D_α^∞.

5.5.2 The non-monotonicity of $\chi_{\mathbf{MAR}}^T(\mu, U, b)$ in b

The explicit form of $\chi_{\mathrm{MAR}}^T(\mu, U, b)$ reveals a surprising non-monotonicity in b. Namely, for fixed $T \geq 2, \mu, U, \chi_{\mathrm{MAR}}^T(\mu, U, b)$ is *not* monotone increasing in b. This is surprising, since one would expect that as the backlogging penalty increases, one would wish to stock higher inventory levels. We also note that for fixed $T, \mu, U, \chi_{\mathrm{IND}}^T(\mu, U, b)$ is monotone increasing in b, i.e. such a non-monotonicity does not manifest in the independent-demand model. We reason as below.

Suppose the inventory manager orders x_1 in the first round and the adversary selects either 0 or U. If the adversary selects U, there will be no cost from period 2 to period T since the manager can always match the demand; if the adversary selects 0,

166

a fixed holding cost x_1 will be incurred in the remaining every period. It means that decreasing x_1 results in more backorder cost just in period 1, but less holding cost from period 1 to period T. Hence if μ is not large such that the adversary does not put too many weights on U, the inventory manager would like to choose x_1 as small as possible as long as the adversary still selects either 0 or U in period 1. Note that this critical (smallest) value of x_1 is decreasing in b, i.e., when b increases, the adversary still has incentive to choose 0 or U for smaller x_1. It explains the non-monotonicity of $\chi^T_{\mathrm{MAR}}(\mu, U, b)$ in terms of b. For example, when $T = 2$, this critical value equals $\frac{U}{b+2}$ (see Appendix, Section 5.7).

An alternative way to explain the non-monotonicity is from the recursion $X^T_{t+1} = \frac{b+T+1-t}{T-t} X^T_t$ discussed in Section 5.5.1. As we mentioned in the last paragraph, it is possible to have $D^T_1 = X^T_2 = U$. In such case, $X^T_1 = \frac{T-1}{b+T} U$, which is strictly decreasing in b.

5.6 Conclusion

In this chapter, we proposed a novel multi-period inventory model by combining the framework of distributionally robust optimization with the theory of martingales. More precisely, we considered the setting in which the joint distribution (over time) must take the form of a martingale with bounded support (called martingale-demand model) and wished to pick the control policy which was optimal against a worst-case distribution belonging to this set. We explicitly computed the optimal policy and derived an interesting interplay between the optimal policy and worst-case distribution. We performed an asymptotic analysis (as the time horizon diverged) and proved weak convergence of the inventory process at optimality. We also compared to an existing model, in which the adversary was restricted to product measures (called independent-demand model). Interestingly, we found that the limiting ratio of the optimal cost under the martingale and independent models to be exactly $\frac{1}{2}$ in the

perfectly symmetric case.

This work leaves several interesting directions for future research. First, we have taken the first step towards establishing a conditional-expectation based theory of dynamic distributional robust forecasting. In the future, it would be interesting to consider more general conditional moment constraints, e.g., the conditional mean is an affine combination of all demands realized. Second, although it is often clear how to specify the marginal distribution in each time period, understanding the effects of positing various joint distributions over time remains an interesting challenge (cf. [1]). It would be interesting to develop a deeper understanding of such "price of correlations" in robust stochastic optimization problems. Third, using moment information is only one among many other ways to construct an uncertainty set. It would be very interesting to connect our framework to phi-divergence ambiguity set (cf. [16]), which uses the empirical probability density function from the historical data; to distributionally robust Bayesian model [117], which dynamically updates the uncertainty set following the Bayes Rule; and to the general theory of risk measures (cf. [182]).

5.7 Appendix

This appendix presents explicit solutions of the martingale-demand model for certain parameters.

5.7.1 Explicit solutions when $T = 1, 2, 3$

This section provides explicit expressions of $\chi^T_{\mathrm{MAR}}(\mu, U, b)$, $\mathrm{Opt}^T_{\mathrm{MAR}}(\mu, U, b)$ and $\mathfrak{q}^T_{x,\mu}$ for $T = 1, 2, 3$.

$T = 1$:

$$\chi^1_{\mathrm{MAR}}(\mu, U, b) = \begin{cases} 0 & \text{if } 0 \leq \mu \leq \frac{U}{b+1}; \\ U & \text{if } \frac{U}{b+1} < \mu \leq U, \end{cases} \qquad \mathrm{Opt}^1_{\mathrm{MAR}}(\mu, U, b) = \begin{cases} b\mu & \text{if } 0 \leq \mu \leq \frac{U}{b+1}; \\ U - \mu & \text{if } \frac{U}{b+1} < \mu \leq U. \end{cases}$$

$\mathfrak{q}^1_{x,\mu}$ is supported on $\{0, U\}$ for all $0 \le \mu \le U$, $0 \le x \le U$.

$T = 2$:

$$\chi^2_{\mathrm{MAR}}(\mu, U, b) = \begin{cases} 0 & \text{if } 0 \le \mu \le \frac{2U}{(b+1)(b+2)}; \\[2ex] \frac{U}{b+2} & \text{if } \frac{2U}{(b+1)(b+2)} < \mu \le \frac{2U}{b+2}; \\[2ex] U & \text{if } \frac{2U}{b+2} < \mu \le U, \end{cases}$$

$$\mathrm{Opt}^2_{\mathrm{MAR}}(\mu, U, b) = \begin{cases} 2b\mu & \text{if } 0 \le \mu \le \frac{2U}{(b+1)(b+2)}; \\[2ex] \frac{2U}{b+2} + (b-1)\mu & \text{if } \frac{2U}{(b+1)(b+2)} < \mu \le \frac{2U}{b+2}; \\[2ex] 2(U - \mu) & \text{if } \frac{2U}{b+2} < \mu \le U. \end{cases}$$

$\mathfrak{q}^2_{x,\mu}$ is supported on

$$\begin{cases} \{0, \frac{U}{b+1}\} & \text{if } 0 \le \mu \le \frac{U}{b+1}, \ 0 \le x < \frac{U}{b+2}; \\[2ex] \{\frac{U}{b+1}, U\} & \text{if } \frac{U}{b+1} < \mu \le U, \ 0 \le x < \frac{U}{b+2}; \\[2ex] \{0, U\} & \text{if } \frac{U}{b+2} \le x \le U. \end{cases}$$

$T = 3$:

$$\chi^3_{\mathrm{MAR}}(\mu, U, b) = \begin{cases} 0 & \text{if } 0 \le \mu \le \frac{6U}{(b+1)(b+2)(b+3)}; \\[2ex] \frac{2U}{(b+2)(b+3)} & \text{if } \frac{6U}{(b+1)(b+2)(b+3)} < \mu \le \frac{6U}{(b+2)(b+3)}; \\[2ex] \frac{2U}{b+3} & \text{if } \frac{6U}{(b+2)(b+3)} < \mu \le \frac{3U}{b+3}; \\[2ex] U & \text{if } \frac{3U}{b+3} < \mu \le U, \end{cases}$$

$$\mathrm{Opt}^3_{\mathrm{MAR}}(\mu, U, b) = \begin{cases} 3b\mu & \text{if } 0 \le \mu \le \frac{6U}{(b+1)(b+2)(b+3)}; \\[2ex] \frac{6U}{(b+2)(b+3)} + (2b-1)\mu & \text{if } \frac{6U}{(b+1)(b+2)(b+3)} < \mu \le \frac{6U}{(b+2)(b+3)}; \\[2ex] \frac{6U}{b+3} + (b-2)\mu & \text{if } \frac{6U}{(b+2)(b+3)} < \mu \le \frac{3U}{b+3}; \\[2ex] 3(U - \mu) & \text{if } \frac{3U}{b+3} < \mu \le U. \end{cases}$$

$\mathfrak{q}^3_{x,\mu}$ is supported on

$$
\begin{cases}
\left\{0, \frac{2U}{(b+1)(b+2)}\right\} & \text{if } 0 \leq \mu \leq \frac{2U}{(b+1)(b+2)}, \ 0 \leq x < \frac{2U}{(b+2)(b+3)}; \\[2ex]
\left\{\frac{2U}{(b+1)(b+2)}, \frac{2U}{b+2}\right\} & \text{if } \frac{2U}{(b+1)(b+2)} < \mu \leq \frac{2U}{b+2}, \ 0 \leq x < \frac{2U}{(b+2)(b+3)}; \\[2ex]
\left\{0, \frac{2U}{b+2}\right\} & \text{if } 0 \leq \mu \leq \frac{2U}{b+2}, \ \frac{2U}{(b+2)(b+3)} \leq x < \frac{2U}{b+3}; \\[2ex]
\left\{\frac{2U}{b+2}, U\right\} & \text{if } \frac{2U}{b+2} < \mu \leq U, \ 0 \leq x < \frac{2U}{b+3}; \\[2ex]
\left\{0, U\right\} & \text{if } \frac{2U}{b+3} \leq x \leq U.
\end{cases}
$$

5.7.2 Explicit solutions when $b = 1$

This section provides explicit expressions of $\chi^T_{\mathrm{MAR}}(\mu, U, b)$, $\mathrm{Opt}^T_{\mathrm{MAR}}(\mu, U, b)$ and $\mathfrak{q}^T_{x,\mu}$ for $b = 1$.

If $\mu = 0$, $\chi^T_{\mathrm{MAR}}(\mu, U, 1) = \mathrm{Opt}^T_{\mathrm{MAR}}(\mu, U, 1) = 0$. Otherwise, if $\frac{j}{T+1} < \mu \leq \frac{j+1}{T+1}$, $j = 0, \ldots, T$,

$$
\chi^T_{\mathrm{MAR}}(\mu, U, 1) = \frac{j(j+1)U}{T(T+1)}, \quad \mathrm{Opt}^T_{\mathrm{MAR}}(\mu, U, 1) = \frac{j(j+1)U}{T+1} + (T - 2j)\mu.
$$

$\mathfrak{q}^T_{x,\mu}$ is supported on

$$
\begin{cases}
\left\{\frac{jU}{T}, \frac{(j+1)U}{T}\right\} & \text{if } \frac{jU}{T} < \mu \leq \frac{(j+1)U}{T}, \ 0 \leq x < \frac{j(j+1)U}{T(T+1)}; \\[2ex]
\left\{0, \frac{(k+1)U}{T}\right\} & \text{if } \frac{jU}{T} < \mu \leq \frac{(j+1)U}{T}, \ \frac{k(k+1)U}{T(T+1)} \leq x < \frac{(k+1)(k+2)U}{T(T+1)}, \ k \geq j; \\[2ex]
\left\{0, U\right\} & \text{if } \mu = 0 \text{ or } x = U.
\end{cases}
$$

REFERENCES

[1] AGRAWAL, S., DING, Y., SABERI, A., and YE, Y., "Price of correlations in stochastic optimization," *Operations Research*, vol. 60, no. 1, pp. 150–162, 2012.

[2] AHMED, S., CAKMAK, U., and SHAPIRO, A., "Coherent risk measures in inventory problems," *European Journal of Operational Research*, vol. 182, pp. 226–238, 2007.

[3] ALLON, G. and MIEGHEM, J. A. V., "Global dual sourcing: Tailored base-surge allocation to near- and offshore production," *Management Science*, vol. 56, no. 1, pp. 110–124, 2010.

[4] ANGELUS, A. and Ö. ÖZER, "Knowledge you can act on: optimal policies for assembly systems with expediting and advance demand information." working paper, 2015.

[5] ARAPOSTATHIS, A., BORKAR, V. S., FERNÁNDEZ-GAUCHERAND, E., GHOSH, M. K., and MARCUS, S. I., "Discrete time controlled markov processes with average cost criterion: a survey," *SIAM Journal on Control and Optimization*, vol. 31, pp. 282–344, 1993.

[6] ARTZNER, P., DELBAEN, F., EBER, J.-M., HEATH, D., and KU, H., "Coherent multiperiod risk adjusted values and bellmans principle," *Annals of Operations Research*, vol. 152, pp. 5–22, 2007.

[7] ASMUSSEN, S., *Applied Probability and Queues*. Springer, Second ed., 2003.

[8] BADINELLIN, R., "The inventory costs of common mis-specification of demand-forecasting models," *The international journal of production research*, vol. 28, no. 12, pp. 2321–2340, 1990.

[9] BARANKIN, E. W., "A delivery-lag inventory model with an emergency provision," *Naval Research Logistics Quarterly*, vol. 8, no. 3, pp. 285–311, 1961.

[10] BEIGLBOCK, M. and NUTZ, M., "Martingale inequalities and deterministic counterparts," 2014. Preprint.

[11] BELLMAN, R., *Dynamic Programming*. Princeton, NJ: Princeton University Press, 1957.

[12] BEN-TAL, A., BOAZ, G., and SHIMRIT, S., "Robust multi-echelon multi-period inventory control," *European Journal of Operational Research*, vol. 199, no. 3, pp. 922–935, 2009.

[13] BEN-TAL, A., GHAOUI, L. E., and NEMIROVSKI, A., *Robust Optimization.* Princeton, NJ: Princeton University Press, 2009.

[14] BEN-TAL, A., GOLANY, B., NEMIROVSKI, A., and VIAL, J., "Retailer-supplier flexible commitments contracts: a robust optimization approach," *Manufacturing & Service Operations Management*, vol. 7, pp. 248–271, 2005.

[15] BEN-TAL, A., GORYASHKO, A., GUSLITZER, A., and NEMIROVSKI, A., "Adjustable robust solutions of uncertain linear programs," *Mathematical Programming*, vol. 99, pp. 351–376, 2004.

[16] BEN-TAL, A., HERTOG, D. D., WAEGENAERE, A. D., MELENBERG, B., and RENNEN, G., "Robust solutions of optimization problems affected by uncertain probabilities," *Management Science*, vol. 59, pp. 341–357, 2013.

[17] BERTSEKAS, D. and SHREVE, S., *Stochastic Optimal Control, The Discrete Time Case.* New York: Academic Press, 1978.

[18] BERTSIMAS, D., BROWN, D., and CARAMANIS, C., "Theory and applications of robust optimization," *SIAM Review*, vol. 53, no. 3, pp. 464–501, 2011.

[19] BERTSIMAS, D., DOAN, X., NATARAJAN, K., and TEO, C., "Models for minimax stochastic linear optimization problems with risk aversion," *Mathematics of Operations Research*, vol. 35, no. 3, pp. 580–602, 2010.

[20] BERTSIMAS, D., IANCU, D. A., and PARRILO, P., "Optimality of affine policies in multistage robust optimization," *Mathematics of Operations Research*, vol. 35, no. 2, pp. 363–394, 2010.

[21] BERTSIMAS, D., SIM, M., and ZHANG, M., "A practicable framework for distributionally robust linear optimization," 2014. Preprint.

[22] BERTSIMAS, D. and THIELE, A., "A robust optimization approach to inventory theory," *Operations Research*, vol. 54, no. 1, pp. 150–168, 2006.

[23] BIELECKI, T., CIALENCO, I., and PITERA, M., "On time consistency of dynamic risk and performance measures in discrete time," 2014. arXiv preprint arXiv:1409.7028.

[24] BIJVANK, M. and VIS, I., "Lost-sales inventory theory: A review," *European Journal of Operations Research*, vol. 215, pp. 1–13, 2011.

[25] BILLINGSLEY, P., *Convergence of Probability Measures.* New York: Wiley, 1999.

[26] BLINDER, A., "Inventories and sticky prices: More on the microfoundations of macroeconomics," *National Bureau of Economic Research*, vol. 72, no. 3, pp. 334–348, 1982.

[27] BODA, K. and FILAR, J., "Time consistent dynamic risk measures," *Mathematical Methods in Operations Research*, vol. 63, pp. 169–186, 2006.

[28] BONNANS, J. F. and SHAPIRO, A., *Perturbation Analysis of Optimization Problems*. New York: Springer-Verlag, 2000.

[29] BOUTE, R. and MIEGHEM, J. V., "Global dual sourcing and order smoothing: The impact of capacity and leadtimes," *Management Science*, 2014. forthcoming.

[30] BRADLEY, J. R., "A brownian approximation of a production-inventory system with a manufacturer that subcontracts," *Operations Research*, vol. 52, no. 5, pp. 765–784, 2004.

[31] BROWN, G., LU, J., and WOLFSON, R., "Dynamic modeling of inventories subject to obsolescence," *Management Science*, vol. 11, no. 1, pp. 51–63, 1964.

[32] CARPENTIER, P., CHANCELIER, J., COHEN, G., LARA, M. D., and GIRARDEAU, P., "Dynamic consistency for stochastic optimal control problems," *Annals of Operations Research*, vol. 200, pp. 247–263, 2012.

[33] CARRIZOSAA, E., OLIVARES-NADAL, A., and RAMIREZ-COBOB, P., "Robust newsvendor problem with autoregressive demand." Working Paper, 2014.

[34] CERAG, P., "Advances in inventory management: Dynamic models," *Erasmus Research Institute of Management (ERIM)*, 2010. No. EPS-2010-199-LIS.

[35] CHEN, F., DREZNER, Z., RYAN, J., and SIMCHI-LEVI, D., "Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information," *Management Science*, vol. 46, no. 3, pp. 436–443, 2000.

[36] CHEN, W., DAWANDE, M., and JANAKIRAMAN, G., "Fixed-dimensional stochastic dynamic programs: An approximation scheme and an inventory application," *Operations Research*, vol. 62, no. 1, pp. 81–103, 2014.

[37] CHEN, W. and SIM, M., "Goal-driven optimization," *Operations Research*, vol. 57, no. 2, pp. 342–357, 2009.

[38] CHEN, X., SIM, M., SIMCHI-LEVI, D., and SUN, P., "Risk aversion in inventory management," *Operations Research*, vol. 55, no. 5, pp. 828–842, 2007.

[39] CHEN, X. and SUN, P., "Optimal structural policies for ambiguity and risk averse inventory and pricing models," *SIAM Journal on Control and Optimization*, vol. 50, no. 1, pp. 133–146, 2012.

[40] CHEN, Z., LI, G., and GUO, J., "Optimal investment policy in the time consistent mean-variance formulation," *Insurance: Mathematics and Economics*, vol. 52, pp. 145–156, 2013.

[41] CHENG, X. and RIEDEL, F., "Optimal stopping under ambiguity in continuous time," *Mathematics and Financial Economics*, vol. 7, no. 1, pp. 29–68, 2013.

[42] CHERIDITO, P. and KUPPER, M., "Recursiveness of indifference prices and translation invariant preferences," *Mathematics and Financial Economics*, vol. 2, no. 3, pp. 173–188, 2009.

[43] CHOI, S. and RUSZCZYŃSKI, A., "A risk-averse newsvendor with law invariant coherent measures of risk," *Operations Research Letters*, vol. 36, no. 1, pp. 77–82, 2008.

[44] CHOI, S., RUSZCZYŃSKI, A., and ZHAO, Y., "A multiproduct risk-averse newswendor with law-invariant coherent measures of risk," *Operations Research*, vol. 59, pp. 346–354, 2011.

[45] DANIEL, K. H., "A delivery-lag inventory model with emergency order," in *Multistage Inventory Models and Techniques* (SCARF, H., GILFORD, D., and SHELLY, M., eds.), Stanford University Press, 1963.

[46] DECROIX, G., SONG, J., and ZIPKIN, P., "A series system with returns: Stationary analysis," *Operations Research*, vol. 53, pp. 350–362, 2005.

[47] DELAGE, E. and YE, Y., "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Operations Research*, vol. 58, no. 3, pp. 596–612, 2010.

[48] DEUSCHEL, J.-D. and STROOCK, D. W., *Large deviations*. Academic Press, 1989.

[49] DOWNS, B., METTERS, R., and SEMPLE, J., "Managing inventory with multiple products, lags in delivery, resource constraints and lost sales: A mathematical programming approach," *Management Science*, vol. 47, no. 3, pp. 464–479, 2001.

[50] DUDLEY, R., *Real Analysis and Probability*. Cambridge University Press, 2002.

[51] DUPACOVÁ, J., "The minimax approach to stochastic programming and an illustrative application," *Stochastics*, vol. 20, pp. 73–88, 1987.

[52] DUPACOVÁ, J., "Stochastic programming: Minimax approach," in *Encyclopedia of Optimization*, vol. 5, pp. 327–330, Kluwer Academic Publishers, 2001.

[53] EDGEWORTH, F., "The mathematical theory of banking," *Royal Statistical Society*, vol. 51, pp. 113–127, 1888.

[54] EPSTEIN, L. G. and SCHNEIDER, M., "Recursive multiple-priors," *Journal of Economic Theory*, vol. 113, no. 1, pp. 1–31, 2003.

[55] ERDOS, P. and KAC, M., "On certain limit theorems of the theory of probability," *Bull. Amer. Math. Soc.*, vol. 52, pp. 292–302, 1946.

[56] ERKIP, N., HAUSMAN, W., and NAHMIAS, S., "Optimal centralized ordering policies in multi-echelon inventory systems with correlated demands," *Management Science*, vol. 36, no. 3, pp. 381–392, 1990.

[57] ETNER, J., JELEVA, M., and TALLON, J.-M., "Decision theory under ambiguity," *Journal of Economic Surveys*, vol. 26, no. 2, pp. 234–270, 2012.

[58] F H.

[59] FEINBERG, E., "Total expected discounted reward mdps: Existence of optimal policies," in *Wiley Encyclopedia of Operations Research and Management Science*, Wiley, 2011.

[60] FELDMAN, R., "A continuous review (s, s) inventory system in a random environment," *Journal of Applied Probability*, vol. 15, no. 3, pp. 654–659, 1978.

[61] FENG, Q., SETHI, S., YAN, H., and ZHANG, H., "Are base-stock policies optimal in inventory problems with multiple delivery modes?," *Operations Research*, vol. 54, no. 4, pp. 801–807, 2006.

[62] FILDES, R. and MAKRIDAKIS, S., "The impact of empirical accuracy studies on time series analysis and forecasting," *International Statistical Review*, vol. 63, no. 3, pp. 289–308, 1995.

[63] FILDES, R., NIKOLOPOULOS, K., CRONE, S. F., and SYNTETOS, A. A., "Forecasting and operational research: A review," *The Journal of the Operational Research Society*, vol. 59, no. 9, pp. 1150–1172, 2008.

[64] FLEISCHMANN, M. and KUIK, R., "An optimal inventory control with independent stochastic item returns," *European Journal of Operational Research*, vol. 151, pp. 25–37, 2003.

[65] FOLLAND, G. B., *Real Analysis: Modern Techniques and Their Applications*. Wiley, Second ed., 1999.

[66] FOX, E., METTERS, R., and SEMPLE, J., "Optimal inventory policy with two suppliers," *Operations Research*, vol. 54, no. 2, pp. 389–393, 2006.

[67] FUKUDA, Y., "Optimal policies for the inventory problem with negotiable leadtime," *Management Science*, vol. 10, pp. 690–708, 1964.

[68] GABREL, V., MURAT, C., and THIELE, A., "Recent advances in robust optimization: an overview," *European Journal of Operational Research*, vol. 235, no. 3, pp. 471–483, 2014.

[69] GALLEGO, G., "New bounds and heuristics for (q, r) policies," *Management Science*, vol. 44, no. 2, pp. 219–233, 1998.

[70] GALLEGO, G. and K. KATIRCIOGLU, B. R., "Inventory management under highly uncertain demand," *Operations Research Letters*, vol. 35, no. 3, pp. 281–289, 2007.

[71] GALLEGO, G. and MOON, I., "The distribution free newsboy problem: review and extensions," *Journal of the Operational Research Society*, vol. 44, no. 8, pp. 825–834, 1993.

[72] GALLEGO, G. and MOON, I., "Distribution free procedures for some inventory models," *Journal of the Operational Research Society*, vol. 45, no. 6, pp. 651–658, 1994.

[73] GALLEGO, G., RYAN, J., and SIMCHI-LEVI, D., "Minimax analysis for finite horizon inventory models," *IIE Transactions*, vol. 33, pp. 861–874, 2001.

[74] GAVER, D., "On base-stock level inventory control," *Operations Research*, vol. 7, pp. 689–703, 1959.

[75] GOLDBERG, D., KATZ-ROGOZHNIKOV, D., LU, Y., SHARMA, M., and SQUILLANTE, M., "Asymptotic optimality of constant-order policies for lost sales inventory models with large lead times," *arXiv preprint arXiv:1211.4063v1*, 2012.

[76] GONG, X., CHAO, X., and ZHENG, S., "Dynamic pricing and inventory management with dual suppliers of different lead times and disruption risks," *Production and Operations Management*, 2014. forthcoming.

[77] GRAVES, S., MEAL, M., DASU, S., and QIU, Y., "Two-stage production planning in a dynamic environment," in *Multi-Stage Production Planning and Inventory Control* (AXSATER, S., SCHNEEWEISS, C., and SILVER, E., eds.), vol. 266 of *Lecture Notes in Economics and Mathematical Systems*, pp. 9–43, Berlin: Springer-Verlag, 1986.

[78] GRUEN, T., CORSTEN, D., and BHARADWAY, S., "Retail out-of-stocks: A worldwide examination of extent, causes and consumer responses," *Grocery Manufacturers of America*, 2002.

[79] GRUNWALD, P. and HALPERN, J. Y., "Making decisions using sets of probabilities: updating, time consistency, and calibration," *Journal of Artificial Intelligence Research*, vol. 42, no. 1, pp. 393–426, 2011.

[80] GUIDE, D. and SRIVASTAVA, R., "Repairable inventory theory: models and applications," *European Journal of Operational Research*, vol. 102, no. 1, pp. 1–20, 1997.

[81] HAIGH, J., *Probability Models.* Springer, 2013.

[82] HAJEK, B., "The proof of a folk theorem on queuing delay with applications to routing in networks," *Journal of the ACM (JACM)*, vol. 30, no. 4, pp. 834–851, 1983.

[83] HALMAN, N., KLABJAN, D., MOSTAGIR, M., ORLIN, J., and SIMCHI-LEVI, D., "A fully polynomial time approximation scheme for single-item inventory control with discrete demand," *Mathematics of Operations Research*, vol. 34, pp. 674–685, 2009.

[84] HALMAN, N., ORLIN, J., and SIMCHI-LEVI, D., "Approximating the nonlinear newsvendor and single-item stochastic lot-sizing problems when data is given by an oracle," *Operations Research*, vol. 60, no. 2, pp. 429–446, 2012.

[85] HANASUSANTO, G., GRANI, A., KUHN, D., WALLACE, W., and ZYMLER, S., "Distributionally robust multi-item newsvendor problems with multimodal demand distributions," *Mathematical Programming*, 2014. forthcoming.

[86] HANSEN, L. and SARGENT, T., "Robust control and model uncertainty," *American Economic Review*, vol. 91, no. 2, pp. 60–66, 2001.

[87] HEATH, D. C. and JACKSON, P. L., "Modeling the evolution of demand forecasts with application," *IIE Transactions*, vol. 26, no. 3, pp. 17–30, 1994.

[88] HEYMAN, D. and SOBEL, M., *Stochastic Models in Operations Research*, vol. II. New York: McGraw-Hill, 1984.

[89] HORDIJK, A. and TIJMS, H., "Convergence results and approximations for optimal (s, s) policies," *Management Sciences*, vol. 20, no. 11, pp. 1432–1438, 1974.

[90] HORDIJK, A. and TIJMS, H., "On a conjecture of iglehart," *Management Sciences*, vol. 21, no. 11, pp. 1342–1345, 1975.

[91] HUA, Z., YU, Y., ZHANG, W., and XU, X., "Structural properties of the optimal policy for dual-sourcing systems with general lead times," *IIE Transactions*, 2014. forthcoming.

[92] HUGGINS, E. L. and OLSEN, T. L., "Inventory control with generalized expediting," *Operations Research*, vol. 58, no. 5, pp. 1414–1426, 2010.

[93] HUH, T., JANAKIRAMAN, G., MUCKSTADT, J., and RUSMEVICHIENTONG, P., "Asymptotic optimality of order-up-to policies in lost sales inventory systems," *Management Sciences*, vol. 55, no. 3, pp. 404–420, 2009.

[94] HUH, W. and JANAKIRAMAN, G., "On the optimal policy structure in serial inventory systems with lost sales," *Operations Research*, vol. 58, no. 2, pp. 486–491, 2010.

[95] HUH, W., JANAKIRAMAN, G., and NAGARAJAN, M., "Average cost single-stage inventory models: an analysis using a vanishing discount approach," *Operations Research*, vol. 59, no. 1, pp. 143–155, 2011.

[96] HUMBLET, P., "Determinism minimizes waiting time in queues," *MIT Technical report*, 1982.

[97] IANCU, D., PETRIK, M., and SUBRAMANIAN, D., "Tight approximations of dynamic risk measures," *Mathematics of Operations Research*, 2014. forthcoming.

[98] IGLEHART, D. and KARLIN, S., "Optimal policy for dynamic inventory process with nonstationary stochastic demands," in *Studies in the Mathematical Theory of Inventory and Production* (ARROW, K., KARLIN, S., and SCARF, H., eds.), Stanford University Press, 1962.

[99] IGLEHART, D. L., "Optimality of (s, s) policies in the infinite horizon dynamic inventory problem," *Management Science*, vol. 9, no. 2, pp. 259–267, 1963.

[100] IGNALL, E. and VEINOTT, A., "Optimality of myopic inventory policies for several substitute products," *Management Science*, vol. 15, no. 5, pp. 284–304, 1969.

[101] IIDA, T. and ZIPKIN, P., "Approximate solutions of a dynamic forecasting inventory model," *Manufacturing & Service Operations Management*, vol. 8, no. 4, pp. 407–425, 2006.

[102] ISII, K., "On sharpness of tchebycheff-type inequalities," *Annals of the Institute of Statistical Mathematics*, vol. 14, no. 1, pp. 185–197, 1962.

[103] IYENGAR, G., "Robust dynamic programming," *Mathematics of Operations Research*, vol. 30, no. 2, pp. 257–280, 2005.

[104] JANAKIRAMAN, G., SESHADRI, S., and SHANTHIKUMAR, J., "A comparison of the optimal costs of two canonical inventory systems," *Operations Research*, vol. 55, pp. 866–875, 2007.

[105] JANAKIRAMAN, G., SESHADRI, S., and SHEOPURI, A., "Analysis of tailored base-surge policies in dual sourcing inventory systems," *Management Science*, 2014. forthcoming.

[106] JANSSEN, F. and KOK, T. D., "A two-supplier inventory model," *International Journal of Production Economics*, vol. 59, pp. 395–403, 1999.

[107] JIANG, Q., FENG, W., PENG, J., HAN, F., LI, Q., ZHANG, W., and SU, H., "Inventory-based resource management in cloud computing," in *Distributed Computing and Applications to Business, Engineering and Science*, pp. 142–145, IEEE, 2011.

[108] JOGLEKAR, P. and LEE, P., "An exact formulation of inventory costs and optimal lot size in face of sudden obsolescence," *Operations Research Letters*, vol. 14, no. 5, pp. 283–290, 1993.

[109] JOHANSEN, S. and THORSTENSON, A., "Pure and restricted base-stock policies for the lost-sales inventory system with periodic review and constant lead times." Technical report, Aarhus University, 2008.

[110] JOHANSEN, S., "Pure and modified base-stock policies for the lost sales inventory system with negligible set-up costs and constant lead times," *International Journal of Production Economics*, vol. 71, pp. 391–399, 2001.

[111] JOHNSON, G. D. and THOMPSON, H. E., "Optimality of myopic inventory policies for certain dependent demand processes," *Management Science*, vol. 21, no. 11, pp. 1303–1307, 1975.

[112] KALYMON, B., "Stochastic prices in a single-item inventory purchasing model," *Operations Research*, vol. 19, no. 6, pp. 1434–1458, 1971.

[113] KARLIN, S. and SCARF, H., "Inventory models of the Arrow-Harris-Marschak type with time lag," in *Studies in the Mathematical Theory of Inventory and Production* (ARROW, K., KARLIN, S., and SCARF, H., eds.), pp. 155–178, Stanford University Press, 1958.

[114] KASUGAI, H. and KASEGAI, T., "Note on minimax regret ordering policy-static and dynamic solutions and a comparison to maximin policy," *Journal of the Operations Research Society of Japan*, vol. 3, no. 4, pp. 155–169, 1961.

[115] KELLEHER, K., "Why fedex is gaining ground," *Business 2.0*, vol. October, pp. 56–57, 2003.

[116] KINGMAN, J., "Some inequalities for the queue gi/g/1," *SIAM Journal on Applied Mathematics*, vol. 49, pp. 315–324, 1962.

[117] KLABJAN, D., SIMCHI-LEVI, D., and SONG, M., "Robust stochastic lot-sizing by means of histograms," *Production and Operations Management*, vol. 22, pp. 691–710, 2013.

[118] KLOSTERHALFEN, S. and KIESM G.

[119] LAM, H. and GHOSH, S., "Iterative method for robust estimation under bivariate uncertainty," in *Proceedings of the Winter Simulation Conference*, 2013.

[120] LANDAU, H., ed., *Moments in mathematics*, vol. 37 of *Proceedings of symposia in applied mathematics*. Providence, RI: American Mathematical Society, 1987.

[121] LAWSON, D. G. and PORTEUS, E. L., "Multistage inventory management with expediting," *Operations Research*, vol. 48, pp. 878–893, 2000.

[122] LEE, H. and V. PADMANABHAN, S. W., "Information distortion in a supply chain: the bullwhip effect," *Management Science*, vol. 50, no. 12, pp. 1875–1886, 2004.

[123] LEVI, R., JANAKIRAMAN, G., and NAGARAJAN, M., "A 2-approximation algorithm for stochastic inventory control models with lost-sales," *Mathematics of Operations Research*, vol. 33, no. 2, pp. 351–374, 2008.

[124] LEVI, R., PA'L, M., ROUNDY, R., and SHMOYS, D., "Approximation algorithms for stochastic inventory control models," *Mathematics of Operations Research*, vol. 32, no. 2, pp. 284–302, 2008.

[125] LEVI, R., PERAKIS, G., and UICHANCO, J., "The data-driven newsvendor problem: new bounds and insights," *preprint*, 2011.

[126] LEVI, R., ROUNDY, R., SHMOYS, D., and TRUONG, V., "Approximation algorithms for capacitated stochastic inventory control models," *Operations Research*, vol. 56, no. 5, pp. 1184–1199, 2008.

[127] LOVEJOY, W., "Stopped myopic policies forsome inventory models with uncertain demand distributions," *Management Science*, vol. 38, no. 5, pp. 688–707, 1992.

[128] LU, X., SONG, J. S., and REGAN, A., "Inventory planning with forecast updates: Approximate solutions and cost error bounds," *Operations Research*, vol. 54, no. 6, pp. 1079–1097, 2006.

[129] LU, Y., SQUILLANTE, M., and YAO, D., "Matching supply and demand – asymptotics and insights," *Preprint*, 2012.

[130] MEYN, S. and TWEEDIE, R., *Markov chains and stochastic stability.* Cambridge University Press, 2009.

[131] MIEGHEM, J. A. V., *Operations Strategy: Principles and Practice.* Belmont, MA: Dynamic Ideas, 2008.

[132] MILLER, B., "Scarf's state reduction method, flexibility, and a dependent demand inventory model," *Operations Research*, vol. 34, no. 1, pp. 83–90, 1986.

[133] MILNER, J. M. and KOUVELIS, P., "Order quantity and timing flexibility in supply chains: The role of demand characteristics," *Management Science*, vol. 51, no. 6, pp. 970–985, 2005.

[134] MINNER, S., "Multiple-supplier inventory models in supply chain management: A review," *International Journal of Production Economics*, vol. 81-82, pp. 265–279, 2003.

[135] MOINZADEH, K. and NAHMIAS, S., "A continuous review model for an inventory system with two supply modes," *Management Science*, vol. 34, no. 6, pp. 761–773, 1988.

[136] MOINZADEH, K. and SCHMIDT, C. P., "An (s-1, s) inventory system with emergency orders," *Operations Research*, vol. 39, no. 2, pp. 308–321, 1991.

[137] MORSE, P., "Solutions of a class of discrete-time inventory problems," *Operations Research*, vol. 7, pp. 67–68, 1959.

[138] MORTON, T. E., "Bounds on the solution of the lagged optimal inventory equation with no demand backlogging and proportional costs," *SIAM Review*, vol. 11, no. 4, pp. 572–596, 1969.

[139] MORTON, T. E., "The near-myopic nature of the lagged-proportional-cost inventory problem with lost sales," *Operations Research*, vol. 19, no. 7, pp. 1708–1716, 1971.

[140] NAHMIAS, S., "Simple approximations for a variety of dynamic leadtime lost-sales inventory models," *Operations Research*, vol. 27, no. 5, pp. 904–924, 1979.

[141] NAIR, J., ADLAKHA, S., and WIERMAN, A., *Energy procurement strategies in the presence of intermittent sources*, 2012. Technical report, Caltech Technical Report.

[142] NATARAJAN, K. and ZHOU, L., "A mean-variance bound for a three-piece linear function," *Probability in the Engineering and Informational Sciences*, vol. 21, no. 4, p. 611C621, 2007.

[143] NEUTS, F., "An inventory model with optimal time lag," *SIAM Journal on Applied Mathematics*, vol. 12, pp. 179–185, 1979.

[144] NICHOLSON, L., VAKHARIA, A., and ERENGUC, S., "Outsourcing inventory management decisions in healthcare: Models and application," *European Journal of Operational Research*, vol. 154, no. 1, pp. 271–290, 2004.

[145] NILIM, A. and GHAOUI, L. E., "Robust control of markov decision processes with uncertain transition matrices," *Operations Research*, vol. 53, no. 5, pp. 780–798, 2005.

[146] NISHIMURA, K. and OZAKI, H., "Irreversible investment and knightian uncertainty," *Journal of Economic Theory*, vol. 136, pp. 668–694, 2007.

[147] PARKER, R. and KAPUSCINSKI, R., "Optimal policies for a capacitated twoechelon inventory system," *Operations Research*, vol. 52, pp. 739–755, 2004.

[148] PERAKIS, G. and ROELS, G., "Regret in the newsvendor model with partial information," *Operations Research*, vol. 56, no. 1, pp. 188–203, 2008.

[149] PIERSKALLA, W., "An inventory problem with obsolescence," *Naval Research Logistics Quarterly*, vol. 16, no. 2, pp. 217–228, 1969.

[150] PINDYCK, R., "Adjustment costs, uncertainty, and the behavior of the firm," *The American Economic Review*, vol. 72, no. 3, pp. 415–427, 1982.

[151] POPESCU, I., "A semidefinite programming approach to optimal moment bounds for convex classes of distributions," *Mathematics of Operations Research*, vol. 50, no. 3, pp. 632–657, 2005.

[152] POPESCU, I., "Robust mean-covariance solutions for stochastic optimization," *Operations Research*, vol. 55, no. 4, pp. 98–112, 2007.

[153] PRASTACOS, G., "Blood inventory management: An overview of theory and practice," *Management science*, vol. 30, no. 7, pp. 777–800, 1984.

[154] PRÉKOPA, A., *Stochastic Programming*. Kluwer Academic Publisher, 1995.

[155] PRIETO-RUMEAU, T. and HERNÁNDEZ-LERMA, O., *Selected topics on continuous-time controlled Markov chains and Markov games*. London: Imperial College Press, 2012.

[156] PUTERMAN, M., *Markov decision processes: discrete stochastic dynamic programming*, vol. 414 of *Wiley Series in Probability and Statistics*. John Wiley and Sons, 2009.

[157] RAO, U., SCHELLER-WOLF, A., and TAYUR, S., "Development of a rapidresponse supply chain at caterpillar," *Operations Research*, vol. 48, no. 2, pp. 189–204, 2000.

[158] REIMAN, M., "A new and simple policy for a continuous review lost-sales inventory model." Unpublished manuscript, 2004.

[159] RIEDEL, F., "Dynamic coherent risk measures," *Stochastic Processes and Their Applications*, vol. 112, pp. 185–200, 2004.

[160] RIEDEL, F., "Optimal stopping with multiple priors," *Econometrica*, vol. 77, no. 3, pp. 857–908, 2009.

[161] ROORDA, B. and SCHUMACHER, J., "Time consistency conditions for acceptability measures, with an applications to tail value at risk," *Insurance: Mathematics and Economics*, vol. 40, pp. 209–230, 2007.

[162] ROSENSHINE, M. and OBEE, D., "Analysis of a standing order inventory system with emergency orders," *Operations Research*, vol. 24, no. 6, pp. 1143–1155, 1976.

[163] ROSSI, R., RIJPKEMA, W. A., and VAN DER VORST, J. G. A. J., "The impact of dual sourcing on food supply chain networks: the case of egyptian strawberries," in *11th Wageningen International Conference on Chain and Network Management*, (Wageningen), 2012.

[164] ROYDEN, H. and FITZPATRICK, P., *Real analysis*. New York: Prentice Hall, 1988.

[165] RUSZCZYŃSKI, A., "Risk-averse dynamic programming for markov decision processes," *Mathematical Programming, Series B*, vol. 125, pp. 235–261, 2010.

[166] RUSZCZYŃSKI, A. and SHAPIRO, A., "Conditional risk mappings," *Mathematics of Operations Research*, vol. 31, pp. 544–61, 2006.

[167] RYAN, J., *Analysis of inventory models with limited demand information.* PhD thesis, Northwestern University, 1997.

[168] SCARF, H., "A min-max solution of an inventory problem," in *Studies in the Mathematical Theory of Inventory and Production* (ARROW, K., KARLIN, S., and SCARF, H., eds.), pp. 201–209, Stanford University Press, 1958.

[169] SCARF, H., "Bayes solution of the statistical inventory problem," *The annals of mathematical statistics*, p. 490C508, 1959.

[170] SCARF, H., "The optimality of (s, s) policies in the dynamic inventory problem," in *Mathematical Methods in the Social Sciences*, pp. 196–202, Stanford University Press, 1960.

[171] SCARF, H., "Some remarks on bayes solutions to the inventory problem," vol. 7, no. 4, pp. 591–596, 1960.

[172] SCHÄL, M., "Average optimality in dynamic programming with general state space," *Mathematics of Operations Research*, vol. 18, no. 1, pp. 163–172, 1993.

[173] SCHELLER-WOLF, A., VEERARAGHAVAN, S., and VAN HOUTUM, G.-J., "Effective dual sourcing with a single index policy." working paper, 2008.

[174] SEE, C. T. and SIM, M., "Robust approximation to multiperiod inventory management," *Operations Research*, vol. 58, no. 3, pp. 583–594, 2010.

[175] SENNOTT, L., "Average cost optimal stationary policies in infinite state markov decision processes with unbounded costs," *Operations Research*, vol. 37, no. 4, pp. 626–633, 1989.

[176] SETHI, S. P., YAN, H., and ZHANG, H., "Inventory models with fixed costs, forecast updates, and two delivery modes," *Operations Research*, vol. 51, no. 2, pp. 321–328, 2003.

[177] SHAPIRO, A., "On a time consistency concept in risk averse multistage stochastic programming," *Operations Research Letters*, vol. 37, pp. 143–147, 2009.

[178] SHAPIRO, A., "A dynamic programming approach to adjustable robust optimization," *Operations Research Letters*, vol. 39, pp. 83–87, 2011.

[179] SHAPIRO, A., *Topics in Stochastic Programming.* CORE Lecture Series, Universite Catholique de Louvain, 2011.

[180] SHAPIRO, A., "Minimax and risk averse multistage stochastic programming," *European Journal of Operational Research*, vol. 219, pp. 719–726, 2012.

[181] SHAPIRO, A., "Rectangular sets of probability measures." `http://www.optimization-online.org/DB_FILE/2014/10/4613.pdf`, 2014.

[182] SHAPIRO, A., DENTCHEVA, D., and RUSZCZYŃSKI, A., *Lectures on Stochastic Programming: Modeling and Theory*. Philadelphia: SIAM, 2009.

[183] SHEOPURI, A., JANAKIRAMAN, G., and SESHADRI, S., "New policies for the stochastic inventory control problem with two supply sources," *Operations Research*, vol. 58, no. 3, pp. 734–745, 2010.

[184] SONG, J. and ZIPKIN, P., "Inventory control in a fluctuating demand environment," *Operations Research*, vol. 41, no. 2, pp. 351–370, 1993.

[185] SONG, J. and ZIPKIN, P., "Inventories with multiple supply sources and networks of queues with over ow bypasses," *Management Science*, vol. 55, no. 3, pp. 362–372, 2009.

[186] SPITZER, F., "A combinatorial lemma and its application to probability theory," *Transactions of the American Mathematical Society*, vol. 82, no. 2, pp. 323–339, 1956.

[187] STROTZ, R., "Myopia and inconsistency in dynamic utility maximization," *Review of Economic Studies*, vol. 23, no. 3, pp. 165–180, 1955.

[188] TAGARAS, G. and VLACHOS, D., "A periodic review inventory system with emergency replenishments," *Management Science*, vol. 47, no. 3, pp. 415–429, 2001.

[189] TOKTAY, L. and WEIN, L., "Analysis of a forecasting-production inventory system with stationary demand," *Management Science*, vol. 47, no. 9, pp. 1268–1281, 2001.

[190] TROJANOWSKA, M. and KORT, P., "The worst case for real options," *Journal of Optimization Theory and Applications*, vol. 146, no. 3, pp. 709–734, 2010.

[191] VEERARAGHAVAN, S. and SCHELLER-WOLF, A., "Now or later: A simple policy for effective dual sourcing in capacitated systems," *Operations Research*, vol. 56, no. 4, pp. 850–864, 2008.

[192] VEINOTT, A., "Optimal policy for a multi-product, dynamic, nonstationary inventory problem," *Management Science*, vol. 12, no. 3, pp. 206–222, 1965.

[193] VEINOTT, A. F., "On the optimality of (s, s) inventory policies: new conditions and a new proof," *SIAM Journal on Applied Mathematics*, vol. 14, pp. 1067–1083, 1966.

[194] VERHOEF, P. and SLOOT, L., "Out-of-stock: reactions, antecedents, management solutions, and a future perspective," in *Retailing in the 21st Century*, pp. 239–253, Springer, 2006.

[195] ŽÁČKOVÁ, J., "On minimax solution of stochastic linear programming problems," *Časopis Pěstování Matematiky*, vol. 91, pp. 423–430, 1966.

[196] WANG, T., *A class of dynamic risk measure.* University of British Columbia, 1999. Working Paper.

[197] WANG, T., ATASU, A., and KURTULUS, M., "A multi-ordering newsvendor model with dynamic forecast evolution," *Manufacturing & Service Operations Management*, vol. 14, no. 3, pp. 472–484, 2012.

[198] WHITT, W., *Stochastic-process limits: an introduction to stochastic-process limits and their application to queues.* Springer, 2002.

[199] WHITTMORE, A. S. and SAUNDERS, S. C., "Optimal inventory under stochastic demand with two supply options," *SIAM Journal on Applied Mathematics*, vol. 32, pp. 293–305, 1977.

[200] WIESEMANN, W., KUHN, D., and RUSTEM, B., "Robust markov decision processes," *Mathematics of Operations Research*, vol. 38, no. 1, pp. 153–183, 2013.

[201] XIN, L. and GOLDBERG, D., "Distributionally robust inventory control when demand is a martingale," 2014. working paper.

[202] XIN, L. and GOLDBERG, D., "Optimality gap of constant-order policies decays exponentially in the lead time for lost sales models," 2014. arXiv preprint arXiv:1409.1499.

[203] XIN, L. and GOLDBERG, D., "Asymptotic optimality of tailored base-surge policies in dual-sourcing inventory systems," 2015. arXiv preprint arXiv:1503.01071.

[204] XIN, L., GOLDBERG, D., and SHAPIRO, A., "Time (in)consistency of multistage distributionally robust inventory models with moment constraints," 2013. arXiv preprint arXiv:1304.3074.

[205] YANG, J., "Inventory and price control under time-consistent coherent and markov risk measure," *preprint*, 2012.

[206] YASPAN, A., "An inclusive solution to the inventory problem," *Operations Research*, vol. 9, pp. 371–382, 1961.

[207] YI, J. and SCHELLER-WOLF, A., "Dual sourcing from a regular supplier and a spot market." Working paper, Carnegie Mellon University, Pittsburgh, 2003.

[208] Yue, J., Chen, B., and Wang, M., "Expected value of distribution information for the newsvendor problem," *Operations Research*, vol. 54, no. 6, pp. 1128–1136, 2006.

[209] Zhang, V. L., "Ordering policies for an inventory system with three supply modes," *Naval Research Logistics*, vol. 43, pp. 691–708, 1996.

[210] Zhou, S. and Chao, X., "Dynamic pricing and inventory management with regular and expedited supplies," *Production and Operations Management*, vol. 23, no. 1, pp. 65–80, 2014.

[211] Zhou, S., Tao, Z., and Chao, X., "Optimal control of inventory systems with multiple types of remanufacturable products," *Manufacturing & Service Operations Management*, vol. 13, no. 1, pp. 20–34, 2011.

[212] Zhu, Z., Zhang, J., and Ye, Y., "Newsvendor optimization with limited distribution information," *Optimization methods and software*, vol. 28, no. 3, pp. 640–667, 2013.

[213] Zipkin, P., *Fundamentals of inventory management.* New York: McGraw-Hill, 2000.

[214] Zipkin, P., "Old and new methods for lost-sales inventory systems," *Operations Research*, vol. 56, no. 5, pp. 1256–1263, 2008.

[215] Zipkin, P., "On the structure of lost-sales inventory models," *Operations Research*, vol. 56, no. 4, pp. 937–944, 2008.